World Wide Web?
A closer look at the transnational online
public discourse on climate change

Inaugural dissertation submitted by Ueli Reber in fulfillment of the
requirements for the degree of Doctor rerum socialium at the Faculty of
Business, Economics and Social Sciences of the University of Bern

Submitted by

Ueli Reber

from Wimmis and Diemtigen

July 14, 2020

The faculty accepted this work as dissertation on September 17, 2020 at the request of the two advisors Prof. Dr. Silke Adam (University of Bern) and Prof. Dr. Annie Waldherr (WWU Münster), without wishing to take a position on the view presented therein.

Acknowledgments

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In 2009, the United Nations Framework Convention on Climate Change (UNFCCC) stated in the Copenhagen Accord that "climate change is one of the greatest challenges of our time" (UNFCCC, 2009, p. 1). It is such a challenge because a massive and coordinated effort is needed to "stabilize greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system" (UNFCCC, 1992, Art. 2). To avoid this, the UNFCCC aims for a temperature rise of less than 2°C relative to pre-industrial levels. However, achieving this goal has become increasingly difficult. The physics of the climate system are complex as is the implementation of effective solutions to tackle the growing number of social, political, and economic problems that result from climate change. This is due to what is known as the tragedy of the commons—a situation in which individual short-term interests are at odds with the long-term interests of the collective (Hardin, 1968). Because it may cause financial losses, it is unattractive for individuals (and states) to cut back on greenhouse gas emissions to limit climate change unless others do so. In addition, climate change is an abstract phenomenon whose causes and effects can only be connected with the help of scientific methods, not personal experience. All of this makes it a highly contested issue.

While responsible and constructive actors advocate—or at least accept—effective climate policies to reach the 2°C goal, obstructive actors try to block all legislation that could possibly entail costs and restrictions for them. In both camps there are actors as varied as governments, parliaments, politicians, political parties, scientists, non-governmental organizations, companies, media outlets, social movements, and individual citizens. It goes without saying that these actors have very different possibilities when it comes to asserting their interests. However, most of them need the support of the public at least sometimes (Kriesi, 2004). Therefore they must inform, convince, and mobilize the public. To this end, they have to communicate their views, arguments, critiques, endorsements, etc. in public statements. The resulting set of statements is what is usually considered *public discourse.*

Public discourses are often seen as national matters in political communication research as they produce public opinion and thus serve as intermediaries between the (national) public and the (national) government (cf. Gerhards & Neidhardt, 1990; Peters, 1994). However, the battle over climate change policies is fought not only at the national level but also in and across political arenas at all levels (Simonis, 2017). This multi-level system of climate governance generates a lot of public communication both at and across the different levels, given that actors need to be active in multiple arenas simultaneously to defend their interests effectively and that they (at least sometimes) require the support of the public to do so.

In addition to this practical aspect, which is important for the actors involved in climate politics, a transnational climate discourse is also desirable for more normative reasons. First, the burdens and responsibilities of climate change are unevenly distributed across the globe (IPCC, 2014). To become aware of this and possibly work out a coordinated response, transnational communication is essential. Second, exclusively national discourses are neither normatively legitimate nor politically efficient when it comes to global issues like climate change (Fraser, 2007; Habermas, 1998). This means that if national discourses are strictly separate, not all people affected by the changing climate can participate in the formation of public opinion. It also means that national institutions—which would be the only ones that could be addressed by an isolated national public—would not be able to respond effectively as the problem goes beyond their competencies. In order to be both normatively legitimate and politically effective, public discourses on global issues like climate change must be transnational. For this reason it is crucial that public communication about climate change is not only studied within the "bounded 'container[s]' of the Westphalian nation state" (Volkmer, 2019, p. 12) but also beyond.

Based on these considerations, I pursue three research objectives in this thesis: (1) to map how transnational the online public discourse on the global phenomenon of climate change is, (2) to understand the role of the (trans)nationalized online public discourses on climate change in today's hybrid media system, and (3) to find, implement, and validate computational methods to study public discourses across different political and language spaces. In the following sections of this introduction, I show why these three objectives deserve our attention and how studying them contributes to the existing literature. I also show how each of the three articles included in the thesis help to achieve the objectives and summarize their findings. The final part of the introduction is devoted to the limitations of this thesis. Building on that, I conclude by discussing possible future research questions.

1.1 Transnationality of the online public discourse on climate change

The web has become an essential platform for many actors when it comes to public communication. This is because websites, blogs, and social media allow everyone the unfiltered distribution of

self-made content. In contrast to traditional mass media, actors can speak for themselves on the web. This does not mean, however, that actors necessarily get more attention when they communicate online (Koopmans & Zimmermann, 2010), but it does mean that the web provides them with a platform to participate in public discourse with relative ease. In addition, the web does not limit communication to a certain territory. This is one aspect that differentiates the web from traditional mass media, where the scope of a medium is often technically or logistically limited to a certain territory (e.g., to the area where you can listen to a radio program or subscribe to a newspaper). In this regard, the web is global. There are some exceptions to this claim, but in most countries, web content can be produced for any (foreign) audience and the audience can access any (foreign) content. In other words, there are no technical barriers on the web that would limit public discourses to national territories and publics. All (trans)transnational communication flows are thus the result of decisions and actions of the actors communicating online. This provides insight not only on how climate change is constructed as a public issue, but also where boundaries are drawn in the collective imagination (Halvais, 2000).

However, little is known about the transnationality of public discourses on the web—in particular climate change. This thesis seeks to reduce this research deficit by analyzing two mechanisms of discourse transnationality: *convergence* and *integration*. Convergence means the harmonization of public discourses on a certain issue in different countries over time. Complete convergence is reached when the same issue is debated by different (national) publics simultaneously using the same arguments (Eder & Kantner, 2000). This is the focus of the second study included in this thesis (chapter 3). In "Overcoming language barriers: Assessing the potential of machine translation and topic modeling for the comparative analysis of multilingual text corpora," I compare how actors from Germany, the United Kingdom, and the United States discuss climate change on their websites and assess whether the online discourses on climate change in these countries converge.

In the first article (chapter 2), "Global climate change or national climate changes? An analysis of the performance of online issue publics in integrating global issues," the focus is on transnational discourse integration. Integration means that foreign actors and their concerns are incorporated in a discourse of a (national) public (Gerhards, 1993; Koopmans & Erbe, 2004). Traditional mass media achieve this by giving foreign actors and their positions a platform in their coverage. The same can be done by all actors when it comes to content on their websites; there, actors can reference other actors, places, and issues. In addition, the web also provides the opportunity to reference actors and their web content via hyperlinks. As hyperlinks directly affect the visibility of content on the web (Park, 2003), this form of integration has a unique quality: it changes the structure of the online public sphere (Koopmans & Zimmermann, 2010). In the first article I therefore examine both link types—hyperlinks and referential links—side by side. Based on these links, I asses the degree and scope of transnational integration for climate change discourses on the web in Germany, Switzerland, the United Kingdom, and the United States.

The findings of the two studies show that the public discourse on climate change is transnationalized to a considerable degree. First, the same topics define the issue in the countries studied. However, some of the topics are of different importance to the actors in these countries. In the United States, the debate revolves largely around the question of whether climate change exists, while in European countries the discussion of possible coping strategies is prevalent. The discourses are therefore largely convergent in terms of the spectrum of what is discussed, but not in terms of the importance that the individual topics have for the actors in the different countries. Second, the discourses in the countries are shaped by both domestic and foreign actors. In the European countries, transnational links are even more frequent than domestic links, which indicates a high level of awareness for the transnational dimension of climate change. However, the scope of transnationalization is restricted to countries of the Global North, with a clear bias towards the United States. In the studied discourses of the Global North, the Global South is thus a blind spot. Communication on the web therefore blurs some boundaries but at the same time deepens others.

1.2 Role of online public discourse on climate change in the hybrid media system

The emergence of the web was a decisive moment in the evolution of the mediated public sphere as it marked not only the transition from a nationally bounded media system but also the transition from a linear and relatively exclusive to a multi-directional and inclusive regime of media discourse production. What we know about the world today we do not necessarily know from the mass media (cf. Luhmann, 1995). We might know it from Facebook, YouTube, Wikipedia, Telegram, a blog, a mailing list, or another online service. Alongside mass media, such online services have become important venues of public communication. Governments, businesses, organizations, groups, individuals, and even the mass media use them to put their messages forward to the public. As a consequence, the mediated public sphere has become more diverse, more inclusive, but also more complex and harder to understand for both citizens and researchers. Chadwick (2013) refers to this decentralization process as a hybridization of the media system.

Although the mass media have lost their unique gatekeeper position in the hybrid media system, their role in public discourse is still crucial: as aggregators of opinions and positions they offer counter-publics a possible gateway into mainstream discourse. Counter-publics are groups of actors who oppose the dominant public opinion and are therefore marginalized in mainstream discourse (Fraser, 1990). For such groups, the web is an ideal space to discuss and determine their position, define their identity and shape their agenda. However, if they want to gain social and political influence, they must also succeed in communicating their messages to people outside their community. This makes mass media coverage a valuable showcase to these groups. Despite this

critical function of mass media for (partly anti-democratic) counter-public spheres, surprisingly little is known about the resonance of their online communication in legacy media.

This is all the more important as there are no borders for counter-publics on the web. They can organize themselves transnationally and thus benefit from each other's experience and resources in their public communication at all political levels. In the case of climate change, this seems to be particularly beneficial for the skeptics in Europe. They deny or talk down man-made climate change and reject any measures to prevent it. In most European countries these positions make them a counter-movement (Dunlap & McCright, 2015). The web, however, allows them to connect with actors in the United States, where skeptics are well funded and part of the mainstream (Brulle, 2014; Dunlap, McCright, & Yarosh, 2016). A transnational network thus benefits the skeptics in Europe in terms of professionally produced content (e.g., movies) and reputation. Whether this helps them to get into the European mass media and thus into mainstream discourse is examined in the third article included in this thesis (chapter 4). In "How climate change skeptics (try to) spread their ideas: Using computational methods to assess the resonance among skeptics' and legacy media" we examine the discursive resonance of online climate change skepticism in traditional media in Germany. Using different computational methods and qualitative case studies, we study whether the transnational climate change skeptics' counter-movement succeeds in using online communication to trigger either continuous or selective resonance in German mass media.

We found no evidence for continuous resonance. However, there were occasions of selective resonance when climate change skeptics manage to exploit specific events to push their perspectives and positions onto the mass media's agenda. Other than in the United States, German skeptics have no allies among the conservative media. Although conservative media play down climate change by giving less space to the issue in their reporting, they do not provide skeptics with a platform in the mainstream discourse. The influence of the skeptical counter-movement on German mainstream discourse is therefore limited. Nevertheless, the transnational network represents a valuable resource from which skeptical actors in all countries can benefit when a window of opportunity opens.

1.3 Methods to study transnational and multilingual public discourses on the web

Despite the transnational reality of online public discourses and the growing importance of the international level in many policy fields, political communication research still tends to focus on the national level alone. This may be explained by the fact that most popular theories, like the Habermasian notion of the public sphere (Habermas, 1962), are based on the premise of a culturally homogeneous and politically sovereign nation state. Yet another factor is certainly that studies involving several countries have long been difficult and expensive (Livingstone, 2003).

This is especially true for studies that deal with communication in more than one language. Even today, when communication researchers have many computational methods at their disposal, most off-the-shelf tools are not suitable for multilingual analyses. However, the drastic increase in computational capacity, the availability of training data, and improved access to software have the potential to change things for the better.

Some efforts have been made in recent years to make topic models—an indispensable tool in the toolbox of every computational communication researcher (Boumans & Trilling, 2016)—suitable for the analysis of multilingual corpora (e.g., Lucas et al., 2015). With the second article of my dissertation (chapter 3), I contribute to these efforts. In "Overcoming language barriers: Assessing the potential of machine translation and topic modeling for the comparative analysis of multilingual text corpora" I show how topic models can be combined with machine translation to support comparative research across countries and languages. In the article, I assess the robustness of different translation methods and test the integrability of the approach into a comparative study design using a real world example: the analysis of transnational discourse convergence in the case of climate change. The results show that the combination of machine translation and topic models is a great option when it comes to the automated analysis of large multilingual corpora. Regardless of whether full texts or only the vocabulary of a corpus is translated, the approach produces reliable and robust results. Moreover, the analysis of transnational discourse convergence (see above) has shown that machine translation and topic models can also be used for comparative research.

This conclusion is supported by the third article included in this thesis (chapter 4). In this article, we used machine translation and topic models to assess the thematic resonance among climate change skeptics' online communication and the reporting of German legacy media and have thereby shown that the approach can easily be integrating into a more complex analysis workflow. Overall, these studies provide the proof of concept that tools for the analysis of multilingual communication are available. So we should start using them to study communication in transnational and multilingual settings more often.

1.4   Limitations and future research

This dissertation sheds light on the geography of online public discourse on climate change, determines its resonance in legacy media, and proposes a methodological approach for its analysis across countries and languages. But it also has its blind spots and raises new questions. The first such question that must be highlighted is that of the motivation and communication strategy of the actors. Since the analyses in the articles are primarily aimed at identifying (trans)national communication patterns, little can be said about what drives the actors to produce these patterns. To gain a better understanding of their reasoning, either the computation of inferential network models (e.g., ERGMs) or a more qualitative approach is necessary. This could include a close

linguistic analysis of the websites or interviews with the actors involved. Both approaches would be helpful to improve our understanding of how public discourses work within the networked public sphere.

The second question relates to the diffusion of ideas on the web. The articles in this thesis are based on the assumption that transnational communication networks foster the transnational flow of knowledge, experiences, and ideas. The analysis of transnational discourse integration even provides evidence for this (see chapter 2). As diffusion is a process, however, it would also be worth looking at transnational information flows over time and in more detail (e.g., individual arguments). This could provide valuable insights into the communication dynamics of transnational counter-movements such as climate change skeptics.

Thinking beyond the web, it must be questioned how relevant the assumed linear flow of information from the web into mass media and eventually into politics in today's hybrid media system still is. With their online communication, counter-movements may also succeed in getting their ideas onto the public agenda and into politics without the support of mass media. To identify the critical processes and mechanisms of the audience democracy (Kriesi, 2004) in the digital age, more research on the mutual interaction of the various arenas is needed.

The final question relates to the struggle for effective climate policies: Does a transnational discourse actually help to meet the 2℃ goal? As mentioned above, a transnationalization of the discourse might be desirable from a theoretical point of view. In practice, however, transnationalization may first and foremost benefit climate change skeptics (as discussed in chapter 4). Thus, we should not only continue our efforts to make skeptics' practices transparent but also spend more time thinking about how to handle the divergence of theoretical and political ideals. This would advance both theoretical and empirical communication research.

Chapter 2

Global climate change or national climate changes? An analysis of the performance of online issue publics in integrating global issues

Abstract: This paper analyzes how the complex global spatiality of climate change is integrated into online national public discourses. Although the Web is an important venue for public discourses, little is known about its capability to integrate transnational issues. By looking at two types of communicative links (hyperlinks and referential links), we assess the degree and the scope of transnational integration for the four cases of Germany, Switzerland, the United Kingdom, and the United States. The findings show that these national discourses are heavily transnationalized. However, the scope of transnationalization is restricted to countries of the Global North, with a clear focus on the United States. This leads to the conclusion that the Web's capability to integrate transnational issues is limited.

2.1 Introduction

Climate change is experienced and caused locally, but the climate system is global (IPCC, 2014). This is why many of those most affected by the changing climate have contributed the least to the greenhouse gases causing the problem (IPCC, 2014). Because of this complexity, climate politics are local, national, regional, and global. In other words, political arenas on different levels are closely intertwined in a multilevel system of climate governance (Di Gregorio et al., 2019; Jänicke, 2017; Simonis, 2017). However, the national arena plays a particularly important role in this system, as national governments are responsible for both negotiating and implementing international treaties, such as the Paris Agreement (UNFCCC, 2015). Since national governments are legitimized by the citizens of their country (i.e., the national public), it is crucial that there are national public discourses on climate change legislation, domestic greenhouse gas emissions, and the local impacts of the changing climate. But it is just as crucial that there is a transnational public discourse on international conferences and treaties, legislation in foreign countries, and the effects of climate change in other parts of the world. After all, climate change is a global problem that requires global awareness and a coordinated effort to solve it. From this point of view, a public discourse in which both national and transnational facets of the issue are simultaneously presented, linked and discussed—that is integrated—is not only desirable, but necessary.

But how well does this integration of both national and transnational aspects into the public discourse on climate change work? Does the complexity of the issue translate into a public discourse that is national, global, or integrated? Looking at public discourses in legacy media, most previous studies found what is best described as *transnationalized national discourses*, meaning that public discourses are first and foremost national because of national media systems but are transnationalized through the frequent reporting about what is happening in other countries and on the international level (Olausson, 2009; Schäfer, Post, Schwab, & Kleinen-von Königslöw, 2018; Wessler, Wozniak, Hofer, & Lück, 2016). It is no surprise that the United Nations Framework Convention on Climate Change (UNFCCC) conferences are major catalysts for the transnationalization of national discourses (Eide & Kunelius, 2010). However, it is not only international politics that trigger transnationalization but also exceptional events (e.g., storms, droughts, rallies, elections) and prominent actors (e.g., politicians, celebrities) in other countries (Schäfer et al., 2018).

This, however, does not necessarily hold true for the public discourse on the Web. Bennett, Lang, and Segerberg (2015) found highly fragmented national discourses when they analyzed the online debate on climate change in Germany and the United Kingdom based on (trans)national hyperlink networks. The authors found that national and transnational actors constitute separate publics around the climate issue, each with a specific political arena in focus (i.e., a country or a transnational body like the European Union). Because these issue publics are only weakly connected, the domestic discourses remain mostly domestic.

By increasing the "spatial scope" (Koopmans & Erbe, 2004, p. 107) of the national discourses, legacy media thus seem to integrate the complex spatiality of the climate change issue better than the seemingly borderless Web. In this study, we examine whether there is empirical evidence to support this statement by analyzing and comparing the integrative performance of the national discourses about climate change on the Web in Germany, Switzerland, the United Kingdom, and the Unites Sates. To do so, we use the actor-centered *issue mapping* approach (Marres, 2015). This approach allows us to determine both degree and scope of transnational integration of the *issue networks* that represent these countries' online public sphere devoted to climate change (Marres & Rogers, 2005). Specifically, we address two research questions:

- To what degree are the climate change issue networks in Germany, Switzerland, the United Kingdom, and the United States transnationalized?

- What is the scope of transnationalization of these countries' climate change issue networks?

Studying and comparing four cases allows us to identify both common and country-specific patterns of transnationalization (Esser & Hanitzsch, 2012). In fact, such a comparative design is the only way to study the geography of the public discourse on climate change beyond the "bounded 'container[s]' of the Westphalian nation state" (Volkmer, 2019, p. 251), without ignoring them as the main political arenas in climate governance. As a result, this study contributes to a better understanding of the public discourse on climate change at both the national and transnational level.

Moreover, our analysis is not only based on hyperlinks—the skeleton of issue networks—but also considers referential links. Referential links are references to objects, persons, places, etc. made in the text of a web page. Just like hyperlinks, they represent communicative flows and therefore have a spatial dimension (Koopmans & Zimmermann, 2010). Although the analysis of referential links is standard when studying the transnationality of discourses in legacy media, they are mostly ignored in Web studies. Especially studies that use the issue mapping approach usually focus on hyperlinks. However, it is unclear whether and how hyperlinks and referential links differ when it comes to transnational communicative flows. We address this gap by examining both types of links side by side. By doing so, we contribute to the extension of the empirical scope of controversy analysis in online settings (cf. Marres, 2015).

2.2 Theoretical Considerations

2.2.1 Issue Networks

The Web is an exceptional platform for public discourse, as there are no technical barriers that would restrict public discourses to a certain space. With a simple website, everyone with internet

access can get involved in any public discourse anywhere in the world at any time. This makes the Web the cornerstone of the *networked public sphere*—a decentralized public sphere that is facilitated by digital communication technologies (Benkler, 2006; Friedland & Rojas, 2006). It builds on multi-directional communicative links among a potentially infinite number of actors. This notion fits the architecture of the Web perfectly, as the Web not only allows actors to easily become their own publishers but also to connect with others via hyperlinks (Park, 2003).

Hyperlinks are the essential structural element of the Web. It directly depends on them, whether webpages are visible to the audience (i.e., the users) or not (González-Bailón, 2009). This is either because users move from one webpage to another via hyperlinks or because the algorithms of search engines, such as Google (Brin & Page, 1998), deduce the importance of a webpage from the hyperlinks it receives (Koopmans & Zimmermann, 2010). Search engines thereby take advantage of the fact that actors with similar interests connect in *issue networks* (Marres & Rogers, 2005).

Issue networks are networks of public communication that form around a common and often contested topic (issue). The network consists of webpages and the hyperlinks connecting them. However, issues do not exist a priori but are created through the different views expressed by the actors in the network (Marres, 2006). The issue, therefore, constitutes the network, and the network, conversely, constitutes the issue. Because issue networks are open to connection and contestation among all interested actors (including citizens), the networks can be understood as full-fledged *issue publics* (Bennett et al., 2015).

Issue networks are thus forums in which public discourses take place (e.g. on climate change). Because they typically consist of actors from different fields (e.g., politics, civil society, science, economy), they are not only a site of contention but also of public opinion formation (Marres, 2006). Because of their unrestricted accessibility, issue networks have the potential to serve as a public source of knowledge and inspiration, as originator of public criticism and support for political actions, as well as indicators of cultural and political change. These processes are not limited to the issue network. In the network public sphere, the views and opinions formed in issue networks may spill over into mass media and politics and thus become eventually socially and politically effective (Pfetsch, Adam, & Bennett, 2013). In order to gain a more comprehensive understanding of social and political developments, it is therefore essential to study the public communication of issue networks.

2.2.2 Degree and Scope of Transnationality

Issue networks are not geography-free but tied to a political space. Otherwise, they could hardly become politically efficient (Fraser, 2007). The spatiality of issue publics is created through the association of its actors with a particular political space. Actors can communicate their scope—that is their sphere of reference—via their website, for example by choosing a country-specific domain ending (e.g., ".us" or ".de"). A national issue network therefore emerges when actors from the

same country connect. By focusing on domestic affairs and actors, such networks produce an issue that is primarily relevant to a national public. We can thus speak of a *national issue public*. By integrating foreign actors (i.e., by linking to them), the network becomes transnationalized. This extends the scope of the issue network, as it now also includes the views of foreign actors. What is debated in the network thus becomes meaningful for a transnational public, which is why we can speak of a *transnational issue public* in this case.

However, to what degree an issue public is transnationalized is relative. Based on the notion of Deutsch (1953), the degree of transnationalization is typically defined as the ratio of communicative interactions within a country to communicative interactions beyond the country (Gerhards & Rössel, 1999; Koopmans & Erbe, 2004). When applied to issue networks, this means that the degree of transnationalization is determined by the ratio of hyperlinks between actors with the same scope (i.e., "their" country) to hyperlinks between actors with different scopes.

Due to the important role of the national arena in climate politics (Simonis, 2017), it is reasonable to proceed from national issue publics also when studying discourses on the technically boundless Web (Cammaerts & Van Audenhove, 2005; Segerberg, 2012). The starting points in this study are therefore national issue networks. In a first step, the networks thus consist of domestic actors (i.e., actors from the same country), but are then "open to connection and contestation" (Bennett et al., 2015, p. 112) among actors from every other political sphere. This allows us to check whether the spatial reach and the boundaries of public communication correspond with national borders or go beyond them (the degree of transnationalization). It also allows us to identify exactly which other spheres are integrated into the national publics (the scope of transnationalization).

2.2.3   Communicative Links: Hyperlinks and Referential Links

Analyzing hyperlinks is a popular approach among scholars who are interested in the transnationality of the Web. Many studies on international hyperlink networks (Barnett, Chung, & Park, 2011; Barnett & Park, 2005; Halvais, 2000; Janc, 2015; Park, Barnett, & Chung, 2011) and "international news flows" (Chang, Himelboim, & Dong, 2009; Himelboim, 2010; Segev, 2010; Segev & Hills, 2014) have measured cross-border communication based on hyperlinks. Studies with a more specific focus on the transnationality of issue publics have likewise focused on hyperlinks in order to measure the issue networks' transnationality (Bennett et al., 2015; Hepp et al., 2016; Koopmans & Zimmermann, 2010; Rogers & Ben-David, 2008; Shumate & Dewitt, 2008).

However, hyperlinks are only one specific form of communicative link. Another, much more common form are links established through references in the text. That is by mentioning an object, a person, or any other entity that could be named in the text of a webpage (i.e., without setting a hyperlink). We call these in-text mentions *referential links*.

Although hyperlinks are the main structural element of the Web and the skeleton of issue networks, referential links should not be ignored when analyzing public discourses online. Mainly

because not all places and actors can be addressed via hyperlinks (e.g., countries, groups of people), and hyperlinks are sometimes avoided (e.g., to avoid that a contradicting view becomes visible). Although the analysis of referential links is not unusual in legacy media research (e.g., Adam, 2007; Koopmans & Statham, 2010; Schäfer et al., 2018; Wessler, Peters, Brüggemann, Kleinen-von Königslöw, & Sifft, 2008b), it has rarely been used to analyze communication on the Web. Exceptions are Koopmans and Zimmermann (2010), who manually coded statements from actors ("claims") on webpages and Segev (2010), as well as Segev and Hills (2014), who automatically extracted country names from online news sites to measure transnationality in public communication online.

Referential links and hyperlinks are thus two kinds of the same phenomenon: *communicative links*. In order to obtain the full picture, both types should be analyzed simultaneously. This is why we assess and compare degree and scope of transnationalization for both hyperlinks and referential links.

2.3 Data and Methods

2.3.1 Case Selection

We concentrate on the national discourses in Germany, Switzerland, the United Kingdom (UK), and the United States (US). Although these are all Western democracies and signees of the UNFCCC (1992), the situation for the domestic actors involved in climate politics are different. While Germany, Switzerland and the UK are making some political efforts to meet the reduction targets defined in the Paris Agreement, this is not the case for the US (Burck, Hagen, Marten, Höhne, & Bals, 2019; Iacobuta, Dubash, Upadhyaya, Deribe, & Höhne, 2018). Climate change is an issue that deeply divides the American society (Dunlap et al., 2016), but less so European societies (McCright, Dunlap, & Marquart-Pyatt, 2015). Moreover, the US is considerably larger than the other three countries and therefore fields far more scientists, policy makers, and other actors with an interest in climate change. Compared to the other countries, the US is also home to more (hyper-partisan) online media and blogs that produce Web content for a domestic public (Heft, Mayerhöffer, Reinhardt, & Knüpfer, 2019). Consequently, the possibilities for political influence and alliances are different in the US and European countries. This is likely to affect the communication behavior of the actors constituting the domestic issue networks (Häussler, Adam, Schmid-Petri, & Reber, 2017). Apart from that, the UK[1] and Germany are members of the European Union, where climate change is a major policy field (Jordan, van Asselt, Berkhout, Huitema, & Rayner, 2012), while Switzerland and the US are not. British and German actors must therefore take a additional political sphere into account, when they decide on their communication strategy. However, the European sphere also offers them additional opportunities to publicly promote their interests (Bennett et al., 2015). Finally, the national languages differ: German is

---

[1] At the time of data collection.

spoken in Germany and Switzerland (in some parts), English in the UK and US. Compared to the "English" Web, the "German" Web is small, which means that the German-speaking actors have few possibilities to link to actors of the same language. The following analysis aims to show whether these contextual factors affect the linking behavior of the actors in the climate change issue networks of these countries.

2.3.2   Generating Issue Networks

The four country-specific issue networks were gathered using a hyperlink-tracing procedure, which consisted of four main steps (Adam, Häussler, Schmid-Petri, & Reber, 2016). First, we identified the most prominent actors in each country based on literature reviews, expert validation,[2] and country-specific Google searches.[3] Prominent actors are known, credible and visible on the Web. The Google searches were used to determine the visibility of the actors. The literature reviews and the expert interviews then helped to determine whether an actor is known and credible. We chose actors who scored well in all three criteria. In this step, the focus was exclusively on national challenger actors. Following Kriesi (2004, p. 189), we understood challengers to be those actors with a (political) goal but no institutionalized access to formal political institutions or the media (e.g., non-governmental organizations, bloggers, universities, think tanks). They are the "champions of online climate communication" (Schäfer, 2012, p. 530), as they make extensive use of the Web to address the public (or parts of the public). However, their motivations and interests in public communication can differ greatly (cf. Stein, 2009). Therefore, also their national respectively transnational orientation potentially differs. In order to capture the whole spectrum of the debate, the four most prominent climate advocates, as well as the four most prominent climate skeptics, were chosen as starting points for the snowball sampling procedure (see list of actors in the Appendix, chapter 2.6).

This second step was done with the help of the crawler software Issue Crawler (Rogers, 2013). We set the crawler to start from the main page dealing with climate change on the seed actors' websites (the URLs are also listed in the Appendix, chapter 2.6), following all internal hyperlinks two levels deep into the respective website, and then collected all outgoing hyperlinks from these pages that were directed to external webpages, no matter whether they were challenger actors or not. The obtained issue networks are therefore not just challenger networks. In a final iteration, the crawler checked for hyperlinks running between the already indexed pages. Pretests have shown that more iterations primarily lead to the inclusion of US actors. The national networks of the European countries thus become quasi US networks. Moreover, an additional iteration produced a large number of webpages unrelated to climate change. Due to capacity constraints, it

---

[2]We asked researchers investigating climate change discourse and policy in the four countries to validate and, if necessary, modify the lists of actors obtained by literature reviews and Google searches.

[3]The English terms "climate change" and "global warming" as well as the corresponding German terms "Klimawandel" and "Globale Erwärmung" were used as search terms. These are often used and identified as appropriate search terms to capture the issue of climate change (Painter & Gavin, 2016; Shapiro & Park, 2015).

was not possible to identify, download and check all these webpages on a monthly basis. Therefore, no further crawling iterations were made.

In the third step, we downloaded all crawled webpages and then automatically indexed them according to our search terms in order to ensure that the climate change issue was actually addressed on these webpages. This is the crucial step in generating issue networks on the basis of hyperlink networks, as it removes a substantial amount of noise (Waldherr, Maier, Miltner, & Günther, 2017).

Finally, we aggregated the crawled webpages and manually coded them. The aggregation was done at the actor level. This means that webpages were not only aggregated on the basis of their domain but that domains of the same actors were also grouped together. Actors could be both individuals (e.g., bloggers) and organizations (e.g., NGOs, companies, governmental bodies). Then, the scope of all actors was manually coded using the information usually found on the "About us" page or a similar section of their website(s). The guiding principle was to take the perspectives of the actors and to determine whether they have a national focus (e.g., Switzerland, Germany) or a transnational one (e.g., EU, United Nations).[4] We have taken into account whether an actor is the national branch of a transnational organization. For example, we coded the national scope for Greenpeace Germany and the transnational scope for Greenpeace international.

Hyperlinks originating from foreign actors were eventually removed from the networks, as the focus of the study is on national issue publics. This is the strictest and most conservative definition of national issue publics, as only the linking behavior of the domestic actors is taken into account. This ensures that the country-specific degree and scope of transnationalization is captured, without any bias caused by the linking of foreign actors who are only indirectly involved (i.e., by being linked into the network) in the national climate change discourse.

Instead of using only one issue network per country, we used 12 (or 11)[5] and reported mean values for all measures. For each country, one issue network was generated at the beginning of each month between June 2012 and May 2013. This was done because issue networks typically have a relatively stable core, but always show a certain fluctuation at the margins. A comparison of snapshots taken at only one point in time can therefore lead to biased findings. To obtain more reliable measurements, we decided to smoothen the networks over a 12-month period. Table 2.1 shows the average number of actors, hyperlinks, and referential links of the issue networks. It also illustrates the variance at the fringes of the networks by reporting the standard deviations.

---

[4]The coding was done by two trained student assistants. The training process consisted of several rounds of coding an identical set of actors, followed by a comparison with the master coding and a discussion of the results. The intercoder reliability was assessed by comparing the coders' classification with the master coding. We used Krippendorff's alpha to measure the agreement between the coders' classification and the master coding as the coefficient is comparable across any number of coders, values, metrics, and sample sizes. For the scope variable (distinguishing 199 categories), Krippendorff's alpha was 0.93 and therefore satisfactory (1 means complete agreement). Throughout data collection, critical cases were discussed and coded in coders' meetings.

[5]A total of 12 for Switzerland and the UK and, due to technical issues in the network gathering process, 11 for Germany and the US.

		Germany	Switzerland	United Kingdom	United States
Actors	Domestic	88 (10)	24 (3)	127 (7)	441 (71)
	Foreign	190 (23)	120 (8)	218 (16)	341 (62)
	Unknown	5 (1)	4 (1)	14 (2)	47 (8)
Hyperlinks	Domestic	2,808 (668)	180 (94)	1,071 (157)	20,511 (3,348)
	Transnational	5,800 (747)	387 (279)	2,518 (532)	14,234 (2,042)
	Unknown	6 (3)	4 (4)	156 (9)	1,121 (370)
Referential links	Domestic	5,830 (464)	127 (30)	11,785 (1,424)	48,277 (6,509)
	Transnational	5,342 (375)	472 (198)	16,308 (2,679)	35,932 (4,676)
	Unknown	2,762 (286)	206 (94)	13,810 (1,289)	28,923 (4,024)

Table 2.1: Key figures of issue networks.

Reported are the number of actors, the number of hyperlinks, and the number of referential links (rounded means, standard deviations in parentheses).

2.3.3 Extracting Referential Links

To identify the referential links in the text of the webpages, automated and manual content analysis methods were used. First, 10,095 actors were identified as important for the climate change discourses in Germany, Switzerland, the UK, and the US. This was done by the manual content analysis of 1,680 webpages, 1,545 newspaper articles, and 774 political documents dealing with climate change. The news articles and political documents were published by newspapers and political institutions located in the four countries and obtained through media database searches and web scraping governmental websites (cf. Häussler, Schmid-Petri, Adam, Reber, & Arlt, 2016).[6] They were collected at the same time as the issue networks, and the manual coded webpages were directly sampled from the issue networks' corpus. Trained coders identified the most important actors in every document and coded their scope if it was indicated in the text.[7]

The manual coded list of climate actors was merged with a generic list of around 1.3 million names of people and organizations (e.g., "Angela Merkel") as well as multiword units (e.g., "market economy"). We then used this extensive list as a dictionary to search for named entities in the text of the webpages. In addition, we trained two probabilistic sequence classifiers—one for German- and one for English-language documents—and used them for the identification of additional, previously unknown named entities. Conditional random field models (CRF) from the Stanford CoreNLP package (Finkel, Grenage, & Manning, 2005) were used as classifier algorithms. Of the additionally identified named entities, the 500 most frequent were then coded by a trained coder.

Whether the reference to a named entity counted as a national, transnational, or unknown referential link was inferred from both the scope of the webpage and the scope of the named entity (e.g. the mention of Angela Merkel on a webpage of Greenpeace Germany would thus count as a domestic referential link, the mention of Barack Obama on the same webpage would count as transnational link). As for the hyperlinks, only referential links from domestic actors' webpages were collected. The obtained numbers are summarized in Table 2.1. What is obvious, compared to

---

[6]The same search terms as for the webpages were used to index the newspaper articles and political documents.

[7]The coding was done by seven trained coders. As for the networks, the intercoder reliability was assessed by comparing the coders' classification with a master coding. For the scope variable (distinguishing 199 categories), Krippendorff's alpha was 0.82 and thus acceptable.

hyperlinks, is the higher number of unknown referential links. This is due to the often unspecific references. For example, actors often referred to "scientists" or "science" in general rather than a person or an institution specifically. In such cases, the scope of the named entity was labelled as unknown.

2.3.4 Measures

To measure the *degree* of transnationalization, we used the *E-I index* (Krackhardt & Stern, 1988). The index is a metric for what Koopmans and Erbe (2004) called "the relative density of public communication within and between political spaces" (p. 102). For the case of the national issue network i, the index is calculated as follows:

$$EI_i = \frac{t_i - d_i}{t_i + d_i}$$

where d_i denotes the number of domestic links and t_i the number of transnational links in the national issue network i. An E-I score of 1 would, therefore, indicate that every link is transnational. A score of -1 would indicate the opposite, that every link is domestic. An E-I score of 0 means that there are the same number of domestic and transnational links in the network. For the degrees reported in the following section, we first calculated the E-I ratios for all the issue networks gathered and then computed the mean ratio per country on this basis.

To measure the *scope* of transnationalization, we differentiated between *horizontal* and *vertical* transnational links (Koopmans & Erbe, 2004). Horizontal-transnational links connect actors/entities from different countries. This means that the primary scope of both source and target is national (but different). This is the case when a German think tank (e.g., Heinrich Boell Stiftung) links to a British NGO (e.g., Oxfam UK). A vertical-transnational link connects an actor with a primarily national scope with an actor/entity with a scope that is transnational. This could be, for example, the European Commission, representatives of the UN, or a globally active NGO (e.g., Greenpeace). The distinction between horizontal and vertical links is useful because it provides us with an indication of the importance of international/supranational institutions in climate politics relative to national ones. The scope of transnationalization of the issue networks was thus determined by the relative frequency of (1) domestic, (2) horizontal-transnational (differentiating between countries), (3) vertical-transnational, and (4) unknown links. The values presented in the following section are based on the mean number of links over all the issue networks per country.

2.4 Findings

2.4.1 Degree of Transnationalization

The different degrees of transnationalization are shown in Figure 2.1. Looking at the hyperlinks first, Figure 2.1 shows that there is a positive degree in all countries except the US. This means that the relative density of transnational hyperlinks in Germany ($EI_{H,DE} = 0.355$), Switzerland ($EI_{H,CH} = 0.272$), and the UK ($EI_{H,UK} = 0.399$) is higher than that of domestic links. In the US, however, the domestic hyperlinks are denser than the transnational ones ($EI_{H,US} = -0.179$).
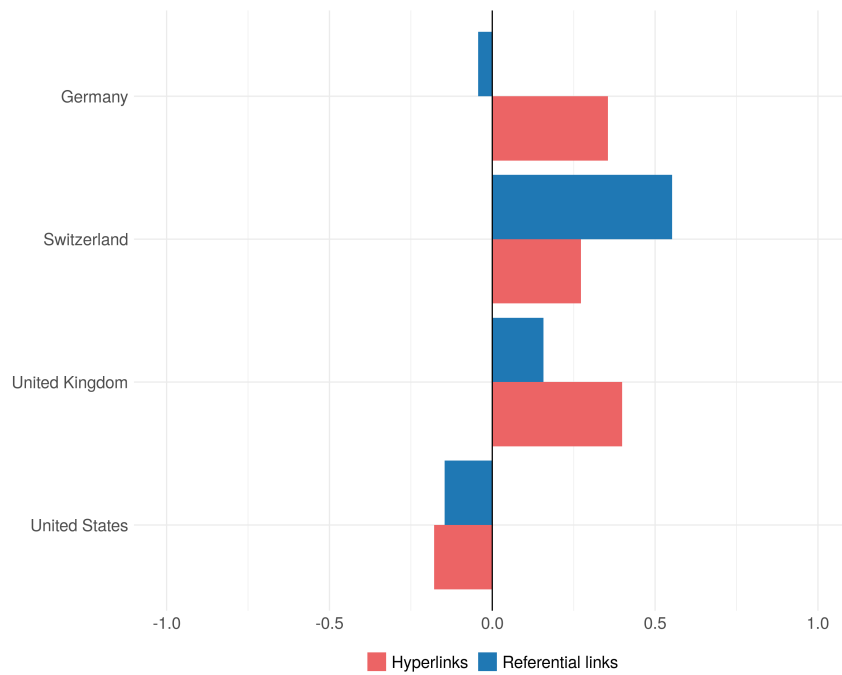
Figure 2.1: Degrees of transnationalization (E-I index) across countries and link type.

The picture is slightly different for referential links. In particular, German domestic actors have a faint preference to link domestically when they use referential links ($EI_{R,DE} = -0.043$). This is in contrast to hyperlinks, where the German actors have a clear transnational orientation. In the other countries, the E-I index for referential links goes in the same direction as for hyperlinks. In Switzerland, there is an even clearer external orientation ($EI_{R,CH} = 0.552$), whereas the UK ($EI_{R,UK} = 0.157$) and the US ($EI_{R,US} = -0.146$) are a little closer to the center in this case.

Overall, these results show that there is a pronounced, if varied, external orientation in all the European countries, with Switzerland being the most externally oriented. The US, on the other hand, shows a slight domestic, or internal, orientation.

2.4.2 Scope of Transnationalization

Figure 2.2 summarizes where the communicative links of the actors in the issue networks go. Looking at the transnational hyperlinks first, the figure shows that German actors link mainly to actors in the US (36.7%). A smaller fraction of the hyperlinks lead to actors from the UK (13%) and actors with a genuine transnational scope (12.7%). Only 2.9% of the hyperlinks in the German issue network point to Swiss actors, and even less (2.1%) to actors from other countries (i.e., Canada 0.5%, Ireland 0.5%, Liechtenstein 0.5%).

The Swiss actors, in turn, link preferably to Germany (26.1%). Less frequently, the hyperlinks of Swiss actors lead to websites from transnational actors (23.7%) and American webpages (16.5%). Only on very rare occasions (2.1%) does a link from a Swiss actor lead to another country (Denmark 0.2%, Austria 0.2%, Russia 0.2%).

British actors most frequently link horizontally to American actors (29.4%) and less often vertically to transnational actors (15.4%). The other transnational hyperlinks lead to actors from the Commonwealth (Canada 10.6%, Australia 5.1%, New Zealand 2.6%, India 0.2%), as well as Norway (3.1%). Other countries include France (0.5%) and Germany (0.1%).

Actors from the US link to British actors most frequently (14.6%) and less often to actors at the transnational level (7.9%), from Australia (6.7%), or Canada (6.4%). Hyperlinks to other countries (Germany 1.4%, New Zealand 1.2%, Czech Republic 0.9%, Denmark 0.2%, Sweden 0.2%) account for only 4.1% of the hyperlinks in the American issue networks.

The picture is only slightly different for referential links. First, there are more links to the transnational level (vertical transnationalization). Second, the scope for horizontal transnationalization is a little wider with referential links than with hyperlinks. German actors link to the transnational level the most (15.6%). Horizontally, they mainly link to actors from the US (8.8%), the UK (3.8%), and Switzerland (2%), followed by Brazil (1.5%), France (1.1%), China (0.6%), and Russia (0.5%). The horizon of Swiss actors mainly encompass the transnational level (19.3%). Horizontally, Swiss actors most often link to the US (12.3%), Germany (10.3%), and the UK (6.7%). Other countries include China (1.4%), Pakistan (1.1%), and Afghanistan (0.8%). On British webpages, users most likely find references to the transnational level (15.6%) or to actors from the US (12.4%). Less frequently, they find references to named entities based in Australia (2.2%), Canada (1.2%), Germany (0.9%), China (0.7%), Denmark (0.5%), France (0.5%), India (0.4%), Russia (0.4%), and Switzerland (0.4%). Finally, American actors refer most frequently to actors with a transnational scope (14%). Horizontally, they frequently refer to actors from the UK (6.1%). Other countries include Australia (1.2%), China (1.1%), Canada (1%), and Germany (0.8%).

Hyperlinks
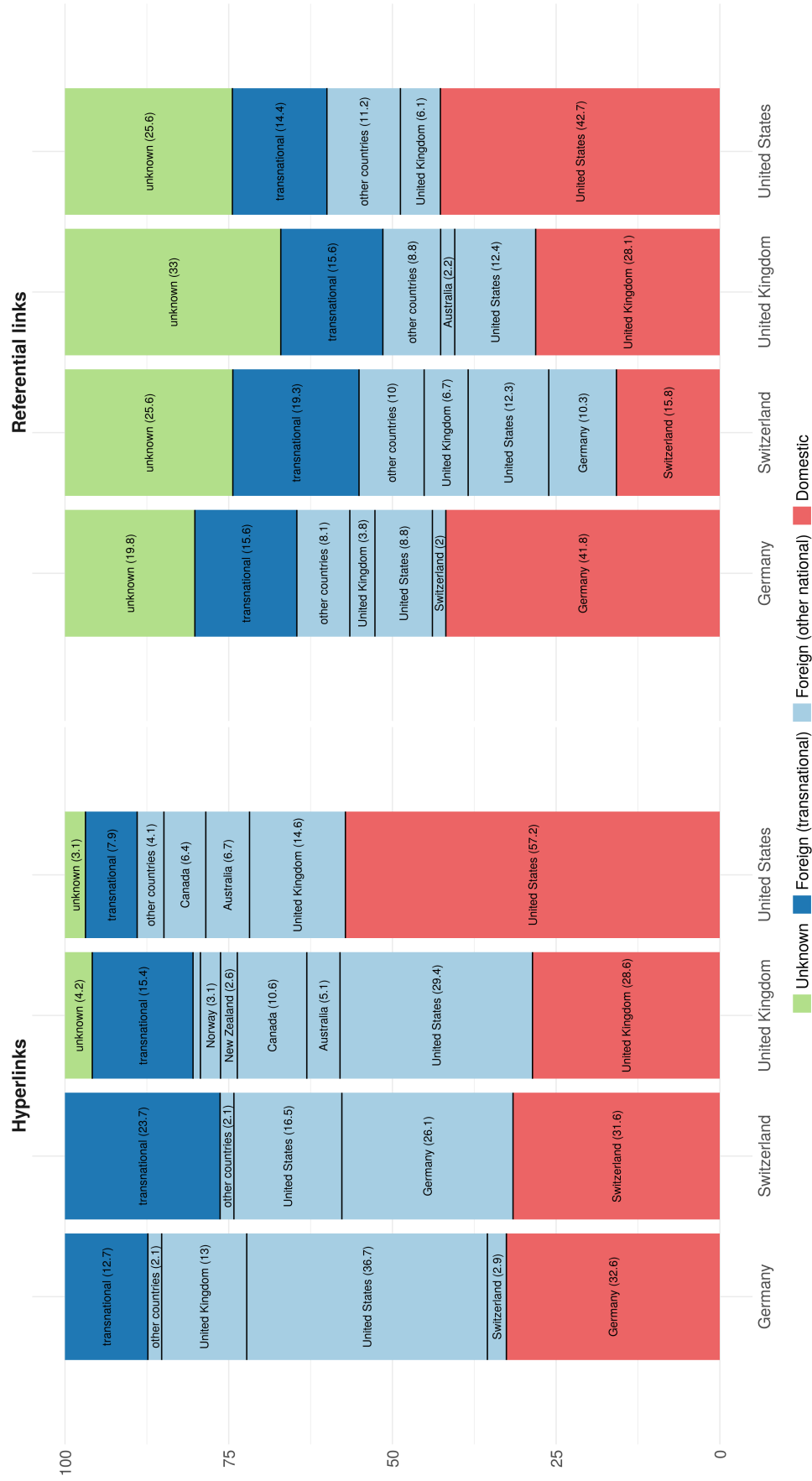
Referential links

Figure 2.2: Scope of hyperlinks and referential links across countries. Relative frequency, percentages in parentheses.

The category "other countries" summarizes countries with a share of less than 2 percent.

Legend: Unknown | Foreign (transnational) | Foreign (other national) | Domestic

2.5   Discussion

In this study, we evaluated the *degree* and the *scope of transnationalization* in the online discussion of climate change in four countries. The goal was to determine whether the discourse is nationally fragmented or transnationally integrated. Based on the findings reported in the previous section, we can say that the online discourses on climate change in Germany, Switzerland, the UK are transnationally integrated. In the US, the picture differs: despite a notable degree of transnationalization, there is a clear internal orientation in the US discourse. As in legacy media (Schäfer et al., 2018), horizontal transnationalization is more pronounced than vertical transnationalization in all countries. Thus, not only are supranational institutions and transnational civil society actors integrated into the national issue publics but also foreign national actors. This shows that even actors whose primary focus is on the domestic national sphere consider climate change a transnational issue. Consequently, transnational links appear to be a central tool in their online communication toolbox. They can be used to support foreign allies, to legitimize the own position, to educate the audience, or to mobilize the public (cf. Schäfer, 2012). Horizontal links allow the actors to refer to peers and political institutions on the national level, while vertical links highlight the specialized civil society actors and institutions at the international and supranational level (Bennett et al., 2015). However, the high number of horizontal links can be indicative that, despite the prominent international and supranational institutions dealing with climate change (e.g., EU, UN, IPCC), the national level is still the most important in climate politics. Whether this is actually the case, a more in-depth analysis of the strategies and motivations of the actors in the issue networks has to show.

Despite the relatively high degrees, the scope of transnationalization is limited to countries and actors of the Global North. This suggests that the focus of the public debate within the studied countries is primarily on those causing the problem, rather than on those affected by it. This is true for both hyperlinks and referential links and is highly problematic, as it prevents the actors of the Global South from speaking for themselves in the national discourses of the Global North. As with other issues and media (Rogers & Ben-David, 2008; Shumate & Dewitt, 2008; Yang, 2012), the Web seems not to eliminate the north-south divide but to deepen it.

Furthermore, as the results have shown, language regions limit the scope of transnational communicative links. Links from non-English-speaking countries to English-speaking ones—especially the US—occur more often than vice versa. However, this pattern is not climate change-specific and has been documented by other studies looking at hyperlink networks (Barnett, Chon, & Rosen, 2001; Barnett et al., 2011; Halvais, 2000; Takhteyev, Gruzd, & Wellman, 2012). The same is true for the fact that actors from smaller countries link to actors from bigger countries more often than the other way round (e.g., Switzerland to Germany or the UK to the US). The issue networks analyzed here also reflect a pattern known from legacy media, where the US typically receives the most attention of any other country in other nations' media (Chang et al., 2009; Himelboim, 2010; Segev, 2010; Segev & Hills, 2014). In return, the degree of transnationalization in the discourse on

climate change in American mass media has been identified as comparatively low (Broadbent et al., 2016; Schäfer et al., 2018). This matches our results, according to which the US is the only country with a negative E-I ratio, indicating a stronger internal than external orientation of the actors in the issue network.

Overall, the Web's capability to integrate the complex spatiality of climate change into national discourses has proven to be limited. Regarding an inclusive and ambitious global climate policy that aims to meet the targets of the UNFCCC, not only the absence of actors from the global South seems problematic, but also the high visibility of US actors in the European discourses. In the US, climate change is a highly controversial issue (Dunlap et al., 2016), which is reflected in a strong online presence of domestic climate skeptics (Elgin, 2015; Schmid-Petri, 2017). In this polarized context, both skeptics and advocates find political allies. Transnational links are therefore less important for US actors (Häussler et al., 2017). The situation is different in Europe, where climate change skeptics constitute a counter-movement in most countries. As politically isolated actors, they use the Web to form transnational advocacy networks with like-minded actors from other countries. Recent research has shown that skeptics from Germany and the UK link to US skeptics for this purpose (Adam, Häussler, Schmid-Petri, & Reber, 2019). Such an integration of skeptical actors into the national discourses of European countries may have a negative impact on the public support of constructive policies in these countries. In order to investigate this hypothesis, future research should not only unravel the linking behavior of the different camps (i.e., skeptics and advocates), but also show how they influence others—like citizens or the mass media—with their online communication behavior.

Moreover, future work should seek explanations for the reported patterns of national (non-)transnationalization. Besides country-specific factors, such as language and size, future studies should also take actor-specific factors, such as their position regarding climate change, into account. Such analyses would not only allow for a statement about which actors transnationalize the national discourses but also provide information about their possible intentions. Further studies may also conduct a combined analysis of network and content data (as in Schmid-Petri et al., 2018), as content provides context information for an accurate interpretation of communicative links (cf. De Maeyer, 2012). Missing context is a major limitation of the approach adopted here. In particular, it is not clear for referential links whether the actors are mentioned as speakers or addressees (cf. Koopmans & Erbe, 2004), nor whether the links are meant to be critical or supportive. Another limitation of this study relates to the number of languages considered. Although German and English are the languages spoken in the studied countries, national issue networks do not have to be monolingual. This holds true for multilingual Switzerland but also for many other countries. Future studies should, therefore, consider more languages, both for the selection of countries and for the indexing of webpages. In addition, more national discourses on climate change should be examined. It would be particularly desirable that the integrative performance of issue networks in the Global South is examined.

Finally, the findings have shown that although hyperlinks and referential links are both communicative links, their use patterns differ. The horizon is wider with referential links, yet at the same time, they are used for more vertical-transnational references, as well as non-specific references, than hyperlinks. Moreover, referential links are much more common than hyperlinks. This suggests that hyperlinks are used more selectively. They are a communicative tool that allows actors to make the Web content of others directly accessible to their readers (Park, 2003). This is neither always possible nor desired. For example, skeptics and advocates avoid linking to each other (Schmid-Petri et al., 2018). Referential links are less problematic in this respect: they do not make the content of the addressee directly accessible. They are also less language-sensitive and allow general references (e.g., to "the people"), as our results have shown. The analysis of hyperlinks therefore reveals not the complete picture. This should be taken into account when communicative links are used to study social and political processes.

2.6   Appendix

Starting points for the snowball-sampling of websites in Germany, Switzerland, the United Kingdom, and the United States.

Germany		
Advocates	Heinrich Boell Stiftung	http://klima-der-gerechtigkeit.boellblog.org
	Greenpeace Germany	http://www.greenpeace.de/themen/klima/nachrichten
	Potsdam Institute for Climate Impact Research	http://www.pik-potsdam.de/aktuelles?set_language=de
	WWF Germany	http://www.wwf.de/themen-projekte/klima-energie
Skeptics	Analyse+Aktion	http://astrologieklassisch.wordpress.com/tag/klimawandel
	EIKE - Europäisches Institut für Klima und Energie	http://www.eike-klima-energie.eu
	Klimaüberraschung	http://www.klima-ueberraschung.de
	Klimaskeptiker	http://www.klimaskeptiker.info
Switzerland		
Advocates	WWF Switzerland	http://www.wwf.ch/de/hintergrundwissen/klima/
	Alliance Sud	http://www.alliancesud.ch/de/cs
	ETH Zürich Klimablog	http://blogs.ethz.ch/klimablog/
	Klimainitiative	http://www.klimainitiativeja.ch/
Skeptics	Bilderberger	http://www.bilderberger.ch/?cat=1686
	Arlesheim reloaded	http://arlesheimreloaded.ch/category/klimawandel/
	We are change	http://www.wearechange.ch/hintergrund/umwelt-und-klima
	Klimaschwindel	http://www.klima-schwindel.com
United Kingdom		
Advocates	Greenpeace UK	http://www.greenpeace.org.uk/climate
	Oxfam UK	http://www.oxfam.org.uk/what-we-do/issues-we-work-on/climate-change
	Friends of the Earth UK	http://www.foe.org/projects/climate-and-energy
	WWF UK	http://www.wwf.org.uk/what_we_do/tackling_climate_change
Skeptics	The Global Warming Policy Foundation	http://thegwpf.org
	Global Warming Hysteria	http://www.globalwarminghysteria.com/
	Climate Resistance	http://www.climate-resistance.org/
	Repealtheact	http://repealtheact.org.uk/
United States		
Advocates	Climate Central	http://www.climatecentral.org
	Greenpeace USA	http://www.greenpeace.org/usa/en/campaigns/global-warming-and-energy
	Worldwatch Institute	http://www.worldwatch.org/climate-energy
	WWF U.S.	http://www.worldwildlife.org/climate/index.html
Skeptics	The Heartland Institute	http://heartland.org/issues/environment
	Climate Depot	http://www.climatedepot.com
	C3 Headlines	http://www.c3headlines.com
	Watts Up With That?	http://wattsupwiththat.com

Table 2.2: Starting points for the snowball-sampling of websites.

Chapter 3

Overcoming language barriers: Assessing the potential of machine translation and topic modeling for the comparative analysis of multilingual text corpora

Abstract: This study assesses the potential of topic models coupled with machine translation for comparative communication research across language barriers. From a methodological point of view, the robustness of a combined approach is examined. For this purpose the results of different machine translation services (Google Translate vs. DeepL) as well as methods (full-text vs. term-by-term) are compared. From a substantive point of view, the integrability of the approach into comparative study designs is tested. For this, the online discourses about climate change in Germany, the United Kingdom, and the United States are compared. First, the results show that the approach is relatively robust and second, that integration in comparative study designs is not a problem. It is concluded that this as well as the relatively moderate costs in terms of time and money makes the strategy to couple topic models with machine translation a valuable addition to the toolbox of comparative communication researchers.

3.1 Introduction

In communication research, comparing instances of public communication has become a popular approach to gain a better understanding of publics and discourses. This is because the comparative perspective allows us to draw conclusions beyond the individual case, either by testing theories in diverse settings or by evaluating the scope and significance of certain phenomena in different contexts (Esser & Hanitzsch, 2012). Even more so in times of digitalization, in which national borders are becoming increasingly permeable for communication flows, the comparative perspective is important to answer the question whether a phenomenon is common across the globe or distinctive to a certain country (Esser, 2013; Livingstone, 2012).

However, comparing instances of public communication is theoretically as well as methodologically challenging (cf. Livingstone, 2003). This is especially true if the compared cases happen to be in different languages. For such comparisons, not only must instruments be developed that work for all studied cases, but also must there be people with the right language skills to collect and analyse the data. This makes analyses of multilingual text corpora particularly complex, labor-intensive, and costly. It is therefore of great interest that now well-established automated content analysis methods can also be used for comparisons of public communication across multiple languages.

One such method, which is now a standard tool in communication research (Boumans & Trilling, 2016; Grimmer & Stewart, 2013), are so-called bag-of-words topic models. These are generative models which allow the identification of underlying thematic structures (i.e., topics) even in large amounts of text (Jacobi, van Atteveldt, & Welbers, 2016). They are called bag-of-words models because the order of the words in a document (i.e., its syntactic structure) is ignored in the modeling process. However, they are not designed for the analysis of multilingual corpora as the resulting topics directly depend on the vocabulary used in the documents (cf. Blei, Ng, & Jordan, 2003). Since languages have different vocabularies, topic models separate topics by language. Standard topic models are thus blind for thematic structures which cross language boundaries.

To solve this "confusion of tongues," there are two possible ways: either compute a topic model on the multilingual corpus and match the topics after or translate the corpus into a common language first and then compute the model. The first approach is challenging because additional internal or external information is needed to bring the topics together (e.g., Wikipedia entries on the same topic in different languages, as used in the Polylingual Topic Model by Mimno, Wallach, Naradowsky, Smith, and McCallum 2009). The second approach is easier, given that a low-cost and reliable translation of the corpus is possible. Thanks to improvements in cheap machine-translation services such as Google Translate (Lotz & van Rensburg, 2014), this seems to be a feasible option. In fact, it has recently been shown that machine translation and topic modeling can be combined to study public communication (De Vries, Schoonevelde, & Schumacher, 2018; Lucas et al., 2015).

This article continues along this path by assessing the potential of such a combined approach for comparative communication research. From a methodological point of view, the robustness of the approach is examined by determining whether the choice of translation service and method matters when it is used for analytical purposes. From a substantive point of view, it is examined whether topic modeling coupled with machine translation can be linked to existing theories and work from the field of comparative communication research in such a way that it is a valuable add-on to the toolbox. For both, the online discourses about climate change in Germany, the United Kingdom and the United States of America are used as test cases. As a global phenomenon, climate change is a widely studied topic in the field of comparative communication research. However, the focus has been on offline media so far. A multilingual comparison of online discourses is uncharted territory and thus an ideal research subject to assess the potential of topic modeling in combination with machine translation for comparative analyses.

3.1.1 Study 1: Robustness

De Vries et al. (2018) have shown that the translations of professional human translators ("gold standard") and Google Translate are quite similar when documents are considered a "bag of words" (i.e., ignoring the syntactic structure) and thus concluded that machine translation services are indeed useful tools when it comes to the analysis of multilingual text corpora with topic models.

But how much do translation results of different machine translation providers differ? Or to put it another way: Does it matter which translation provider is used for the translation of a multilingual text corpus when topic models are used for its analysis? To answer this question, this study compares two full-text translations of the same corpus—one by Google Translate and one by DeepL. This is the current top dog (Google Translate) compared to the rising star (DeepL) in the field of online machine translation providers. According to their own surveys, DeepL currently provides the best translations of all online translation services available (DeepL, 2017). However, it remains to be seen whether the better results of DeepL—in terms of readability—are also of importance in bag-of-word analyses or if Google Translate does the job equally well.

In their assessment, De Vries et al. (2018) focused on translations of whole documents. However, since most machine-translation services charge fees based on the number of translated characters, it would be significantly more cost-effective (as well as time-saving) if every word in a corpus has to be translated only once. The good thing about topic models is that they perceive documents just as vectors containing the count of each word within the document, ignoring the order in which they appear (i.e., bag of words). For the modeling process, all documents of a corpus are bound together in a so-called document-term matrix (DTM), containing the frequency of each word (term) in each document. The translation of the unique terms of a DTM (i.e., its vocabulary) would thus be an extremely efficient way to obtain a monolingual corpus, as every word must be translated only once instead of multiple times for every document. One potential problem with such a term-by-term translation, however, is that the absence of the syntactic structure may affect

the outcome of the translation. The question thus is whether this shortcut is a valid alternative to the more expensive full-text translation. Or more precisely: Does it matter which translation method is used for topic model analyses?

Thus far, only Lucas et al. (2015) compared a DTM translation to a full-text translation of the same corpus in combination with topic modeling. After translating Chinese and Arabic tweets into English, they computed topic models on both corpora and assessed their similarity by comparing the content of the topics (i.e., their word probabilities). They concluded that "two investigators using different translation methods might have reached similar substantive conclusions" (Lucas et al., 2015, p. 274). This result is checked here by systematically comparing the term-by-term translation of a corpus (DTM translation) to the full-text translations of the same corpus mentioned above. Google Translate was chosen for the DTM translation. There are two reasons for this choice: First, Google Translate was also the choice of Lucas et al. (2015) for their DTM translation. This allows the translation service to be kept constant. Second, DeepL restricts the use of their API in a way that a DTM translation does not save time (and money) compared to a full-text translation.[1]

3.1.2 Study 2: Integrability

The number of socioscientific studies that have explicitly used machine translation is still limited. Thus far, most studies have had a methodological focus, as they either proposed and evaluated concrete analytical procedures that build on machine-translated texts (Agarwal, Xie, Vovsha, Rambow, & Passonneau, 2011; Balahur & Turchi, 2014; Pennings, 2011), evaluated the quality of machine-translation services (e.g., Hampshire & Porta Salvia, 2010; Lotz & van Rensburg, 2014), or assessed their usefulness for analyses with text models (De Vries et al., 2018; Lucas et al., 2015). Studies that have used machine translation to examine substantive questions are especially rare. Benoit, Schwarz, and Traber (2012) translated parts of their corpus of parliamentary speeches for an analysis of legislators' policy preferences and Zhou, Cristea, and Roberts (2015) used Google Translate for a sentiment analysis of war-related Wikipedia articles. In any case, the possible value of machine translation and topic modeling for comparative communication research has hardly been shown (the study by Lucas et al., 2015 is the notable exception). For this reason, a substantive question is examineded here. Specifically, it is examined whether actors in Germany, the United Kingdom, and the United States debate the same aspects of climate change on their websites and whether they accord them similar relevance.

Since climate change is a global problem that might only be solved globally, it is important to know whether the public discourse about it is rather national or transnational in nature. To assess that, issue networks are studied here. That are networks of public communication, consisting

---

[1] This changed with the introduction of the DeepL Pro plan. For this study, however, an earlier and more restricted version of the API was used. Due to the restrictions (i.e., limited number of translation requests per time), the translation of the DTM would have taken considerably longer than the full-text translation.

of the websites as well as the hyperlinks connecting the contents of these websites (Marres & Rogers, 2005). Bennett et al. (2015) argue that such issue networks are full-fledged issue publics, as they are open to connection and contestation among all kinds of actors. One way to measure the transnationality of such issue networks is through the analysis of hyperlink connections (e.g., Bennett et al., 2015). Another way is to analyze the contents of the websites. The second approach was chosen here, as discourse convergence has been identified as one of the main indicators for the transnationality of public discourses (Eder, 2000; Eder & Kantner, 2000; Kantner, 2004). At the core stands the question of "whether speakers in different national public spheres identify the same issue as important, accord them similar relevance, and employ similar problem definitions" (Wessler, Peters, Brüggemann, Kleinen-von Königslöw, & Sifft, 2008a, p. 11). From that point of view, a public discourse is transnational "if within an anonymous mass public the same issues are discussed at the same time under similar criteria of relevance" (Eder & Kantner, 2000, p. 315, translation by the author). For the this study, it means that actors in Germany, the United Kingdom, and the United States have to discuss the same aspects of climate change on their websites and accord them similar relevance for it to be a transnationalized debate.

Although there is a rich and growing body of literature regarding online climate change communication (for an overview, see Schäfer 2012), comparative studies are still the exception (e.g., Häussler et al., 2017; Jang & Hart, 2015). The content of the public debate about climate change has thus far only been compared in regard to Twitter (Kirilenko & Stepchenkova, 2014) and the mass media (e.g., Grundmann & Scott, 2014; Ivanova, 2017; Ivanova, Schmidt, & Schäfer, 2014; Schmidt, Ivanova, & Schäfer, 2013) across countries and languages.

This study therefore addresses three questions: First, how much do translation results of different machine translation providers differ? Second, is the translation of a DTM a valid alternative to the full-text translation of a corpus? Third, is the online climate change discourse the same in Germany, the United Kingdom, and the United States? All three questions aim to clarify whether the combination of machine translation and topic modeling are a valuable addition to the methodical toolbox of comparative communication researchers. In the following section, the data-collection procedure as well as the steps taken for the preprocessing and translation of the corpus is described. Then, the comparison of the translation providers as well as the translation methods is explained (study 1), and the procedure for the cross-country analysis of the climate change discourse is laid out (study 2). The results of both studies are then presented. The paper concludes with a discussion of the results as well as the promises and limitations of the applied approach for comparative communication research.

3.2 Data and Method

3.2.1 Web Scraping and Building of the Corpus

The corpus for both studies—the comparisons of translation providers/methods and the comparison of the climate change discourse in Germany, the United Kingdom, and the United States—was collected in June 2014 in the course of a larger project conducted with collaborative partners (Adam et al., 2016). That project proceeded in four steps to gather all sorts of websites with an interest in climate change, which includes, but is not limited to, websites of civil society actors (such as NGOs, Blogs, universities, churches), the media, and governmental bodies.

A snowball-sampling strategy was used to harvest hyperlink-networks originating from carefully selected starting points in four countries (four climate advocates and four climate skeptics in Germany, the United Kingdom, and the United States; see the list in Appendix A, chapter 3.5.1). This was done with the help of the crawler software Issue Crawler (Rogers, 2013). The starting points were chosen based on literature reviews, expert interviews, and country-specific Google searches.

The crawled websites were then indexed according to keywords to ensure that they related to climate change, and they were downloaded if at least one keyword was mentioned once.[2]

To obtain additional information about the actors included in this sample, the country of activity (national, transnational) as well as the name of the actor was coded manually by two coders using information found on the "about us" or a similar page.[3]

The final step was the extraction of plain text from the archived HTML documents. To do this, all HTML markups were deleted using the Java-based content extraction library Apache Tika. Then, meaningless terms used on most web pages (e.g., navigation elements, copyright information) were filtered out by deleting all sentences containing "regular expressions" (i.e., character sequences defining search patterns) from a blacklist. The remaining documents were then categorized according to their language (German or English) by a language-detection algorithm and marked as duplicates if their similarity, defined by the Jaccard index on their word set, was above a threshold of 0.95.

For the analyses conducted in this study, the corpus was reduced by removing all duplicates, all websites from non-German, non-British, and non-U.S. American actors, and all websites not written in German or English. For June 2014, this resulted in a total of 875 unique web pages published by 95 German actors, 2,172 web pages published by 181 British actors, and 3,896 web pages published by 539 U.S. American actors. From the total of 6,843 web pages, 633 were written

---

[2]Keywords: climate change, global warming, Klimawandel*, globale* Erwärmung

[3]The intercoder reliability was measured by comparing the coders' classification with a master coding. Krippendorff's alpha was .93 for the country of activity (distinguishing 199 categories). The intercoder reliability was thus satisfactory.

in German and 6,310 in English because some German actors also publish in English. For the assessment of the translation providers/methods, only the 633 web pages written in German were considered.

3.2.2 Translation and Preprocessing

The translations process of the 633 German full-texts was straightforward: The documents were translated via Google Translate's and DeepL's API. However, due to length restrictions by the DeepL API, the texts were segmented into sentences, then translated, and finally put back together. The same was done for Google Translate to avoid a bias. All the textual data manipulation as well as the translation itself was done in *R*. For translations with Google Translate, the *translateR* package (Lucas & Tingley, 2014) was used. The package provides easy out-of-R access to the Google Translate API. A similar function was built to translate the documents with DeepL via their API.[4]

The process of the term-by-term translation was less straightforward, as the corpus had to be converted to a DTM first, requiring the documents to be preprocessed. The problem with most preprocessing steps is that they are not language-independent. This is true for common procedures like stop word removal, stemming, decompounding, or lemmatization. An attempt was made to include as much grammatical information as possible in the DTM for the translation. Thus, as little preprocessing as possible was done before the translation. No decompounding or stemming procedures were applied, and all the letters were kept as they were (uppercase or lowercase). However, all 633 German documents were split into individual words (tokenization), and then all punctuation marks (Unicode "Punctuation" [S] class), symbols (Unicode "Symbol" [S] class), and tokens consisting only of numbers were removed. Hyphens were not removed, as that might change the meaning of terms (e.g., "EU-Abgeordneter" [Member of the European Parliament]). From these tokens, a DTM was created. This step, as well as the rest of the preprocessing, was done in *R* using the *quanteda* package (Benoit et al., 2017). For the actual translation, the vocabulary (i.e., the unique terms) was extracted from the German DTM, translated with Google Translate using the *translateR* package (Lucas & Tingley, 2014) and placed back into the DTM as English terms.

A specific feature of the German language is the concatenation of multiple words into a single word. Because such compounds (e.g., "Klimawandel") are less common in English, they were often translated as ngrams (e.g., "climate change"). To avoid systematic differences in the vocabulary, all the newly created ngrams in the "German" DTM were split into unigrams (e.g., "climate" and "change") before all further steps. In addition, duplicate terms were removed.

Apart from counting the words, the aim of preprocessing is to remove noise and to extract as much useful information as possible by "treating words with very similar properties identically

---

[4]The R code produced for the analyses done here is available on GitHub: https://github.com/ureber/mt-paper

and removing words that are unnecessary to our interpretation and our model" (Lucas et al., 2015, p. 257). Hence, the final preprocessing steps—which were conducted on all the translated DTMs for both studies—included the conversion of all letters to lowercase, the removal of terms with fewer than three characters, lemmatization, the removal of stop words, and relative pruning (for further reflection on preprocessing, see Maier, Waldherr, Miltner, Wiedemann, et al. 2018). The preprocessing steps were conducted in this order. It is important to note that a different order would lead to different results (Denny & Spirling, 2018). The conversion to all lowercase was done for the sake of term unification. This was necessary because of words used, for example, at the beginning of a sentence. Stop word removal and the removal of terms with fewer than three characters (e.g., "to," "on," "dr," "mp") was done to remove terms that are extremely frequent or unspecific and thus not helpful as indicators for a document's content (Salton, 1991). For stop word removal, a look-up list with the most common stop words (e.g., "from," "after," "the") as well as stop words specific to websites (e.g., "click," "login," "comments") was used. The list was put together specifically for this project. Lemmatization was done to convert inflected words to their base forms (e.g., "warming" to "warm"). Since the declination or conjugation of a word is usually not indicative of its meaning, lemmatization is used to combine words with the same meaning (Lucas et al., 2015). This reduces the dimension of the model input and generally improves the results (Jacobi et al., 2016). A comprehensive look-up list with base forms of inflected words was used for the lemmatization. Finally, relative pruning was done to remove extremely rare as well as extremely frequent terms. All terms were removed that occurred in less than 0.5% or more than 99% of the documents, two common thresholds (e.g., Grimmer & Stewart, 2013). If there is a theoretical interest in finding patterns of words that are used across documents, the removal of very infrequent terms is recommended because they do not contribute much information to the identification of document similarities (Denny & Spirling, 2018, p. 8). The removal of very frequent terms is advisable for the same reason as the removal of stop words—they are unspecific for a document's content and therefore add no helpful information to the topic model.

3.2.3 The Structural Topic Model

For the analyses done here, four topic models were needed: three to compare the translation providers/methods (study 1) and one for the cross-country comparison of the climate change discourse (study 2). The first three models are based on the 633 differently translated web pages. The fourth model is based on the entire corpus of 6,843 German, British and U.S. American web pages.

For modeling, the structural topic model framework was used (STM; Roberts, Stewart, & Airoldi, 2016; Roberts, Stewart, Tingley, & Airoldi, 2013).[5] It builds on the same principles as the latent Dirichlet allocation (LDA; Blei et al., 2003)—the most common topic model in the social

[5]See Roberts et al. (2016, 2013) and the R package's vignette (Roberts, Stewart, & Tingley, 2017) for a comprehensive description of the framework's technical details.

sciences—but allows users to incorporate additional information and covariates into the model (Roberts et al., 2016). These covariates can affect either the topical prevalence or the topical content. Topical prevalence refers to how much of a document is associated with a topic (e.g., German actors are more likely to speak about topic 1 than British actors). Topical content relates to the words used to discuss a topic (e.g., German-speaking actors are more likely than English speaking actors to use a particular word when they discuss topic 1). The advantage of the STM is that topical prevalence or topical content is not assumed to be constant across all documents, but rather may vary depending on its properties (Roberts et al., 2014). The STM framework, therefore, allows both the identification of topics and the estimation of their relationships to document properties.

Moreover, topical content covariates allow the model to "condition away systematic differences within the corpus that are not of primary interest" (Lucas et al., 2015, p. 263). It could be that the translation led to minor but systematic differences in the vocabulary of the documents. Since the primary interest is in the actors' use of different topics based on their country of action (i.e., the big picture), small differences in the vocabulary should not influence the model. In this case, the inclusion of the documents' original language as a topical covariate allowed the STM to capture systematic differences in the frequency a word was used by German and English speaking actors without affecting the overall outcome of the model. Technically this is done by defining "the distribution over the terms associated with the different topics as an exponential family model, similar to a multinominal logistic regression" (Roberts et al., 2016, p. 989). Thus, a topical content covariate indicating the document's original language was included in the multinational model for the comparison of the climate change discourse, but not for the three models used for the comparison of the translation providers/methods.

The modeling procedure itself was done in R using the *stm* package (Roberts et al., 2017) and included the following steps. First, the optimal number of topics was evaluated. This was not a straightforward routine, as there is not a "right" number of topics for a given corpus or research question (Grimmer & Stewart, 2013, p. 285). For both studies several models were estimated, each with a different number of topics. To find the optimal number of topics, the models were then compared by three diagnostic indicators: their held-out likelihood (Wallach, Murray, Salakhutdinov, & Mimno, 2009), their residuals dispersion (Taddy, 2012), and their lower bounds. For the analysis of the climate change discourse, however, diagnostic indicators were not the only criteria since the overall objective was to obtain the best solution in terms of interpretability (Maier, Waldherr, Miltner, Wiedemann, et al., 2018). The models' outputs were thus also compared qualitatively by looking at the topics' top words (determined by their highest marginal probability). Considering diagnostic indicators and the models' overall interpretability, the 30-topics version of the model was ultimately chosen for study 2. For study 1, the 40-topics solution was chosen based on models' diagnostic indicators.

One challenge of topic models is that topics are abstract objects consisting of clusters of words that likely co-occur in the corpus's documents. To make use of them, they must be interpreted in theoretical terms (Jacobi et al., 2016). In the climate change case studied here, topics were examined through the theoretical lens of "same issue at the same time under similar criteria of relevance" (Eder & Kantner, 2000, p. 315, translation by the author). The issue in this case is climate change. Thus, this study approaches topics as "criteria of relevance" of the climate change issue. This means that a topic theoretically represents a particular aspect of the climate change discourse that can be interpreted and named. It does not mean that word clusters necessarily represent a coherent position or frame, but they do represent at least some kind of sub-issue (Maier, Waldherr, Miltner, Jähnichen, & Pfetsch, 2018). However, not all topics are interpretable. Some do not represent a coherent concept or meaningful aspect of the debate and are thus hard to describe. Others are so-called boilerplate topics with no substantive meaning at all (Mimno & Blei, 2011). Such topics were excluded from the analysis because they do not help to answer the research questions. For the exclusion as well as for the labeling, the topic's top words and 10 randomly sampled documents with a relatively high probability (> 0.6) of the topic were considered.

Label	Top words
Causes / effects of climate change	co2, atmosphere, carbon, increase, effect, ocean, temperature, greenhouse, gas, water
Climate change evidence	warm, global, temperature, chart, climate, co2, emission, year, trend, change
Climate change research	climate, research, change, science, university, institute, work, policy, study, impact
Climate modeling	climate, model, ipcc, prediction, study, predict, scientist, scientific, science, evidence
Climate politics	climate, change, government, green, emission, country, carbon, guardian, policy, action
Climate politics and science	science, climate, political, change, scientific, argument, make, claim, public, debate
Climate scepticism	solar, climate, radiation, watt, sun, wuwt, model, vapor, anthony, pinterest
Doubting climate research	paper, science, gore, publish, skeptic, write, review, medium, mann, michael
Economy and climate politics	company, industry, group, fund, oil, report, project, airport, business, plan
Energy consumption	energy, green, save, make, recycle, home, heat, reduce, waste, efficient
Energy sector	energy, gas, power, wind, fuel, coal, emission, price, cost, electricity
Environmental activism	wwf, work, change, campaign, climate, centre, live, business, people, sustainable
Extreme weather	climate, change, report, weather, research, publish, storm, impact, extreme, read
Food / health	food, science, movement, crop, environmental, ddt, permaculture, malaria, year, issue
Humanity	world, people, human, life, society, change, social, future, live, idea
Melting ice	ice, sea, arctic, level, rise, melt, ocean, year, glacier, polar
Science	physic, space, earth, science, scientist, planet, sun, nasa, system, year
Scientific results / concensus	warm, global, climate, change, scientist, report, temperature, year, world, ipcc
Temperature	temperature, warm, period, year, change, climate, record, global, trend, past
U.S. environmental politics	state, u.s, obama, american, epa, president, tax, federal, rule, regulation
U.S. fiscal policy	obama, spend, house, climate, democrat, year, president, republican, make, time
Wildlife protection	read, specie, climate, environment, fish, animal, green, wildlife, forest, energy

Table 3.1: Labels and top words of the multinational climate change model.

Top words are determined by their marginal highest probability.

The decision process was guided by two questions (Maier, Waldherr, Miltner, Wiedemann, et al., 2018): (1) Does the topic depict a coherent (possibly controversial) aspect of the climate change discourse? (2) How can this aspect be described most comprehensively? If the top words and the documents pointed in a different direction or only one could be interpreted meaningfully, the topic was excluded. For the model used in study 2, eight topics were excluded, and 22 were labeled. Table 3.1 summarizes the labels and top words for the analysis of the online climate

change discourse in Germany, the United Kingdom, and the United States. For this model, the German documents were translated as full-texts with DeepL.

3.2.4 Study 1: Robustness

Following De Vries et al. (2018), the vocabulary of the translated and preprocessed DTMs were compared to obtain a detailed impression of how much the results of the two machine-translation providers differ and whether the translation of a DTM is a valid alternative to the full-text translation of a corpus. The first comparison focused on the level of documents. Using the cosine similarity, it was measured how similar the term vectors of a document are when the document was translated as full-texts with DeepL (DL/FT), as full-texts with Google Translate (GT/FT), and term-by-term with Google Translate (GT/DTM). Second, at the corpus level, the comparison of the unique terms gave an impression of both how comprehensive as well as how exclusive the vocabulary in the corresponding DTMs is.

De Vries et al. (2018) further assessed the similarity of translated corpora by comparing the results of topic models computed on the translated corpora. This makes sense, since it ultimately depends on the similarity of the topic models whether the conclusions drawn from the corpora are the same. Therefore the focus here was primarily on the topics' interpretations as well as on their prevalence in the documents.

DL/FT	GT/FT	GT/DTM
ipcc	ipcc	climate
change	co2	ipcc
catastrophe	catastrophe	lie
co2	lie	co2
lie	state	world
state	world	catastrophe
world	people	front
people	report	year
year	earth	kyoto
panel	year	people

Table 3.2: Top words of matched topics describing climate change as lie.

The first step was to see which topics can be matched with each other. To do so, the probability values of 30 top words were compared all labeled topics. Building on Niekler and Jähnichen (2012) approach to match topic model results, the cosine similarity was used to measure the topics' similarity. All topic pairs with a relatively high cosine similarity (> 0.4) were then examined in depth. This means that the previously assigned labels as well as the top words of the topics were compared qualitatively with each other. If two topics described the same thematic aspect of the climate change discourse, they were considered a match.

Table 3.2 shows an example for three matched topics. When looking at the top words, it can be stated that they describe the same thematic aspect of the discourse. This impression is confirmed

by the measured cosine similarities. The resulting similarity of 0.859 for the Google Translate (GT/FT) and the DeepL (DL/FT) translations indicates a very good match. The similarity of the GT/DTM's topic with the others, however, not quite as good. The cosine similarity is 0.565 for the match with the DL/FT's topic and 0.626 for topic of the GT/FT model.

The second step focused on the bigger picture by looking at the number of matched topics as well as at the average cosine similarity. The more topics could be matched, the more similar two models are. Ideally, each topic could only be matched with one topic from the other model (unique match). In practice, however, a topic was often matched multiple times (multi-match). The interpretation of such multi-matches is difficult, as they indicate both similarity and difference of the compared models at the same time. Since the reason for multi-matches lies not only in different corpora (due to different translations), but also in the topic models' generative process (cf. Maier, Waldherr, Miltner, Wiedemann, et al., 2018), only those pairs out of the multi-matches with the highest cosine similarity (first matches) as well as the unique matches were considered in the final step of the study.

This step concentrated on the question whether the matched topics have the same prevalence in the documents of the translated corpora. Or in other words: Is the distribution of the matched topics the same across the documents? This question is central, since most studies using topic models just rely on the topic proportions to draw conclusions from a corpus. As suggested by De Vries et al. (2018), the matched topics' distributions over all documents were correlated with each other to determine their similarity.

3.2.5   Study 2: Integrability

To measure the discourse convergence of the climate change discourses in Germany, the United Kingdom, and the United States, the multinational topic model was analyzed in two ways: (1) a comparison of the mean topic proportions per country, (2) a comparison of the relative frequency of topics as the top topics per country. The top topic is the topic with the highest average probability over all web pages of an actor. In the first case, differences between countries were tested for significance using ANOVA. In the second case, Fisher's exact test was used due to the occasionally small number of cases in which a topic was the top topic.

The topic model results are generally discussed directly at the individual document level (i.e., web page level) because they are the main unit for the calculation of the model (i.e., "bag of words"). In this study, however, the question is whether actors in Germany, the United Kingdom, and the United States discuss the same aspects of climate change on their websites. Thus, the focus is on the actors' websites rather than on the individual web pages. In this case, the actor definition includes all individuals or organizations who are represented in the sample with at least one web page. Actors can therefore be political actors (e. g. governments, offices, parties), civil society actors (e.g., environmental organizations, research institutions, churches, foundations,

citizens' initiatives, blogs), media, or companies. For the analysis, the individual web pages were aggregated to actors based on the manual coding of the "about us" pages. The aggregation was done by calculating the mean topic probability over all web pages of an actor. In contrast to summing up the topic probabilities, calculating the mean maintains the bag-of-words logic on the actor level and therefore allows a direct comparison of actors further in the analytical process.

3.3   Results

3.3.1   Study 1: Robustness

The first comparison to assess whether the different translation providers/methods produce similar results is that of the translated documents. For each document, the cosine similarity between the word vectors (i.e., "bag of words") of the different translated versions was calculated. For two identical translations, the cosine similarity would be 1. If no words in both translations are identical, the similarity is 0.
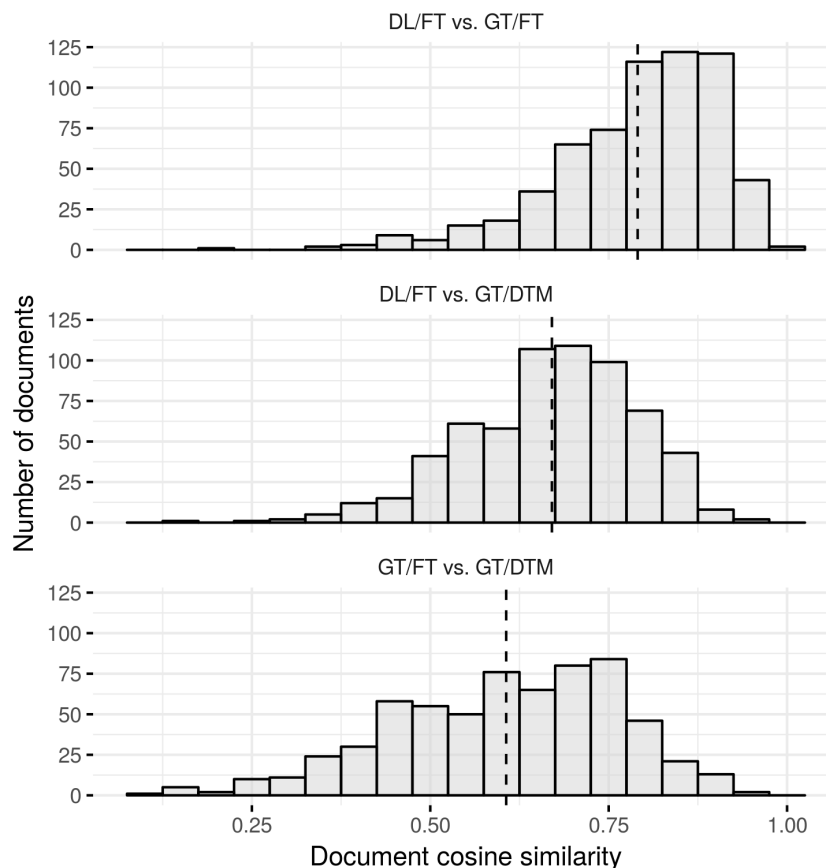

Figure 3.1: Distributions of cosine similarity between documents per translation method pair.

Figure 3.1 shows the distribution of cosine similarities between document pairs (n = 633) for every translation method pair. The figure shows that Google Translate (GT/FT) and DeepL (DL/FT)

produce quite similar results when whole web pages are translated (M = 0.791, SD = 0.117). However, if only the single words of the DTM were translated (GT/DTM), the similarities between the documents are smaller, although still respectable. An ANOVA showed that the differences between the pairings' distributions are significant (F(1, 1897) = 216.6, p < 0.001).[6] Interestingly, the results of the DL/FT and the GT/DTM translation are closer together (M = 0.670, SD = 0.120) than the GT/FT and the GT/DTM translation (M = 0.606, SD = 0.158), even though the DTM translation was also done with Google Translate.
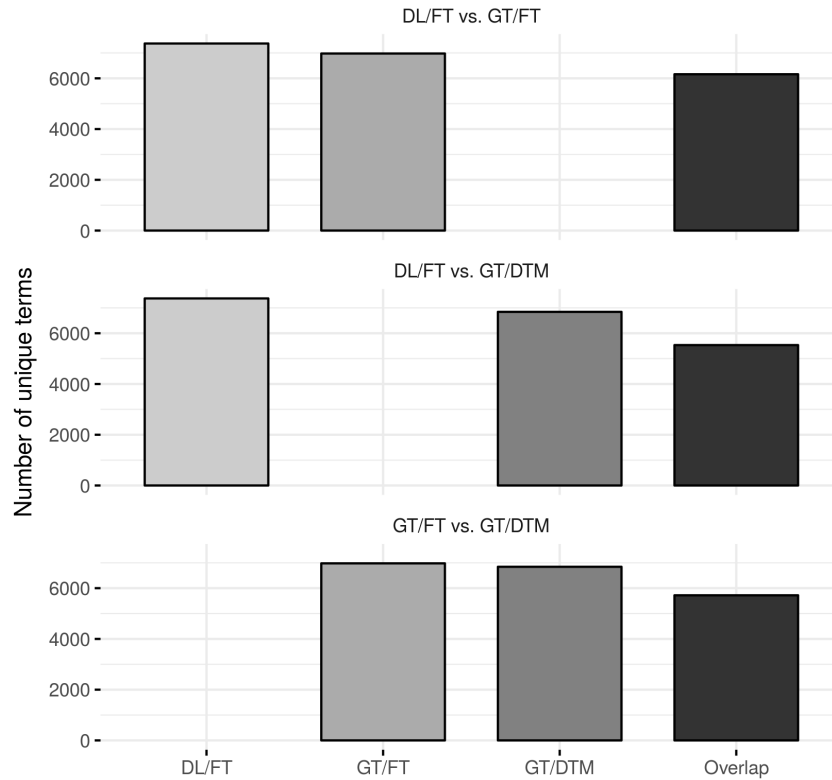
Figure 3.2: Unique terms in the translated and preprocessed DTMs and their overlap.

A second way to compare the translations on a corpus level is to compare the vocabulary in the respective DTMs. A larger vocabulary means that thematic structures are represented in a more detailed and thus probably more accurate way. Furthermore, if large parts of the vocabulary match, the texts have been translated correspondingly. As shown in Figure 3.2, the vocabularies in the full-text translations are slightly larger than in the term-by-term translation of the DTM. The full-text translation with DeepL (DL/FT) produced the vocabulary with the most unique terms (features; 7,372), followed by the full-text translation with Google Translate (GT/FT: 6,978). The term-by-term translation with Google Translate (GT/DTM) produced the result with the smallest number of unique terms (6,843). The number of overlapping features is also larger for the two full-text translations (6,161 terms) than for the full-text translations versus term-by-term translation (GT/FT vs. GT/DTM: 5,719; DL/FT vs. GT/DTM: 5,533 terms). Not surprisingly, but

---

[6]The documents at the lower end of the left tails are mainly longer blog posts with many special characters and unconventional punctuation. In such cases, sentence decomposition often led to sentence fragments, which were then translated differently by the translation services.

other than above, the results of the GT/FT and the GT/DTM translations are closer to each other than to the DL/FT translation. However, the DeepL translation is the most detailed and thus probably the most accurate.

The apparent reason for the smaller vocabulary of the GT/DTM translation is that words with several meanings (depending on the context) are lumped together and then translated with one and the same term. Part of the semantic information is therefore lost in the process of a DTM translation. Whether this causes major divergences in the corpus can be checked by comparing the results of the topic models.

The first comparison of the topic model results denotes to the number of matched topics. For every translated corpus a model with 40 topics was computed. The models were then interpreted as described above. For the DL/FT model a total of 29 topics could be interpreted and labeled based on their top words and a selection of documents. The same was true for a total of 28 topics in the GT/FT model and 29 topics in the GT/DTM model. Thus, there were no notable differences in the interpretability of the three models.
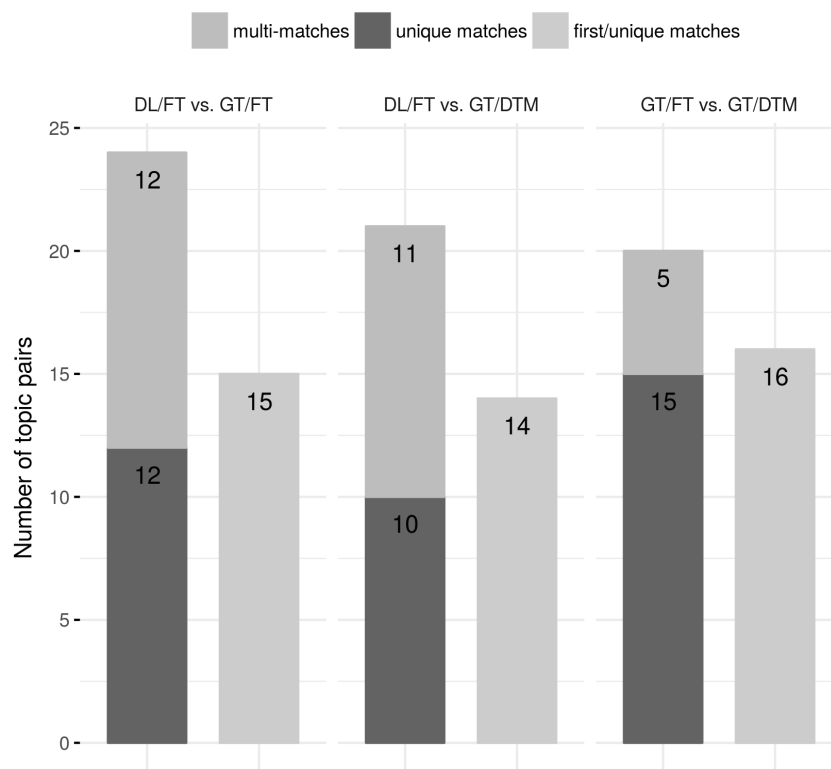
Figure 3.3: Matched topics.

Of the 29 respectively 28 topics a total of 24 could be matched in the DL/FT versus GT/FT comparison. However, 12 pairings were multi-matches (i.e., one or both topics of the pair is also involved in another match). If only the pairings with the best cosine similarity (first matches) as well as the unique matches were considered, 15 topics could be matched. For the DL/FT versus GT/DTM comparison a total of 21 topics could be matched, but only 14 as first/unique matches.

Finally, for the GT/FT versus GT/DTM comparison 20 topics could be matched, 16 of them as first/unique matches. Figure 3.3 illustrates these findings.

If the multi-matches are included in the picture, a considerable part of the interpretable topics could be matched in all three cases. Comparing the two full-text models (DL/FT vs. GT/FT), only 4 respectively 5 topics have no equivalent in the other model. No matter which of the two models is used, the overall picture regarding topical content is very similar. Not quite as good is the result when comparing the full-text models with the GT/DTM model. However, in both cases more than two thirds of the labeled topics could be matched. The majority of the topics therefore point in the same substantive direction. If only the first matches are considered, the differences between the models vanish, as only about half of the topics could be matched in all three comparisons.
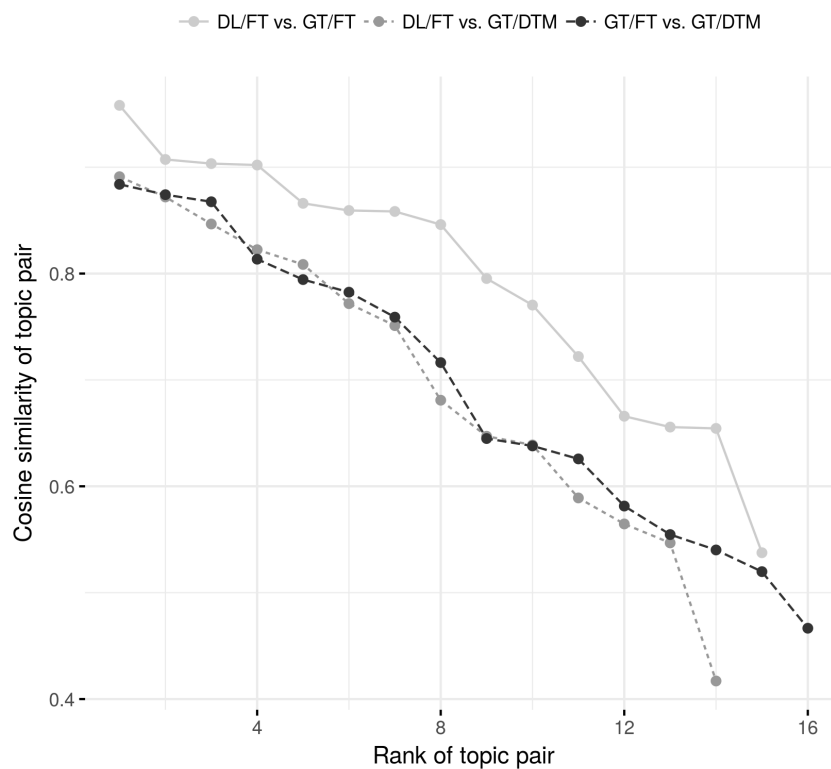
Figure 3.4: Cosine similarity of matched topics (first matches).

When looking at the first matches more closely, however, it becomes clear that those of the full-text models (DL/FT vs. GT/FT) have a higher mean cosine similarity (M = 0.793, SD = 0.121) than when they are compared to the GT/DTM model (DL/FT vs. GT/DTM: M = 0.703, SD = 0.142; GT/FT vs. GT/DTM: M = 0.691, SD = 0.137). Figure 3.4 illustrates this. With regard to topical content it can therefore be said: The two full-text models are more similar to each other than they are to the GT/DTM model.

But are the matched topics equally important for the same documents? To measure this, the individual topics' proportions over all web pages were correlated per translation method pair. Figure 3.5 shows the overall picture. The descriptive figures are summarized in Table 3.3. With

mean correlation coefficients between 0.6 and 0.7 in all three cases, most matched topics correlate quite well. Although the resulting coefficients are wider distributed in the comparisons of the full-text models with the GT/DTM model, the results do not differ significantly from the comparison of the two full-text model (ANOVA: $F(1, 43) = 0.405$, $p = 0.569$). Looking at the topical prevalence of the topics, it can thus be said that the conclusions drawn from the different models would be quite similar.
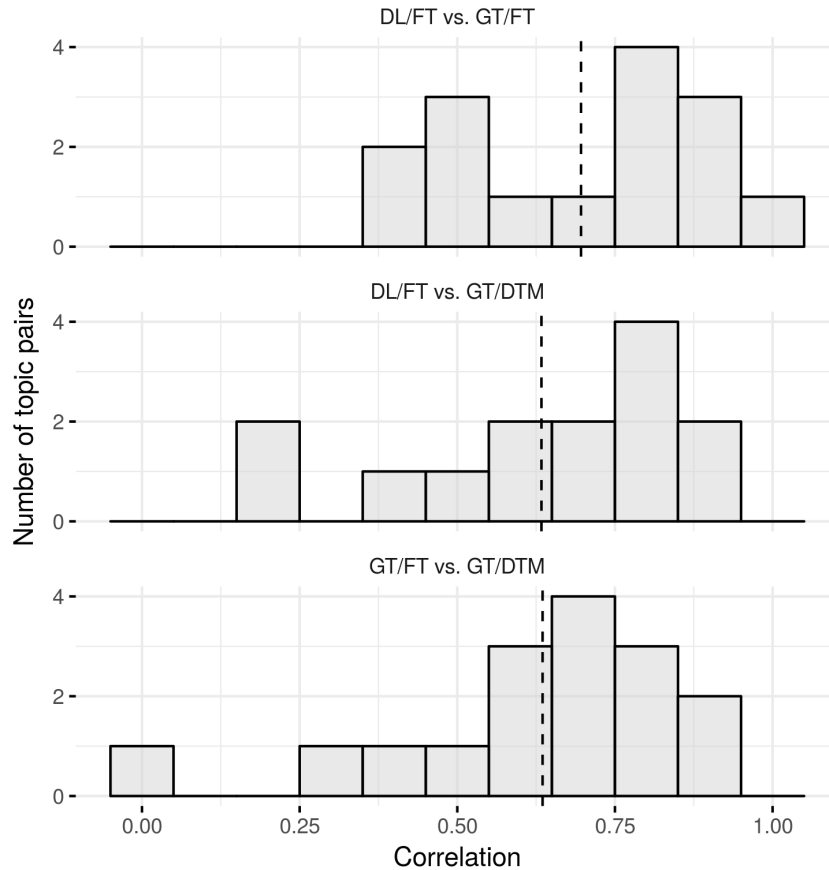
Figure 3.5: Topical prevalence correlations.

Translation method pair	N	Mean	SD	Min	Max
DL/FT vs. GL/FT	15	0.696	0.209	0.384	0.990
DL/FT vs. GL/DTM	14	0.633	0.231	0.174	0.907
GL/FT vs. GL/DTM	16	0.635	0.233	0.032	0.922

Table 3.3: Descriptive figures for topical prevalence correlations.

To sum things up, the differences between the two full-text translations are small. The accuracy of DeepL and Google Translate seems to be similar for full-text translations. Due to the larger vocabulary, DeepL may be slightly more precise than Google Translate, but it is safe to say that the choice of translation service plays a minor role. More important is the choice of the translation method, as the differences between the full-text translations and the DTM translation are bigger than between the full-text translations. However, the differences are not of a fundamental nature. Both the majority of documents as well as topics point in the same substantive direction for DTM

and full-text translations. The conclusion of Lucas et al. (2015) can therefore be confirmed that two researchers using a DTM and a full-text translation would reach the same substantive conclusions. Thus, the translation of the individual terms of a DTM can be a useful shortcut for the translation of larger corpora. Wherever possible, however, the whole texts should be translated. The smaller vocabulary of the DTM translation is a clear indicator that information is lost due to the previous preprocessing of the documents. The size of the vocabulary is also the reason, why the DeepL translation was used for the following analysis of the climate change discourse.

3.3.2 Study 2: Integrability
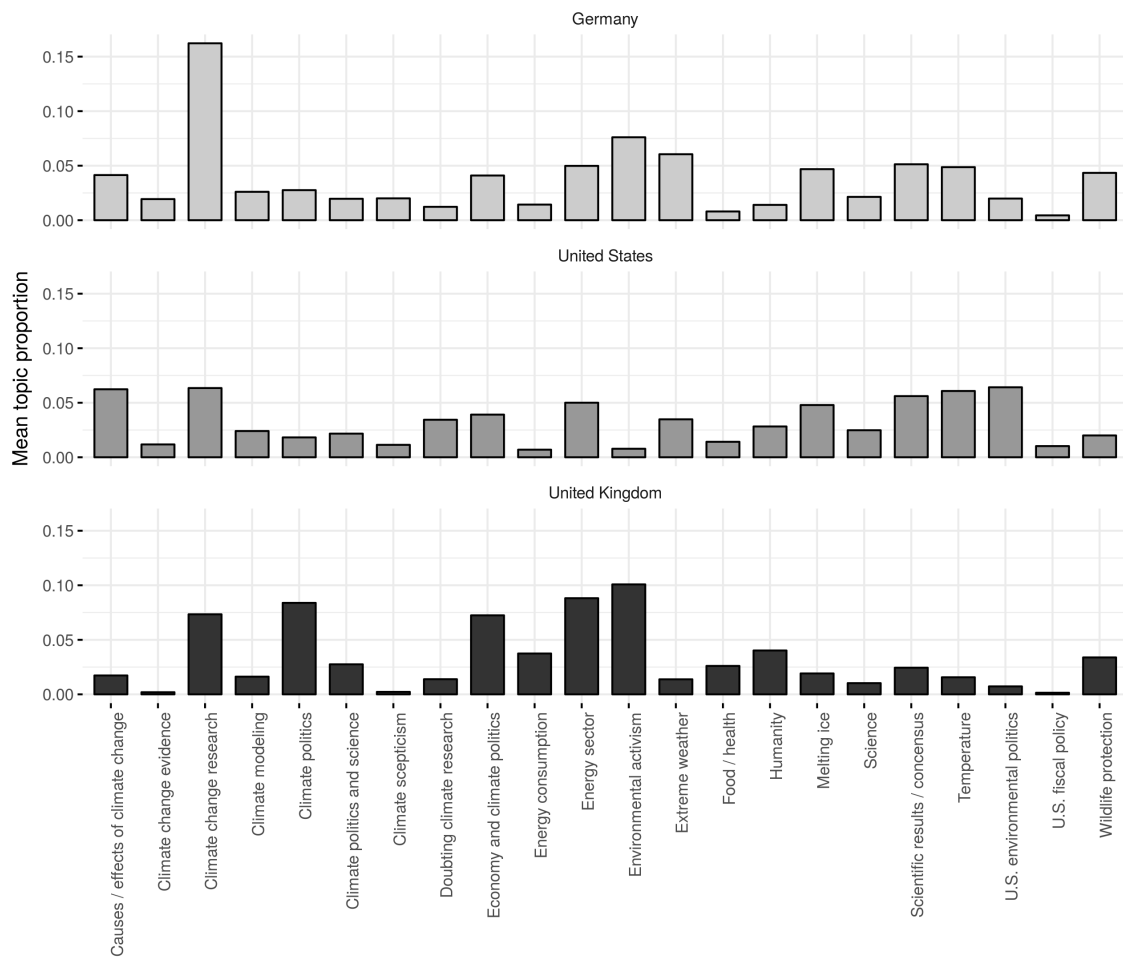
Figure 3.6: Mean topic proportion (over all actors).

The first comparison to measure whether actors in Germany, the United Kingdom, and the United States discuss climate change "under similar criteria of relevance" uses the mean topic proportions. Figure 3.6 gives an overview of the topic proportion based on the mean topic proportions of the actors. Looking at the German actors, it is noticeable that climate change research plays an important role in their communication. Not only because the corresponding topic has by far the highest probability, but also because other "scientific" topics, such as the causes and effects of climate change, the melting of ice, the temperature, or scientific results have high mean values.

Other topics frequently encountered by German actors include environmental activism as well as economic issues (especially the energy industry). U.S. American actors also give particular weight to scientific aspects of the climate change issues (causes and effects of climate change, climate change research, the temperature as well as scientific results, and the scientific consensus). Other important sub-issues are the U.S. environmental politics and the energy sector. British actors, finally, tend to place more emphasis on politics. Topics such as climate politics, economy and climate politics, the (regulation of) the energy sector as well as environmental activism have remarkably high average proportions. Climate research also plays a crucial role for British actors, albeit not a very important one.

With regard to differences between the countries, an ANOVA shows that most of the expected topic probabilities are actually different (see Appendix B, chapter 3.5.2 for a table with the test results). British actors emphasize political as well as economic aspects (climate politics, environmental activism, economy and climate politics, energy sector) significantly more than actors in Germany and the United States. Climate science, on the other hand, appears to be a rather German sub-issues (climate change research, extreme weather). Not surprisingly, U.S. environmental politics is on average more important to U.S. actors than to others.

Figure 3.7: Relative frequency of topics as top topic.

A clearer picture results if only the top topics are considered that have the highest mean proportion over all web pages of an actor. This simplification is helpful, as each document theoretically consists of each topic. Figure 3.7 shows the relative frequency o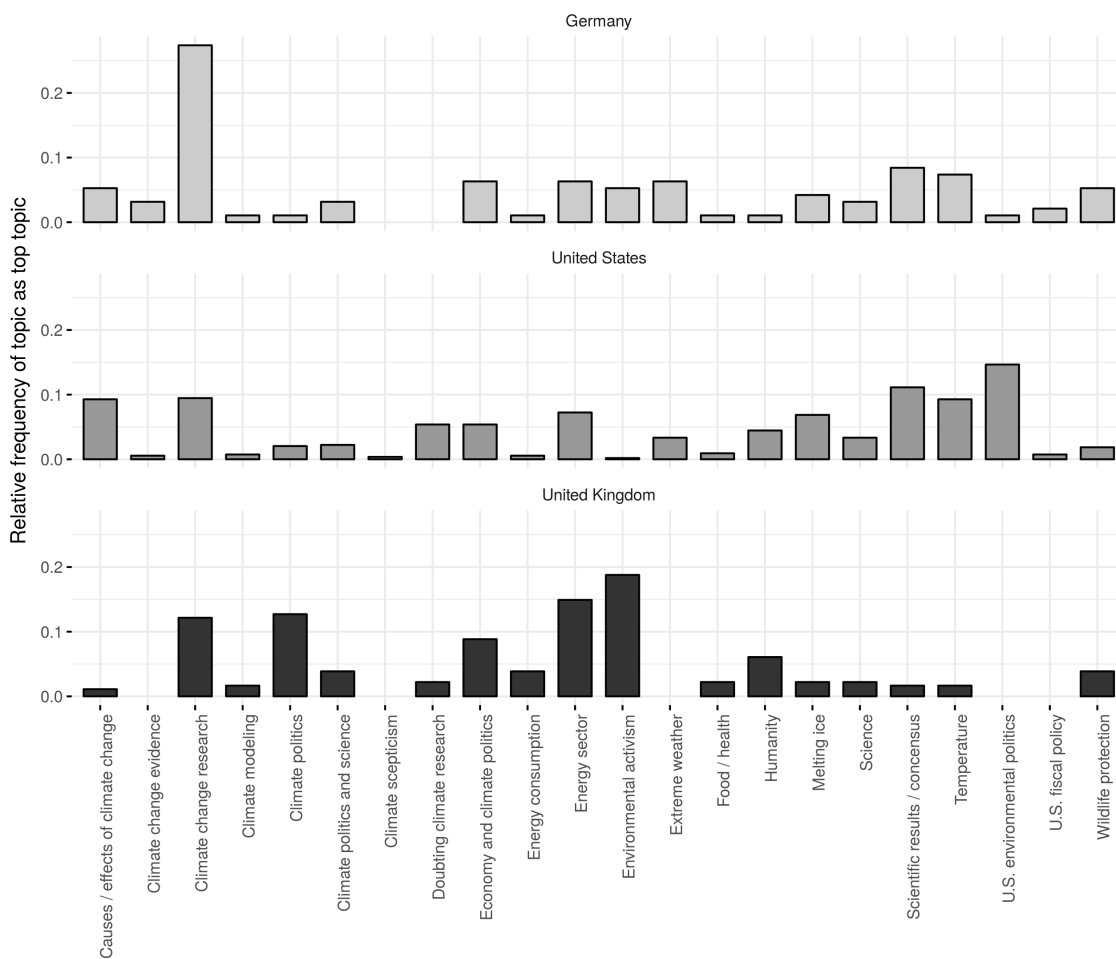f topics as the top topics for the three countries. Germany again shows the highest interest in climate change research. The impression gained before is also confirmed by the British actors, whose main concern is often political in nature (climate politics, energy sector, environmental activism, economy and climate politics, humanity). Particularly interesting is the picture that arises for U.S. actors. Clearer than in the analysis of the mean values, the top topics show a broad interest in fundamental questions on climate change. This holds true for topics such as causes and effects of climate change, scientific results and consensus as well as the temperature (trend). A substantial proportion of the actors are mainly concerned with U.S. environmental politics.

Not every topic that is typical of a country's discourse is also a unique feature. The sub-issue of economy and climate politics is relatively important in the United Kingdom, but it is not more important than in the two other countries. However, using Fisher's exact test (see Appendix C, chapter 3.5.3 for a table with all the test results), some differences in the distribution can be identified. According to the test results, climate change research can be defined as a German topic, whereas climate politics, the energy sector, and environmental activism appear to be typically British. Typical U.S. American topics are the causes and effects of climate change, the scientific consensus, and the U.S. environmental politics. Nevertheless, the public discourses about climate change in Germany, the United Kingdom, and the United States are not fundamentally different. Since there are very few topics (U.S. environmental politics, climate skepticism) that solely belong to actors of only one country, the observed differences relate mostly to the salience with which a topic is discussed, and less to the question of whether a topic is discussed at all. Actors in both countries therefore discuss climate change "under similar criteria of relevance" (Eder & Kantner, 2000, p. 315), but they do not always attribute the same importance to them. It can therefore be said that the online climate change discourse in Germany, the United Kingdom, and the Unites States partly converges, but each of them also has its clear national characteristics.

3.4 Discussion

The availability of affordable and quickly improving machine-translation services introduces new opportunities for comparative communication research. With online translation services such as Google Translate and DeepL, multilingual text corpora can be easily transformed into monolingual corpora, which can then be analyzed with, for example, topic models. In this paper, the potential of such a combined approach has been assessed, both from a methodological and substantive point of view.

From a methodological point of view, it can be said that when texts are translated for an analysis with topic models, similar results can be obtained no matter which translation service (Google

Translate vs. DeepL) or method (full-text vs. DTM) is used. Looking at the big picture—and given that the topic modeling procedure is kept unchanged—the combination of topic models and machine translation therefore appears to be quite robust. Looking closer, however, it turns out that the choice of translation service is less of a factor than the choice of method. While the full-text translations of Google Translate and DeepL are relatively similar, they differ more from the translation of the individual terms of a DTM. This shows that machine translation services, just like humans, use the context in which a word is written in a text to determine its exact meaning. Nevertheless, the comparisons have shown that the results of the DTM translation are not too far from those of the full-text translations. To translate the individual terms of a DTM is thus an acceptable and cost-effective alternative to obtain a monolingual corpus from a multilingual one. For better results, however, the first choice should always be the translation of entire documents.

To examined whether topic modeling coupled with machine translation can be linked to existing theories and work from the field of comparative communication research (i.e., the substantive point of view), the online climate change discourses in Germany, the United Kingdom and the United States of America was compared with regard to their transnationality. For the analysis, the German web pages were translated with DeepL as a whole. The results indicate that there are parallels as well as divergences between the debates in the countries. This is consistent with what is known about the climate change discourse in offline mass media. Ivanova (2017) has shown that the media agenda for climate change is similar in all three countries, but there are differences in the importance of different topics. Similarities to the findings described here are particularly apparent in the greater emphasis on scientific descriptions in Germany and the stronger accentuation of political topics in the United Kingdom (cf. Ivanova, 2017). One potential explanation for the different national priorities as well as for potential parallels between national online and offline agendas could be the national political agenda (cf. Grundmann & Scott, 2014). But it must be left to others to investigate this further.

This study showed that machine translation and topic models are a powerful couple when it comes to the analysis of multilingual corpora. Even in large amounts of text, a combined approach makes it possible to identify and compare thematic structures across language boundaries with relative ease. The cross-sectional comparison of two cases made here, is of course only one possible application. Longitudinal comparisons, as well as comparisons between several cases (i.e., languages, countries, and document types), are also possible. Moreover, content information obtained from topic models can also be combined with other information, such as network data. This could contribute to a better understanding of underlying discursive mechanisms, especially in the case of online communication. Topic models coupled with machine translation thus is a valuable addition to the toolbox of comparative communication researchers.

However, a combined use topic models and machine translation is not a cure-all. First, topic models—with or without translation—are hardly the right tool for in-depth analyses of discourses. Their strength is the detection of relatively coarse thematic structures in large text corpora, but

they are blind to more complex thematic structures, such as sentiments or arguments. Thus, if the focus is on more complex linguistic structures, other, possibly manual methods of content analysis are required. Second, topic models must be put on a solid theoretical foundation in order to be interpreted meaningfully.

Third, it must always be asked whether the analyzed texts should be sent to servers of online translation services. When public communication is studied, as here, this is usually not a problem, but when it comes to private communication, such as interviews, letters or e-mails, protecting the privacy of the people involved must be an issue. In such cases, online translation services should be avoided.

With regard to the two studies presented here, there are further limitations which must be addressed. In particular, it is not possible to determine whether the different translations providers/methods have caused systematic biases (e.g., by translating into British/U.S. American English). The validity of the translation can not be assessed with the corpus of web documents used here, as there is no benchmark available. The same applies for the accuracy of the translation. It is not certain that the larger vocabulary of the DL/FT translation actually means a more accurate translation, although a direct but unsystematic comparison of randomly selected documents from both full-text translations supports this assumption. In order to be sure, however, a comparison with a reference translation ("gold standard") is necessary (cf. De Vries et al., 2018).

Another restriction relates to the languages studied here. German to English translations are believed to be quite good in comparison to translations between other languages, as they are both Germanic languages and because there is much training data available (e.g., human-translated transcripts of parliament debates). Although Lucas et al. (2015) have reported similar findings for documents written in Chinese and Arabic, further studies have to show whether the results reported here are also valid for other languages.

It should also be borne in mind that other languages may pose other challenges—such as the ngrams in the German to English translation done here—and that the proposed procedure may therefore require some adjustments if it is used with other languages. It is also important to note again that changes in the preprocessing as well as in the modeling procedure may well affect the results (cf. Maier, Waldherr, Miltner, Wiedemann, et al., 2018).

The final restriction relates to the types of models that can be reliably computed with machine-translated texts. Possible inaccuracies caused by the machine translation are less of a concern with bag-of-words topic models due to the relatively rigid preprocessing procedure and their focus on coarse semantic structures. Whether machine translation is also useful for models that focus on finer semantic structures (e.g., arguments), and thus depend on correct syntactical structures, is uncertain. Therefore, future research should also focus on the usefulness of machine translation for linguistically more sophisticated methods.

3.5 Appendix

3.5.1 Appendix A: Webcrawler Starting Points

Starting points for the snowball-sampling of websites in Germany, the United Kingdom, and the United States.

Germany		
Advocates	Heinrich Boell Stiftung	http://klima-der-gerechtigkeit.boellblog.org
	Greenpeace Germany	http://www.greenpeace.de/themen/klima/nachrichten
	Potsdam Institute for Climate Impact Research	http://www.pik-potsdam.de/aktuelles?set_language=de
	WWF Germany	http://www.wwf.de/themen-projekte/klima-energie
Skeptics	Analyse+Aktion	http://astrologieklassisch.wordpress.com/tag/klimawandel
	EIKE - Europäisches Institut für Klima und Energie	http://www.eike-klima-energie.eu
	Klimaüberraschung	http://www.klima-ueberraschung.de
	Klimaskeptiker	http://www.klimaskeptiker.info
United Kingdom		
Advocates	Greenpeace UK	http://www.greenpeace.org.uk/climate
	Oxfam UK	http://www.oxfam.org.uk/what-we-do/issues-we-work-on/climate-change
	Friends of the Earth UK	http://www.foe.org/projects/climate-and-energy
	WWF UK	http://www.wwf.org.uk/what_we_do/tackling_climate_change
Skeptics	The Global Warming Policy Foundation	http://thegwpf.org
	Global Warming Hysteria	http://www.globalwarminghysteria.com/
	Climate Resistance	http://www.climate-resistance.org/
	Repealtheact	http://repealtheact.org.uk/
United States		
Advocates	Climate Central	http://www.climatecentral.org
	Greenpeace USA	http://www.greenpeace.org/usa/en/campaigns/global-warming-and-energy
	Worldwatch Institute	http://www.worldwatch.org/climate-energy
	WWF U.S.	http://www.worldwildlife.org/climate/index.html
Skeptics	The Heartland Institute	http://heartland.org/issues/environment
	Climate Depot	http://www.climatedepot.com
	C3 Headlines	http://www.c3headlines.com
	Watts Up With That?	http://wattsupwiththat.com

Table 3.4: Starting points for the snowball-sampling of websites.

3.5.2 Appendix B: ANOVA Results

ANOVA results for the comparison of the expected topic probabilities over all actors by country.

Topic	F(1,813)	p
Causes / effects of climate change	8.895	0.003 **
Climate change evidence	20.003	< 0.001 ***
Climate change research	17.994	< 0.001 ***
Climate modeling	5.455	0.020 *
Climate politics	89.937	< 0.001 ***
Climate politics and science	3.663	0.056
Climate scepticism	24.673	< 0.001 ***
Doubting climate research	2.093	0.148
Economy and climate politics	18.649	< 0.001 ***
Energy consumption	27.207	< 0.001 ***
Energy sector	14.213	< 0.001 ***
Environmental activism	34.808	< 0.001 ***
Extreme weather	34.937	< 0.001 ***
Food / health	12.698	< 0.001 ***
Humanity	24.584	< 0.001 ***
Melting ice	9.949	0.002 **
Science	4.301	0.038 *
Scientific results / concensus	16.189	< 0.001 ***
Temperature	13.234	< 0.001 ***
U.S. environmental politics	17.94	< 0.001 ***
U.S. fiscal policy	1.721	0.190
Wildlife protection	0.051	0.823

Table 3.5: ANOVA results.

3.5.3 Appendix C: Fisher's Exact Test Results

Fisher's exact test results for every topic as top topic by country.

Topic	p
Causes / effects of climate change	< 0.001 ***
Climate change evidence	0.029 *
Climate change research	< 0.001 ***
Climate modeling	0.463
Climate politics	< 0.001 ***
Climate politics and science	0.401
Climate scepticism	1
Doubting climate research	0.013 *
Economy and climate politics	0.232
Energy consumption	0.005 **
Energy sector	0.008 **
Environmental activism	< 0.001 ***
Extreme weather	0.003 **
Food / health	0.321
Humanity	0.144
Melting ice	0.044 *
Science	0.819
Scientific results / concensus	< 0.001 ***
Temperature	< 0.001 ***
U.S. environmental politics	< 0.001 ***
U.S. fiscal policy	0.153
Wildlife protection	0.069

Table 3.6: Fisher's Exact Test results.

Chapter 4

How climate change skeptics (try to) spread their ideas: Using computational methods to assess the resonance among skeptics' and legacy media

Abstract: We study the discursive resonance of online climate skepticism in traditional media in Germany, a country where climate skeptics lack public prestige and thus form a political counter-movement. We thereby differentiate two temporal dynamics: resonance can be continuous or selective, based on the exploitation of specific events. Beyond, we test whether such resonance is higher within the conservative media. We rely on news value theory to shed light on the mechanism facilitating or hindering such resonance and identify three indicators for resonance: frames, positions and actors. Using various computational methods as well as qualitative case studies, we examine the skeptical and traditional media discourses over a period of two years. Our analysis shows that there is no continuous resonance. However, our data reveal selective resonance: skeptics' manage to exploit specific events pushing their frames and positions onto traditional media's agenda. Thereby, conservative media did not give greater resonance to climate skeptical voices whereas they resort to downplaying the issue by allocating less space to it.

Data and code: The data underlying the results presented in the study is available to all interested researchers upon request via the open repository of GESIS (https://dx.doi.org/10.4232/1.5183). Restrictions apply because the online text data examined cannot be completely anonymized. The R code produced for the analyses is publicly available on GitHub (https://github.com/ikmb-unibe/coab_so2).

4.1   Introduction

Anthropogenic climate change is one of the most fundamental problems the world is facing, and it presents a threat to the global community (IPCC, 2014). Although there is scientific consensus that climate change is occurring, that the rise in temperature is predominantly due to human activity, that it has severe consequences for ecological systems, and that only a sharp reduction in CO2 emissions can limit climate change, political efforts have been lacking. One reason for the impasse is the counter-mobilization of climate skeptics (Dunlap & McCright, 2015), who have run campaigns that cast doubt on the scientific consensus and those responsible for it—the climate scientists. Climate skeptics and their spokespersons in the media and politics have been stunningly successful in the US, where not only conservative citizens forcefully reject climate politics (Tesler, 2017) but also the Trump administration. By contrast, in most European countries those defending the climate consensus, the climate advocates, (still) dominate political institutions (e.g., Häussler et al., 2016), traditional mass media (e.g., Schmid-Petri, 2017), and public opinion (e.g., Metag, Füchslin, & Schäfer, 2017; Tranter & Booth, 2015) with climate skeptics being in the role of counter-movements (Fraser, 1990).

Such counter-movements, however, use another venue, the internet, to push their ideas (Schäfer, 2012). Here, climate skeptics are more active and visible compared to climate advocates in European countries, building connections with conservative media and transnational allies (Adam, Häussler, et al., 2019; Elgesem, Steskal, & Diakopoulos, 2015), and even imposing their frames onto the agenda of climate advocates (Adam, Schmid-Petri, Reber, & Häussler, 2019). Yet, so far, research has hardly analyzed (for an exception, see Farrell, 2016b) whether there is any connection between climate skeptical online communication and the coverage in traditional media and politics (for this research deficit, see Adam, Häussler, et al., 2019; Sharman, 2014). The neglect of studying discursive resonance among venues is all the more surprising as research for the US convincingly shows that the discourse in the traditional channels influences public perceptions (Tesler, 2017): polarized media discourses lead to polarized public perceptions. Beyond, research so far has primarily focused on the US where climate skeptics are strong and prestigious. We, however, focus

on climate skeptics as a political counter-movement and ask: How is the discursive resonance of online climate skepticism shaped in traditional media when climate skeptics lack public prestige?

This study seeks to make three contributions. First, our study connects venues that have mostly been examined separately, as research has focused either on climate skeptics' online communication (e.g., Adam, Häussler, et al., 2019; Elgesem et al., 2015; Sharman, 2014; for a summary, see Schäfer, 2012), or on their visibility in traditional media (e.g., Boykoff & Boykoff, 2004; Grundmann & Scott, 2014; Schmid-Petri, 2017; Vu, Liu, & Tran, 2019), or on their strength in parliamentary settings (e.g., Fisher, Waggle, & Leifeld, 2013; Häussler et al., 2016). By comparing the discourses across venues, we contribute to a better understanding of the conditions that lead to greater resonance of counter-movement ideas in mainstream discourses. Second, as climate skeptics are organized across national borders (Dunlap & McCright, 2015), we take into account the potential transnational flow of ideas, via the web, into national arenas. Finally, we apply computational approaches to the analysis of the core of political contests. They allow—in contrast to manual techniques—to study discursive resonance among different venues, outlets and over longer-time periods and thus grasp different temporal dynamics of resonance as well as different patterns of resonance among outlets.

To answer our research question, we first look at climate change skeptics as political counter-movement that lacks public prestige. We then turn to the theoretical mechanisms by which these political counter-movements may resonate with traditional media and then develop indicators to measure such resonance. Before turning to the results, we elaborate on the methods employed and elaborate why we focus on Germany. The paper concludes with discussing the implications of our study.

4.1.1 Climate skeptics as political counter-movement

Following Rahmstorf (2004), climate skeptics either fully deny or cast doubt on the fact that global warming is taking place (trend skepticism), that humans are the main drivers of it (attribution skepticism), and/or that climate change is leading to severe consequences (impact skepticism). In addition, climate skeptics may also cast doubt on those producing scientific evidence (consensus skepticism, see e.g., Engels, Hüther, Schäfer, & Held, 2013) or may question the relevance of binding policy regulations (policy skepticism, see e.g., Capstick & Pidgeon, 2014; Van Rensburg, 2015). Previous research has shown that climate skepticism is often linked to certain worldviews and beliefs, such as conservatism and the support for free unregulated markets (Cook, 2016; Dunlap & McCright, 2011). Climate skepticism, however, is more than an individual attitude. Climate skeptics in the US have built what Dunlap and McCright (2011) call a "climate change denial machine" (p. 147), in which conservative politicians, media, and bloggers work hand in hand with think tanks and interest groups.

While climate skepticism has been characterized as a counter-movement (Boussalis & Coan, 2016; Brulle, 2014; Dunlap & McCright, 2015)—or counter-coalition (Adam, Häussler, et al., 2019) based on its counter-status to scientific results (Oreskes, 2004)—the political status of climate skeptics varies greatly between countries. They are far from being a minority in the US, which has been cited as a country where climate change denial has progressed to a point of becoming the official presidential doctrine; but other countries such as Norway or Australia also show strong forms of climate skepticism (e.g., Engels et al., 2013; Tranter & Booth, 2015). However, in countries like Germany, climate change skeptics are more marginal. In such countries, they are political counter-movements in the sense of Fraser (1990): that are groups in society that put forward a minority position and, as a consequence, are excluded from the mainstream debate.

Unsurprisingly perhaps, most research on climate skepticism focus on the US case (for a critique, see Engels et al., 2013)—the country in which climate skeptics are strongest and hardly excluded from mainstream debates. For the US, the climate change denial machine is well described (Brulle, 2014; Dunlap & McCright, 2015; Farrell, 2016a, 2016b), with detailed findings on climate skeptics' appearances in the media (e.g., Boykoff & Boykoff, 2004; Carmichael, Brulle, & Huxster, 2017; Schmid-Petri, Adam, Schmucki, & Häussler, 2017), in the (English-speaking) online world (e.g., Boussalis & Coan, 2016; Elgesem et al., 2015; Sharman, 2014), and in politics (e.g., Fisher et al., 2013). If we turn, however, to those countries in which climate skeptics are still political counter-movements, our knowledge is limited. While some research focused on appearances of skeptics in traditional media (often in comparison to the US; e.g., Gavin & Mashall, 2011; Grundmann & Scott, 2014; Kaiser & Rhomberg, 2015; Painter & Ashe, 2012; Schmid-Petri, 2017), only few studies analyzed the presence of climate change skeptics in parliamentary arenas (Häussler et al., 2016). Other research has focused on skeptics' online communication in these countries (Adam, Häussler, et al., 2019; Adam, Schmid-Petri, et al., 2019; Gavin & Mashall, 2011; Schäfer, 2012). Yet, to our knowledge, research has failed to study how separated the political counter-movement still is or whether we can observe discursive resonance between counter-movements debates (mostly conducted online) and mainstream discourses.

Research so far has remained primarily on a descriptive level, focusing on the degree of climate skepticism in different venues. Bringing these venues together seems all the more necessary as Fraser (1990) has already pointed out that counter-movements follow two goals: Through their internal communication they develop a group identity, whereas their external communication is directed towards the mainstream public. Investigating the connection between climate skeptics' online discourses and the wider public is all the more important as research has shown that climate skeptics fully exploit the affordance provided by the internet, regardless of their status in traditional arenas (e.g., Adam, Häussler, et al., 2019; Adam, Schmid-Petri, et al., 2019; Boussalis & Coan, 2016; Gavin & Mashall, 2011; Sharman, 2014); they can bypass journalistic gatekeepers, and directly connect with like-minded others across national borders. However, whether their strength in online communication matters for other arenas is still to be shown. This requires the study of the intersection of discourses between different venues (for this desideratum, see also Etling,

Kelly, Faris, & Palfrey, 2010; Grundmann & Scott, 2014). Theoretically, thereby three scenarios are possible: if discourses remain separated, the public is fragmented into mainstream and counter-public. If discourses resonate which each other, it might well be that counter debates take up ideas of the mainstream. With the omnipresence of the mainstream debates, such resonance is quite likely. However, what is politically more relevant and thus of utter interest to us, is the more unlikely case that political counter-movements manage to resonate within the mainstream. Research clearly shows that such actors do not have routine access to established channels. Their success therefore depends largely on their resonance in the mainstream discourse (Bennett, 1990; Kriesi, 2004; Wolfsfeld, 1997).

4.1.2 Political counter-movements and mechanisms for media resonance

To better understand the mechanisms that lead to discursive resonance we employ a dual perspective that takes into account the activities of the climate skeptical counter-movement and the working routines of media coverage. Counter-movements such as climate skeptics seek public visibility to gain a voice in political debates and affect their outcome (Fraser, 1990). Even in hybrid communication environments (Chadwick, 2013), traditional media coverage still plays a crucial role in amplifying voices of actors, their positions and viewpoints. Being at a disadvantage vis-à-vis established actors, counter-movement actors rely on discursive strategies to pursue their goals, knowing that they fit all the better into the news cycle the more they adapt to the narrative needs and working routines of the media (Gamson & Wolfsfeld, 1993).

According to news value theory, journalists evaluate events based on specific professional selection criteria—the news factors—with regard to their worthiness of publication and prominence in the coverage (e.g., Galtung & Ruge, 1965; Kepplinger, 2008; Schulz, 1976; Shoemaker, 1996). In the case of climate change, one of the strategies pursued by climate skeptics is to voice doubt about fundamental aspects such as climate science, as this ties readily to the news factor "conflict", increases the newsworthiness of the coverage and thus their chance of being included in news reports (Boykoff & Boykoff, 2004). It has already been argued that the news factor "conflict" eases the resonance of skeptic ideas within mainstream journalistic debate (Boykoff & Boykoff, 2004). Beyond this, we may argue that the news factor "surprise" pushes skeptics' ideas onto the agendas. The downside of this strategy is that it risks losing much of its newsworthiness after a while, unless climate skeptics succeed in introducing new, unexpected perspectives to the debate that emphasize the news factor "surprise". Finally, climate skeptics can attempt to rely on the prestige of some of the members of the counter-movement and thus emphasize the news factor "status/elite".

In countries where climate skeptics are still a minority movement and do not command the necessary status, getting media visibility is less likely for them. This prestige factor is all the more important as research strongly shows that legacy media have a bias towards the elites (Bennett, 1990; Wolfsfeld, 1997). With regard to climate change skeptics, we can thus ask:

RQ1: How is the discursive resonance of online climate skepticism shaped in traditional media coverage when climate change skeptics lack public prestige?

As we have seen above, news value theory suggests that the more news factors are attached to an event or issue and the stronger they are, the higher the chance that it is taken up by the media. Thereby, media coverage is driven by two different temporal dynamics. First, some news factors grant continuous media resonance. This is most likely the case for actors commanding a high level of prestige as the news factor "status/elite" describes a continuous (social and) discursive quality. Furthermore, prestigious actors on the climate skeptical side would allow the counter-movement to promote the position and their perspectives. Second and in contrast to this, counter-movements might rely on the creation of specific events—or their exploitation, to generate selective media resonance. The news factor "surprise" is clearly related to this discursive strategy. We are thus interested which types of resonance we observe in countries where climate skeptics lack prestige and thus ask:

RQ2: Is there rather a continuous or a selective congruence between the online communication of climate skeptics and traditional media coverage?

Finally, researchers have started to question whether news factors are perceived similarly among different news outlets (e.g., Kepplinger, 2008). This so-called "two-component" theory assumes that news factors as characteristics of events are perceived differently by journalists depending on the political orientation of the outlets. In the US for instance conservative media ascribe greater news value to the ideas of climate skeptics (Boussalis & Coan, 2016; Dunlap & McCright, 2011; Feldman, Maibach, Roser-Renouf, & Leiserowitz, 2012; Painter & Ashe, 2012). This conservative alliance structure has also been shown for countries in which counter-movements still have a minority status. Here, specific conservative newspapers publish skeptics' claims (Schmid-Petri, 2017), and skeptics closely connect via hyperlinks to these right-wing media outlets (Adam, Häussler, et al., 2019). Consequently, our study examines the following question:

RQ3: Is there a special connection with regard to the discursive resonance between climate skeptics and conservative media?

4.1.3   Political counter-movements and indicators for media resonance

To determine the degree of congruence between the online communication of the climate skeptical counter-movement and traditional media coverage, and how it develops over time, we distinguish three different discursive dimensions (Pfetsch et al., 2013): (1) issues and frames, (2) the positions articulated, and (3) the visibility of actors. The more we see frames, positions and actors converge between the two venues, the more we can speak of a discursive resonance. If we see that skeptic frames, positions and actors resonate with the mainstream, the plausibility is high that the counter-movement has succeeded in spreading their ideas; whereas if frames, positions and actors from the mainstream turn more prominent in skeptics' discourse, the opposite is the case. However,

the methods employed in this paper, do not allow us to draw any direct conclusion as to causal direction.

Skeptical issues and frames: Following Fraser (1990), counter-movements are excluded from mainstream debate. Consequently, one of their central goals is to make their issues and frames visible to a wider public. From a classical agenda-building perspective (Cobb, Ross, & Ross, 1976) this means that counter-movements seek to raise awareness of those issues that are important for them and try to frame debates from their viewpoint. In well-established issues, where the agenda is largely set by traditional media and political institutions, counter-movements primarily seek to re-frame the debate by promoting alternative views on an issue such as climate change. They attempt to shift "central organizing idea[s] or story line[s] that provide [...] meaning to an unfolding strip of events [...] The frame suggests what the controversy is about, the essence of the issue" (Gamson & Modigliani, 1987, p. 143).

On the thematic level, climate skeptics do this by sowing doubt where there is (scientific) consensus, with the aim to draw climate advocates into a debate where contrarian positions might gain the upper hand (Lewandowski, Oreskes, Risbey, Newell, & Smithson, 2015). Studies for the US (Farrell, 2016b) have shown that a positive semantic relation between climate skeptic's online communication and traditional media coverage may occur, documenting the responsiveness of the media to contrarian ideas and thus the success of this strategy. In this perspective, an increase of skeptics' frames on the media agenda might be taken as an indicator for discursive resonance independent of the fact, whether media counter-argue these frames or not. Pure visibility of frames matters (Ellinas, 2010).

Skeptical positions: Frames are only part of the discursive structure of a debate; equally important is the question whose position is being covered by the media. After all, climate skeptical frames can just as well be reported from a critical viewpoint that effectively undermines their credibility. When for instance climate skeptics succeed in provoking a debate about the uncertainty of scientific results (which is a classic skeptics' frame, see Adam, Schmid-Petri, et al., 2019), mass media may decide to give voice to mainstream scientists who contest this frame. Similarly, journalists might not follow climate skeptical framing, but include their counter-position.

Climate skeptical participants in debates: In addition to frames and positions, counter-movements attempt to promote specific representatives of their cause, in an attempt to expand the range of legitimate participants in the debate. These actors gain visibility in two ways, either as speakers or addressees. In the first case, they may be spokespersons who relate the view of a skeptical think tank. This form of visibility is closely associated with the positional dimension introduced above. In such a setting, journalists use the counter-movement's representatives as "opportune witnesses" (Hagen, 1993) to make climate change skepticism more prominent. Second, climate skeptics may gain visibility as objects of reporting. In this role, they are ratified by other actors positively or negatively. Yet, in both instances, the counter-movement succeeds in gaining visibility in mainstream debates, which helps turn it into a legitimate actor.

Table 4.1 shows the different dimensions of climate skeptical discourse. To simplify the analytical framework, we distinguish four core types according to the thematic and positional dimensions. We use the actor dimension to further specify the typology.

		Frame resonance	
		Yes	No
Positional resonance	Yes	**Full resonance:** Counter-movement frames and positions become more central in the mainstream coverage	**Positional resonance:** counter-movement positions become more central in the mainstream coverage
	No	**Frame resonance:** Counter-movement frames become more central in the mainstream coverage	**No resonance:** Counter-movement discourse remains marginal in the mainstream coverage

Table 4.1: Types of discursive resonance.

If frames and positions of counter-movements are mirrored in the coverage of mainstream media, we speak of full resonance. Pure positional resonance occurs if only the positions of counter-movements are reported, yet within the mainstream discourse. For climate skeptics, this would, for example, mean that their skeptical positions receive attention within the larger mainstream debate about the role of renewable energy in the transition from fossil fuels to a greener society. This indicator looks for the pure attention of skeptical positions without taking into account whether journalists counter-argue. However, research points out that visibility matters independent of the evaluations surrounding it (Ellinas, 2010). By contrast, we speak of pure thematic congruence if the frames promoted by counter-movements become more prominently part of the relevance structure of the mainstream debate around climate change without, however, finding a parallel increase in their positions. Such a setting occurs if climate skeptical frames such as the credibility of scientific studies, receive attention in mainstream debates, though the skeptical position associated with it is dismissed. The media would cover the perspective promoted by climate skeptics but explicitly reject it, arguing for instance that climate science is credible, and scientific results confirm the trends established by previous research. Finally, if neither skeptical counter-movement frames nor positions resonate in the media, the counter-movement discourse remains segregated from the mainstream (Fraser, 1990).

Full, positional and thematic resonance can be further classified according to whether skeptical actors gain visibility in the coverage and thus are ratified as legitimate participants in the debate. The more climate skeptical actors become visible together with their position and/or their frames, the more they succeed in becoming part of the debate on their own terms. Conversely, the less climate skeptical actors are mentioned in the coverage together with frames and/or positions, the more the debate is shaped by journalist and other actors such as mainstream politicians.

4.2 Case selection, data, and methods

4.2.1 Case selection

To study the discursive resonance between the (potentially transnational) online communication of climate skeptics and traditional media, we focus on Germany. Germany represents an ideal case as climate skepticism was a small but significant phenomenon in traditional media (Lörcher & Taddicken, 2017), parliamentary arenas (Häussler et al., 2016; Schäfer, 2016), and public opinion (Engels et al., 2013; Tranter & Booth, 2015) during the period of analysis (June 2012 – June 2014). However, this wide-spread acceptance of man-made climate change does not mean that there was an equally high level of agreement on political measures. Anyhow, our design allows to search for initial resonance increase among venues. Thereby, research for Germany also shows that climate skeptics successfully exploit the online world; although fewer in number, climate skeptics are more visible and more active online and more strongly connect to transnational allies compared to climate advocates (Adam, Häussler, et al., 2019).

In such a setting, it is possible to observe whether and how ideas from the political counter-movement may flow into mainstream debates, whereas such flows of ideas are hard to detect in those countries where climate skeptics are not a minority movement anymore as their ideas are already visible in all channels. The German setting, thus, may allow us to understand the spread of ideas, respectively, the conditions under which policy monopolies are destructed (Baumgartner & Jones, 1993) by redefining the issue, by changing the positions formulated, or by changing the participants of the debate. This is even more interesting as some researchers have claimed that climate skeptics seem to be gaining attention and terrain in Germany: Brunnengräber (2018) posited that climate skepticism is increasingly gaining societal acceptance, and Schmid-Petri and Arlt (2016) show that climate skeptical arguments have slightly increased in mainstream German media over recent years.

4.2.2 Collecting data on the counter-movements' communication

To study skeptics' communication, we studied their online communication—the field in which they are most active. Hereby, we relied on hyperlink issue networks that originate from prominent counter-movement actors. We thereby followed the logic of snowball sampling—a method employed if researchers have limited knowledge about the overall population. Such method allows to detect also those skeptics that are less-known.

To identify relevant skeptics, we relied on a six-step procedure shown in Fig 4.1. First, we selected the four most important civil society actors of the climate skeptical counter-movement in Germany as starting points based on expert interviews, literature reviews, and country-specific Google searches (with deleted search histories). These are Analyse+Aktion, EIKE—Europäisches Institut für Klima und Energie, Klimaskeptiker, and Klimaüberraschung. We chose civil society

actors as they show the broadest linking behavior (Rogers & Marres, 2000) and because they are the "champions of online communication" (Schäfer, 2012, p. 530). Second, starting from the actors' main climate pages (list of URLs in S1 Appendix, chapter 4.5.1), crawling software (called Issuecrawler; see Rogers, 2013) collected all hyperlinks two levels deep within the websites and all of those that pointed to other websites. We limited our snowball crawling to go only one step "out" as pre-studies have shown that further crawling substantially increases the number of pages that do not deal with climate change. Third, to make sure that only pages that were relevant to the climate debate remained in our network, we indexed all pages according to our keywords (i.e., "Klimawandel," "globale Erwärmung," "globaler Erwärmung," "globalen Erwärmung"). We only indexed content that was publicly available, i.e. not password-protected. Also, we respected the robots exclusion standard (robots.txt). This standard allows website owners to define areas of their website that should not be scanned or indexed by robots (e.g. search engines, web crawlers). In this way, we made sure that the content was permitted for download by the website owners.

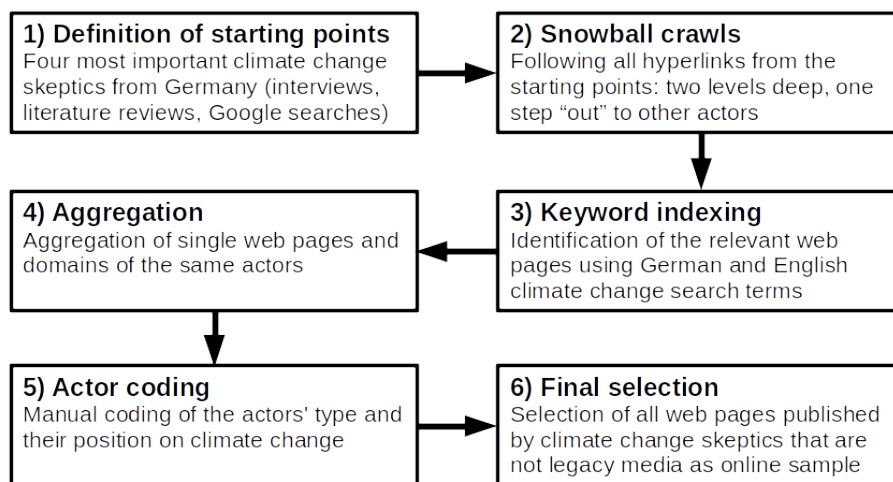| 1) Definition of starting points
Four most important climate change skeptics from Germany (interviews, literature reviews, Google searches) | 2) Snowball crawls
Following all hyperlinks from the starting points: two levels deep, one step "out" to other actors |
| 4) Aggregation
Aggregation of single web pages and domains of the same actors | 3) Keyword indexing
Identification of the relevant web pages using German and English climate change search terms |
| 5) Actor coding
Manual coding of the actors' type and their position on climate change | 6) Final selection
Selection of all web pages published by climate change skeptics that are not legacy media as online sample |

Figure 4.1: Online sampling procedure.

Fourth, some data preparation was necessary. We aggregated single web pages and domains that belonged to the same actors. This allowed us, in a fifth step, to attribute a position on climate change and a type of actor to each of the identified actors. To do so, we applied a manual content analysis, based on information found on the "About us" pages or similar sections of the website and conducted by two trained coders. They distinguished, position-wise, between climate skeptics, climate advocates, and climate neutrals (with no clear-cut position); separated legacy media from other actors; and coded their scope/country of activity (distinguishing a total of 199 geographical areas). Note that according to our definition, a climate skeptic is someone who explicitly questions at least one of the following: the existence of climate change, the human contribution to it, the science of climate change and/or its findings, projected trends/consequences of climate change, and/or the adaptation to it. By this definition, a climate skeptic is also someone who endorses

the science but negates, for instance, its political and economic consequences. In turn, a climate advocate explicitly supports at least one of these points without doubting any of the others. The actors' attributes were classified by two trained coders. The reliability of their classification was assessed by comparing them with a master coding and was calculated using Krippendorff's alpha, which is a common statistical measure of the agreement between different coders when coding the same texts. The agreement is usually measured on the level of individual variables. The test revealed satisfactory results with Krippendorff's Alpha of .90 for the position variable, .90 for the actor type variable, and .93 for the scope variable. In the final step, only those actors and their communication that put forward climate skeptical positions and were not rated as legacy media remained in our sample, which serves as a proxy for climate skeptics' communication (154 different actors with a total of 13,009 unique web pages). Climate skeptics' online communication here originates from prominent German climate skeptics from civil society, but then it is extended by snowball sampling to include text material from all types of skeptical non-legacy-media actors. As shown in Table 4.2, however, non-profit civil society actors and individual private persons account for the vast majority of the web pages in the final sample. Out of the 154 different actors, 33 are domestic actors from Germany with a primarily national scope, and 121 are foreign actors (either with a foreign national or transnational scope).

		Actors	Webpages
Domicile	German actors	33	3,549
	Foreign actors	121	9,490
Type	Politicians, political actors	3	3
	Socioeconomic pressure groups, companies	4	33
	Non-profit civil society actors	45	3,779
	Genuine online media	26	216
	Citizens / private persons	74	8,977
	Other actors	2	12
Total		154	13,009

Table 4.2: Key figures of the skeptics' online sample.

We collected such online communication data of climate skeptics once a month over a period of two years (2012–14). We are aware that this proxy captures only part of climate skeptics' communication, omitting, for example, online communication via social media (e.g., Gerlitz & Helmond, 2013). However, as actors are active on all venues and as research has shown that hyperlinking, i.e. referring to another actor by linking to its web presence, is closely related to social media and the interactions there (Fu & Shumate, 2017), we are confident that such a partial approach might capture the relevant content of skeptics' online communication (see for a similar argument, Benkler, Roberts, Faris, Solow-Niederman, & Etling, 2015).

4.2.3 Collecting data on mainstream debates in traditional media

We regard the national print media landscape (whether online or offline) as a good proxy for the mainstream debate. Research has shown that mass media have an elitist focus (Bennett, 1990; Wolfsfeld, 1997) and, as such, are likely to reflect the mainstream debate. For Germany, we have selected the 15 most important daily and weekly newspapers as well as magazines with a national audience reach (see S2 Appendix, chapter 4.5.2). All of these outlets are regarded as national opinion-leaders. Within these outlets, we identified all relevant articles on climate change by searching for the above-mentioned keywords in the databases of Factiva and LexisNexis. This resulted in 4,111 articles about climate change in the observed period in German legacy media.

To find out whether discourse resonance is especially strong as regards the conservative media landscape, we finally classified the newspapers and magazines in our sample according to their ideological position. Following the work of Beck (2018, pp. 153-155), Begenat (2016, pp. 98-99), Lüter (2004), Maurer and Reinemann (2006, pp. 129-130), Pew Research Center (2018), Schwarz-Friesel (2013, p. 52), and Wessler and Rinke (2013, p. 640), we identified the following nine legacy media as right of the center: *Bild, Bild am Sonntag, Die Welt, Frankfurter Allgemeine Zeitung, FAZ am Sonntag, Financial Times Deutschland, Focus, Handelsblatt, and Welt am Sonntag.* As shown in Table 4.3, they account for 1,495 articles in our sample (see S2 Appendix, chapter 4.5.2 for exact number for each outlet). However, it is important to note that unlike in other countries (e.g., Great Britain), there are no hard right-wing newspapers in Germany, as even the most pronounced right-leaning paper (i.e., *Die Welt*) mostly adheres to liberal-pluralist principles (Beck, 2018).

		Outlets	Articles
Type	Conservative media	9	1,495
	Other media	6	2,616
Total		15	4,111

Table 4.3: Key figures of the offline sample.

The raw data collected from legacy media as well as from the counter-movements' online communication is available to all interested researchers upon request via the open repository of GESIS (https://dx.doi.org/10.4232/1.5183). The R code produced for the data collection as well as for the analyses is publicly available on GitHub (https://github.com/ikmb-unibe/coab_so2).

4.2.4 Measuring discursive resonance

Our analysis is based on digital text on climate change of the skeptical counter-movement (online) and of legacy media (offline) over the course of two years. We aggregated the data on the level of single months and examined the discursive resonance. We acknowledge that this approach only captures the monthly correlation of the agendas and we therefore make no claims as to any strict causal relationships between counter- and mainstream-arenas. However, a more detailed look reveals whether discursive resonance pushes skeptics' frames, positions and actors or those of

the mainstream. Of course, such patterns of increased discursive resonance may well be caused by third, unobserved factors, which play an important role in creating opportunities of discursive resonance. Such factors are hard to find in the continuous analysis whereas it is one strength of our qualitative case studies that such factors are detected.

To reveal the general trends, we employ a correlation approach on the aggregate level of the data, examining discursive resonance between counter-movement discourse and media coverage month by month. For single critical instances of heightened resonance, in turn, we take a more qualitative close-up view that is able to reveal some of the reasons behind the sudden changes Table 4.4 provides an overview of the analytical techniques applied.

Type of resonance	Approach	Discursive indicator	Method	Measure
Continuous	Quantitative	Frames	Topic model (STM, unsupervised): identification of frames used in documents	Jensen-Shannon divergence (ordinary least squares regression)
		Positions	Classifier (SVM, supervised): categorization of sentence as advocate, skeptical, or irrelevant	Difference of skeptical sentences rates on-/offline (ordinary least squares regression)
		Participants	Named-entity recognition (supervised/manual content analysis): identification of (non-)skeptical actors	Difference of skeptical actors rates on-/offline (ordinary least squares regression)
Selective	Qualitative	Frames, Positions, Participants	Identification of critical moments based on quantitative indicators, followed by an in-depth study of the factors triggering such discursive resonance.	

Table 4.4: Methods and techniques used to measure discursive resonance.

To conduct this analysis, we employed a variety of computational methods. We used a bag-of-words *topic model* to detect shifts in the framing competition, employed a purpose-trained *classifier algorithm* to distinguish positions on climate change, and relied on a semi-automated *named entity recognition* procedure to show which actors participated in the debate. In the following paragraphs, we will shortly introduce each of these methods.

To identify *frames*, we relied on probabilistic *topic modeling* (Blei, 2012), an unsupervised method of automated content analysis (Grimmer & Stewart, 2013), that helps unravel latent or hidden thematic structures of text material using Bayesian statistics. Topic models are mixed-membership models, meaning that every document consists of a mixture of different topics. Each document can therefore be understood as probability distribution over a set of topics and is best described by those topics with the highest probabilities. A topic, on the other hand, is defined by a "probability distribution over the entire corpus' vocabulary" (Maier, Waldherr, Miltner, Wiedemann, et al., 2018, p. 97) and represents a latent pattern of word (co-)occurrences. Those words that have a high probability within a topic are the ones that define it thematically.

In order to become meaningful, topics must be interpreted in a qualitative process and against the backdrop of a theory. Depending on both the corpus and the theory, the resulting thematic structures (i.e., probability distribution over words) may either be interpreted as actual topics (e.g., environment versus economy), as issues (e.g., climate change), or as frames (e.g., the emphasis on the scientific consensus), depending on the underlying type of text corpus (Jacobi et al., 2016). As explained before, we understand frames as "central organizing idea[s] or story line[s]" (Gamson & Modigliani, 1987, p. 143) that represent the "particular ways [in which] issues are presented" (Price & Tewksbury, 1997, p. 184). In combination with the fact that we have keyword-cleaned texts, all dealing with climate change, and thus a relatively coherent corpus, we consider it valid to interpret the resulting latent patterns of word (co-)occurrences as frames or interpretative packages (see also DiMaggio, Nag, & Blei, 2013).

For our analysis, we rely on the commonly used and well documented structural topic model framework (STM; Roberts et al., 2016, 2013). As we removed duplicate web pages to compute our STM, it is based on a total of 17,120 text documents. This was done to avoid a bias in favour of the duplicated documents. As topic models only work with monolingual text material, we translated all the vocabulary of the English web pages into German before calculating the model (Reber, 2019). We applied several common preprocessing steps in order to extract as much information from the corpus as possible. This included the removal of punctuation, conversion to all lowercase, removal of words with less than three characters, removal of stop words (e.g., "und", "oder", "auf", "der"), stemming, and the removal of words that appear in less than 0.5% and in more than 99% of all documents (relative pruning). To decide on the number of topics (K), we combined data-driven indicators with a qualitative assessment of the interpretability of different solutions (Maier, Waldherr, Miltner, Wiedemann, et al., 2018). To do so, we first calculated 8 models with 5 to 40 topics. We then compared these models based on standard measures (i.e., held-out likelihood, semantic coherence, residuals). The models with 20 and 30 topics were selected for the final interpretation step. Based on both the topic top words and particularly relevant documents, the topics were interpreted by two people. The guiding question was whether top words and documents represent an interpretable and coherent frame. If both persons came to the same conclusion, the topic was labelled, otherwise it was excluded from the analysis. In the end, we chose the model with 20 topics, whereby five uninterpretable topics were excluded and two similar topics were merged (all labels and top words are shown in S3 Appendix, chapter 4.5.3). The calculation was done in R using the *stm* package (Roberts et al., 2017).

To compare the similarity of frames used by climate skeptics in their online communication and in mass media reporting, we relied on the Jensen-Shannon divergence (JSD). This is a smoothed and symmetric derivative of the Kullback-Leibler (KL) divergence, which is a common measure when comparing distributions (Niekler & Jähnichen, 2012). The normalized outcomes of the JSD can be used as measure of similarity between two probability distributions and is therefore well suited for the comparison of the topic distributions of our online and offline samples. A JSD of 0 would indicate complete congruence of two distributions (i.e., the same frames used by the

climate change skeptics and the legacy media). A JSD of 1 would mean completely different distributions and thus completely different frames. If climate skeptics succeed in influencing the thematic relevance structure of legacy media in a continuous way, we would therefore expect a declining JSD over the course of our two-year period of analysis. We use ordinary least square regression to test whether the potential trend is significant.

To identify *positions*, we relied on a trained *classifier algorithm*, or, more precisely, on a linear support vector machine (Joachims, 1998). This classifier algorithm was used to categorize each single sentence in our documents as either advocate-leaning, skeptical-leaning, or irrelevant. In contrast to topic modeling, a classifier follows the logic of automated supervised content analyses (Grimmer & Stewart, 2013). This means that it follows the logic of a pre-defined coding scheme. In a first step, this coding scheme guides a manual content analysis of a text sample, distinguishing skeptical from advocative sentences. In a second step, the manually coded material serves as learning material for the computer algorithm (Scharkow, 2013). To account for the two languages in our data set we trained two models, one for English documents and one for German documents.

To train the models, we used an active learning scenario. This means that we trained and checked the two models in several iterations, using manually coded sentences as training material and benchmark. The initial training set consisted of sentences from advocates and skeptics as well as of sentences that have nothing to do with climate change. The inclusion of such random sentences is crucial in order to be able to detect also irrelevant sentences in the data set. This initial training sets were used to train a first model for each language that was then applied to 10'000 random uncategorized sentences of our corpus in the respective language. After this first classification, the result was evaluated by a team of three human coders. The evaluated sentences were then added to the initial training data and the models were trained once again (hyperparameter C optimized by cross validation to avoid overfitting). We repeated this process three times, until we could not improve the classifiers' performance anymore (measured by k-fold cross validation). To measure the accuracy of the classifiers, we treated the manually coded sentences as gold standard and compared them with the machine coded sentences. After the third iteration, the overall accuracy measured by the harmonic mean of precision and recall (micro-average) was $F1 = 0.83$ for the English model and $F1 = 0.85$ for the German model (see S4 Appendix, chapter 4.5.4 for macro-average $F1$ of different categories). These are satisfying accuracy values. Classification was done in R using the $LibLineaR$ package (Fan, Chang, Hsieh, Wang, & Lin, 2008).

To assess the discursive resonance as regards positions, we analyzed whether the increase in skeptical sentences online is correlated with the share of skeptical sentences in the media. To do so, we first compare the share of skeptical sentences online and offline and study whether the differences of these shares increases or decreases. To detect potential trends, we use linear regression models.

Finally, to identify *skeptical actors* within the debate, we relied on *named-entity recognition* (NER). It is a set of procedures to extract categories, like people's names, organizations, and locations,

from unstructured texts (for a short description, see Scharkow, 2013). To identify all actors in our sample of climate change web pages and media articles, we first used a list of around 1.3 million known named-entities as a look-up list (a lexical approach). On the one hand, the list consisted of names of prominent individuals (e.g., "Angela Merkel") and multi-word units (e.g., "market economy") that were not necessarily related to the climate change issue. On the other hand, it contained 10,095 names of actors that we had identified as important for the public discourse on climate change using manual content analysis (Schmid-Petri et al., 2013). In order to identify actors who were not on the list, we used two probabilistic sequence classifiers (English and German), which we trained specifically for this purpose using the list. We used conditional random field models from the *Stanford CoreNLP* package (Finkel et al., 2005) as classifier algorithms.

To show discursive resonance as regards debates' participants, we needed to add information to the named entities extracted. Thus, we relied on the list of 10,095 manually coded actors. For each actor, we coded the position on climate change—more precisely, whether the actor thinks that climate change is occurring and whether he/she sees it as a problem. This allowed us to identify the most important climate skeptics in our corpus. Using Kripendorff's alpha again, the reliability scores were 0.69 for the first variable (occurrence of climate change) and 0.75 for the second (climate change seen as problem). This was measured as master-coder reliability based on a random sample of 30 actors. Both manually coded variables achieved satisfactory reliability scores.

Using these procedures, we identified a total of 46,901 skeptical actors, 443,452 advocative actors, and 411,955 actors without a clear position on climate change in our corpus. As before, we analyzed whether an increase in mentions of skeptical actors online was associated with an increase in mentions in the media. Again, we first compare the shares of skeptical actors used by both the skeptics and the legacy media. We then use the difference of the shares as divergence measure. Simple linear regression models are used to check for a significant convergence or divergence respectively.

Finally, we used a qualitative approach to identify selective resonance. In a first step, we identified critical moments in the time series. These are moments when the frames found online and offline were more similar than usual, when there were an increase of skeptical sentences in the media coverage, or when an unusual number of skeptical actors were mentioned in offline reporting. We then searched for documents that showed the characteristic patterns. A thorough reading of these documents finally allowed us to identify and describe the factors for discursive resonance and put them into the wider context of legacy media coverage on climate change.

4.3   Results

As a basis for our further analysis of the three indicators for discursive resonance, we first looked at the salience of climate change in legacy media (Fig 4.2 A)—a prerequisite to study the resonance

of different venues in this issue. Fig 4.2 A clearly shows that legacy media reported frequently on climate change. However, thereby two observations are worth mentioning: first, conservative media are responsible for only a third of the articles in our sample (although they constitute more than half of our sample). The other two thirds of the articles on climate change are published by left-leaning media and media without a clear political profile (hereinafter referred to as "other media"). Second, the volume of reporting on climate change decreased in both media categories over the course of the two years analyzed.
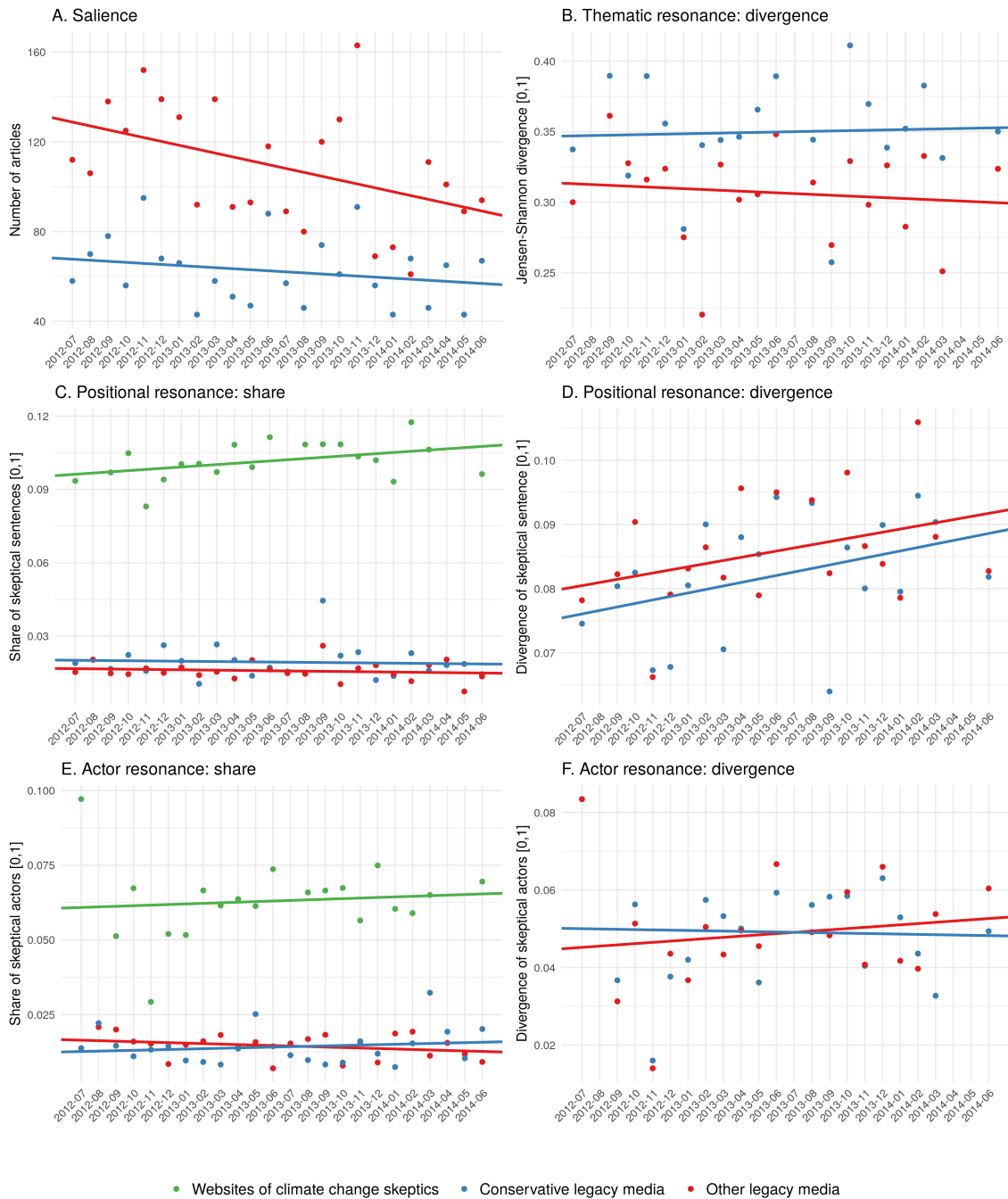
Figure 4.2: Salience of climate change (A) and resonance measures (B-F).

4.3.1 Continuous resonance

Frame resonance: We speak of a continuous frame resonance if the JSD becomes significantly smaller in the two-year period. This would indicate a convergence of online and offline agendas. However, the regression models reported in Table 4.5 show that there was no significant decrease of the JSD for both conservative and other legacy media. As the coefficients indicate, the distance between online and offline agendas remain almost the same over the whole time period. Therefore, neither of the two legacy media types can be attributed a continuous thematic resonance.

	Thematic resonance		Positional resonance		Actor resonance	
Media	Conservative	Other	Conservative	Other	Conservative	Other
Constant	0.3467	0.3137	0.0750	0.0795	0.0502	0.0446
	(0.0176)	(0.0162)	(0.0044)	(0.0042)	(0.0073)	(0.0074)
Time	0.0003	-0.0006	0.0005 .	0.0005	-0.0001	0.0003
	(0.0013)	(0.0012)	(0.0003)	(0.0003)	(0.0005)	(0.0005)

Table 4.5: Regression coefficients β for resonance measures.

Standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

Beyond, a comparison between conservative and other media reveals that there is no indication that conservative media are more open for the skeptics' frames. In contrast Fig 4.2 B even indicates that the Jensen-Shannon divergence (JSD) between the aggregated online and offline frame agendas is slightly bigger for conservative media than for other media. This means that the climate change framing in left-leaning media and media without a clear political profile is, overall, closer to that of the climate change skeptics on the web compared to that of conservative newspapers and magazines which clearly contradicts the expectations about the special role of right-wing media in giving voice to climate skeptics as regards frames.

Positional resonance: With regard to positional resonance, Fig 4.2 C shows that climate change skeptics became more radical over time. However, the degree of verbal radicalization is marginal ($\beta = 0.0005$) and only weakly significant ($p = 0.0687$). Nevertheless, every tenth sentence on an average skeptic's website explicitly expresses a skeptical position on climate change. The traditional media, however, remain unaffected by this development. The proportion of skeptical sentences is almost identical for both the legacy media categories and remains low over the entire time period. Accordingly, it is no surprise that the divergence (Fig 4.2 D) between online and offline increases. In case of the conservative media, the increase ($\beta = 0.0005$, $se = 0.0003$) is even weakly significant ($p = 0.0904$). Despite their radicalization, the climate change skeptics have therefore not succeeded in provoking a continuous resonance regards their positions.

Resonance as regards participants of a debate: As shown in Fig 4.2 E and 4.2 F, no resonance can be shown as regards participants. Legacy media hardly give skeptical actors a platform, regardless of the media's political profile. Moreover, there is also no visible convergence. The difference

remains similar for both conservative and all other media over the whole two years. Thus, there are no obvious continuous resonance effects here either.

Overall, no continuous discursive resonance can be identified. There is no evidence that German legacy media are increasingly following climate skeptics over the course of time. Interestingly, hardly any difference is found between conservative and other legacy media as regards frames, positions and actors, with some counter-intuitive exceptions as regards the framing of the issue. For the German conservative media, we can therefore not confirm what has been shown for their US counterparts—that conservative media in general are allies to climate change skeptics. German legacy media do not seem to offer a platform for climate change skeptics.

The lack of continuous resonance between legacy media and climate skeptics in Germany as regards frames, positions and actors, results in a debate that is fragmented. On a thematic dimension, this fragmentation becomes visible in Fig 4.3. It makes clear that German legacy media frame the issue of climate change almost exclusively in terms of the German economy/consumption patterns. This one frame captures most of the mainstream debate, whereas climate skeptics employ a variety of frames with "doubting the climate science consensus," "climate science skepticism," and "measuring climate change" being the most prominent. On a positional dimension, this fragmentation is shown in the radicalization of climate skeptics' communication that is not at all reflected in mainstream debate (Fig 4.2 C, 4.2 D). In sum, climate skeptics still form a separated counter public in Germany in the time of analysis, that is occupied with its own frames, radicalizes its positions and is not taking-over the economy frame dominating the mainstream debate. The mainstream public, in contrast, is occupied with discussing the economic consequences of climate change while largely ignoring the skeptics' frames, positions and actors.
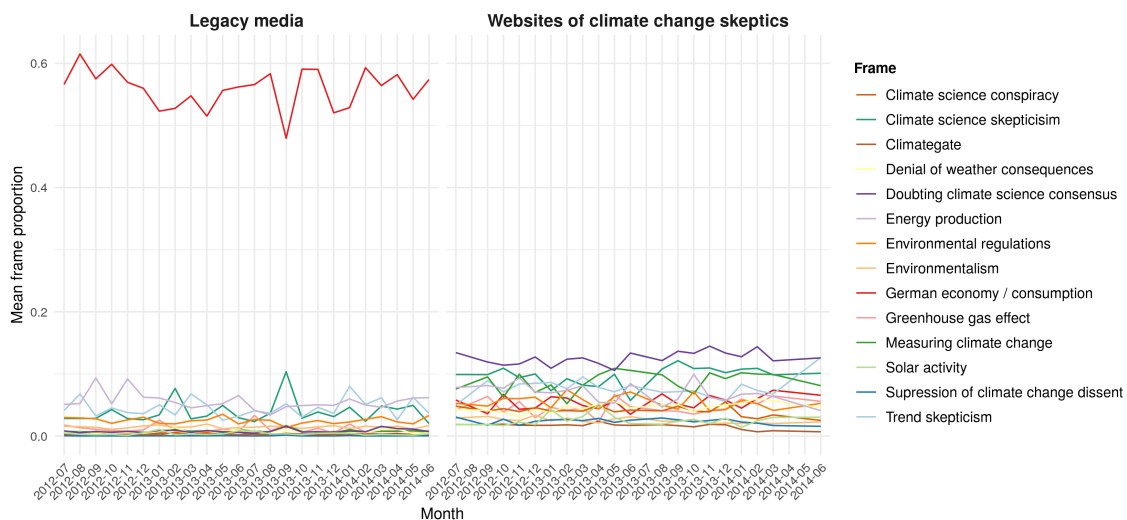

Figure 4.3: Mean topic proportions in legacy media and on websites of climate skeptics.

4.3.2 Selective resonance

Does this, however, mean that German legacy media are immune to climate skeptics' frames, positions, and allies, or do we find indicators of selective resonance among the different venues? Our data indicate that, in September 2013, a selective resonance might have occurred. In this month, the Jensen-Shannon divergence was noticeably low, which indicates a thematic resonance (Fig 4.2 B). Hereby, the frame, "climate science skepticism," peaked in the legacy media, reducing the attention on the dominating frame as regards the German economy, both in the conservative as well as in the other legacy media. This thematic resonance was accompanied by an above average share of skeptical sentences in the legacy media (Fig 4.2 C).

Looking at the articles that have both a high probability of the "climate science skepticism" frame and a high proportion of skeptical sentences, it can be seen that their publication was triggered by two events: the publication of the final draft of the 5th assessment report of the Intergovernmental Panel on Climate Change (IPCC) and the publication of an oceanographic study in the journal Nature. While the assessment report contains some more conservative projections than previous versions, the Nature study (Kosaka & Xie, 2013) addresses the lower than predicted temperature increases in recent decades. Both aspects—the IPCC's supposed failure/fraud and the plateau in the temperature curve—can be regarded as critical events, changing visibility structures within the climate change debates. The skeptics have successfully used these events to put forward their alleged counter-evidence for climate change referring to the news value "surprise".

These scientific publications have opened up a window of opportunity for skeptics' frames and positions to reach out to legacy media's agenda. A closed-up reading of the relevant articles, however, shows that such punctual visibility of skeptics' frames and positions in the legacy media have been opposed by the journalists themselves: they criticize skeptics' interpretations and, thus, accompany this selective resonance with their own critical examination of skeptics' arguments.

4.4 Discussion

In this article, we examined how different venues—the online venue of climate change skeptics and the mainstream discourse—resonate. Given that counter-movements are by definition oriented towards the mainstream of political debates (Fraser, 1990), this paper has investigated the opposite relationship, more challenging from the point of view of democratic theory: do the online counter-movements of climate skeptics resonate in traditional media? We have studied three types of discursive resonance: the media's adoption of climate change skeptics' frames (thematic resonance), the inclusion of skeptical positions (positional resonance), and the mentioning of skeptical actors (actor resonance). In addition to this, we have distinguished continuous from selective resonance patterns.

As a first major result of our study, we find no evidence for continuous resonance. We find neither an increase of skeptical frames in the legacy media's coverage, nor an increase of skeptical positions or actors. Second, however, this does not mean that skeptical voice are invisible in traditional media, as resonance may occur only selectively at single points in time. Our data reveal that specific events—in our case, the publication of scientific reports—can open a small window of opportunity that sees skeptical frames resonate to a greater degree in the media. While even in these cases skeptical positions remain contested by the media, their mere inclusion serves to ratify their views as legitimate contributions to the debate. This can be seen by counter-movements as a first step towards greater possibilities of participation in political debates.

Third, the ideological profile of newspapers and magazines in Germany did not turn out to be relevant: Conservative legacy media in Germany did not give greater resonance to climate skeptical voices. This strongly contrasts with the US, where conservative outlets play an important role in amplifying climate skepticism and are part of the "climate change denial machine" (Dunlap & McCright, 2011, p. 147). This finding also underlines the necessity to move research beyond the US context to better understand the role played by different context factors. In Germany, for example, the landscape of traditional media is substantially less polarized than in the US, and – contra the two component theory—journalists across the political spectrum appear to display no difference in their orientation towards news factors. However, while there is no qualitative difference between media outlets in terms of news factor emphasis, there is a telling quantitative distinction between them: while our sample includes more conservative than centrist or left-leaning media, climate change is much less salient in them. Lacking established skeptical actors on the national stage they could use as opportune witnesses in their reporting, conservative media resort to downplaying the issue by allocating less space to it.

How can we explain the lack of continuous resonance in the German case? News value theory hints at three factors (e.g., Schulz, 1976). First, the network of German climate change skeptics is rather weak. Most of the web pages in the sample are from foreign actors, in particular from the United States, and included in the German issue network by German skeptics. However, but neither the scope nor the language of these actors is necessarily German, and, as such, they are no points of reference for journalists, whose reporting is primarily oriented towards domestic politics. In news value terminology, they lack both the "closeness" and the "prestige". Second, German skeptics lack prominent speakers, whose status would guarantee continuous resonance. Third, the climate skeptical discourse is largely uncoupled from the policy cycle and its coverage by the media. Whereas news outlets followed the political that revolved around the economic aspects of climate change, skeptics discussed unrelated, more fundamental questions (e.g., whether it occurs or whether one can trust climate scientists), which did not fit into the relevance structure of news reporting. This leaves few options for climate skeptics and their resonance opportunities are mostly selective: without prestigious actors and disengaged from the policy cycle, they rely on event-driven news factors. The "surprise" news factor becomes an important element, and

our study shows that skeptics rely on external events like the publication of scientific reports that—from a mainstream perspective—reveal surprising findings.

Our results point out that German legacy media have not played a prominent role in spreading the ideas of climate skeptics. On the contrary, they seemed to fulfill their task of informing the public about climate change in accordance with the scientific consensus (e.g., Oreskes, 2004), and criticized the climate skeptical positions on which they reported, as our qualitative case study shows. While this conforms to the normative standards associated with journalism, the media largely reduced their climate change coverage to the economic aspect, neglecting its political dimension or its international scope. This one-sided focus may hamper the ability to act politically and actually solve the climate crisis.

Moreover, the media's exclusionary practice towards the skeptical counter-movement has its drawbacks: those supporting the counter-positions might well develop a feeling of alienation and misrepresentation. As a consequence, these people may turn to alternative digital information sources, while their distrust of traditional media increases—a trend that has occurred in recent years in Germany on the political far-right. The media's decision to refrain for instance from questioning the scientific consensus might well result in losing their role as legitimate gatekeepers in the debate in the eyes of some parts of the population.

Do our results mean that the engagement of counter-movements' online has no effect? Although skeptics lack a continuous resonance on legacy media agendas, they still have a selective one. Counter-movements' online communication serves as a "reservoir of ideas" from which they can draw as soon as a window of opportunity opens (Adam, Schmid-Petri, et al., 2019). However, further research needs to show whether counter-movements' online communication fulfills more functions. It might be that climate skeptical ideas and positions are disseminated without the help of traditional media. This raises the question of diffusion patterns of climate skepticism and the role played by traditional media, digital outlets, blogs and social media—and how they differ between countries. And depending on the national context, can skeptics bypass the media and directly influence politics and, in this way, gain legitimacy and a wider access to citizens?

Beyond this, future research needs to dig deeper into the conditions under which media resonance occurs. To further understand continuous resonance and the relevance of the news factor "prominence/elite," Germany would certainly be an interesting case. Here, the recent electoral gains of the right-wing "Alternative für Deutschland" (AfD) have also brought some climate change skeptics into parliament, which in turn might help skeptical voices to become more prominent in traditional media. In addition, more research is needed to understand which specific types of events and conditions make selective resonance more likely. In our research, climate skeptics have relied on external events to promote their ideas. However, can counter-movements create events themselves to advance their agenda (see Staggenborg, 1993)? Knowledge about these events and conditions might help make journalists more sensitive in their reporting. Finally, more comparative research is needed to gain a better understanding of digital discourse strategies

employed by climate skeptics and their resonance in media coverage. Our study already shows that the patterns we observe in Germany are not comparable to the ones in the US. We thus need to systematically study context factors.

Finally, there are a number of limitations that we need to address: First, we speak of the resonance of climate skeptical discourse in the media. Statistically, this corresponds to correlations models and we are aware that we cannot draw conclusions about causal relationships in a strict sense. Particularly with regard to frames, it might be climate skeptics whose discourse is influenced by the mainstream debate covered by the media. Second, empirically, the distinction between frames and positions is not as clear as presented by the theory. Frames define the relevance structure of an issue, regardless of the position of the actors. However, the results of our topic model suggest that some frames are aligned with a specific position (e.g., the denial of weather consequences reflects the skeptical position). At the same time, none of the frames is solely used by one side of the debate (Adam, Schmid-Petri, et al., 2019), and we are confident that our analysis captures the three crucial components of political debates: frames, positions, and actors. Third, our named-entity approach did not allow us to distinguish between speakers and addressees. However, this distinction is important to assess the role of the actors in the media reports. Beyond this, the automated approaches failed to reveal how journalists reacted to an increase in skeptical frames and positions. Here, only a qualitative case study can show journalists' critical reactions. To this end, additional methodological work is needed.

4.5 Supporting information

4.5.1 S1 Appendix: Starting points for the snowball-sampling of websites

Actor	Start URL
Analyse+Aktion	http://astrologieklassisch.wordpress.com/tag/klimawandel
EIKE - Europäisches Institut für Klima und Energie	http://www.eike-klima-energie.eu
Klimaüberraschung	http://www.klima-ueberraschung.de
Klimaskeptiker	http://www.klimaskeptiker.info

Table 4.6: Starting points for the snowball-sampling of websites.

4.5.2 S2 Appendix: List of German legacy media

Newspaper/magazine	Conservative	Number of articles
Bild	x	33
Bild am Sonntag	x	15
Der Spiegel		88
Die Welt	x	347
Die Zeit		175
Frankfurter Allgemeine Zeitung	x	622
FAZ am Sonntag	x	101
Financial Times Deutschland (until 12/2012)	x	47
Focus	x	43
Frankfurter Rundschau		466
Handesblatt	x	172
Stern		40
Süddeutsche Zeitung		1,225
Taz, die Tageszeitung		622
Welt am Sonntag	x	112

Table 4.7: List of German legacy media.

4.5.3   S3 Appendix: Topic labels and top words

Label	Top words
Climate science conspiracy	wissenschaft, klima, unterlag, herzland, wissenschaftl, institut, peer, arbeit, mensch, falsch
Climate science conspiracy	e-mail, freitas, papier-, wissenschaft, mann, klima, mannschaft, versuch, sorg, geschicht
Climate science skepticisim	prof, jahr, bitt, ipcc, co2, wissenschaft, global, klimawandel, klima, erwarm
Climategate	wissenschaft, tagebuch, freitas, aktion, papi, hinweis, rezension, klima, phil, person
Denial of weather consequences	klima, co2, ansteig, erhohen,, verander, erwarm, global, mensch, wissenschaft, wett
Doubting climate science consensus	klima, wissenschaft, global, erwarm, verander, ipcc, wissenschaftl, bericht, polit, forschung
Energy production	energi, wind, leistung, vereinigt, regier, polit, land, wirtschaft, europa, kost
Environmental regulations	pflanz, kohlenstoff, mensch, kohl, umwelt, epa, jahr, emission, welt, zustand
Environmentalism	mensch, natur, bericht, vereinigt, klima, konigreich, umgeb, leb, beispiel, auswirk
German economy/consumption	jahr, deutsch, klimawandel, deutschland, gross, prozent, land, energiew, europa, hoh
Greenhouse gas effect	atmosphar, temperatur, co2, oberflach, erhoht, strahlung, erd, energi, luft, ansteig
Measuring climate change	zeigt, temperatur, dat, jahr, erwarm, modell, bedeut, global, trend, zahl
Solar activity	solar, klima, sonn, erd, zyklus, aktivitat, wolk, kosmisch, strahl, planet
Supression of climate change dissent	wissenschaft, klima, seeland, aufzeichn, temperatur, niwa, verander, position, bemerk, debatt
Trend skepticism	jahr, eis, erhebt, meer, ansteig, erhohen, global, erwarm, wett, temperatur

Table 4.8: Topic labels and top words.

4.5.4   S4 Appendix: Performance indicators for the semi-automated classifier

	Macro F1			Micro F1
Language	Skeptics	Advocate	irrelevant/neutral	overall
English	0.62	0.68	0.91	0.82
German	0.92	0.89	0.88	0.85

Table 4.9: Performance indicators for the semi-automated classifier.

Reported are the F1 scores (harmonic average of precision and recall; macro F1 for the different categories and micro F1 overall) after the third iteration of active learning. The scores are calculated by comparing the machine coding against the manual coding (gold standard).

References

Adam, S. (2007). *Symbolische Netzwerke in Europa: Der Einfluss der nationalen Ebene auf europäische Öffentlichkeit. Deutschland und Frankreich im Vergleich.* Köln: Herbert von Halem.

Adam, S., Häussler, T., Schmid-Petri, H., & Reber, U. (2016). Identifying and analyzing hyperlink issue networks. In G. Vowe & P. Henn (Eds.), *Political communication in the online world: Theoretical approaches and research designs* (p. 233-247). New York: Routledge.

Adam, S., Häussler, T., Schmid-Petri, H., & Reber, U. (2019). Political contestation online: Analyzing coalitions and their online strength in the field of climate change. *New Media & Society*, *21*(11-12), 2671-2690. doi: 10.1177/1461444819855966

Adam*, S., Reber*, U., Häussler*, T., & Schmid-Petri, H. (under review). How climate change skeptics (try to) spread their ideas: Using computational methods to assess the resonance among skeptics' and legacy media. *PLOS ONE*.

Adam, S., Schmid-Petri, H., Reber, U., & Häussler, T. (2019, May). *An analysis of online framing dynamics between climate advocates and skeptics in the UK.* Paper presented at the 69th Annual Conference of the International Communication Association (ICA). Washington, DC.

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011). Sentiment analysis of Twitter data. In *Proceedings of the Workshop on Language in Social Media* (p. 30-38). Retrieved from http://www.cs.columbia.edu/~julia/papers/Agarwaletal11.pdf

Balahur, A., & Turchi, M. (2014). Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. *Computer Speech and Language*, *28*, 56-75. doi: 10.1016/j.csl.2013.03.004

Barnett, G. A., Chon, B.-S., & Rosen, D. (2001). The structure of the internet flows in cyberspace. *Networks and Communication Studies*, *15*(1-2), 61-80.

Barnett, G. A., Chung, C. J., & Park, H. W. (2011). Uncovering transnational hyperlink patterns and web-mediated contents: A new approach based on cracking.com domain. *Social Science Computer Review*, *29*(3), 369-384. doi: 10.1177/0894439310382519

Barnett, G. A., & Park, H. W. (2005). The structure of international internet hyperlinks and bilateral bandwidth. *Annales des telecommunications*, *60*(9-10), 1110-1127. doi: 10.1007/BF03219838

Baumgartner, F. R., & Jones, B. D. (1993). *Agendas and instability in American politics.* Chicago: University of Chicago Press.

Beck, K. (2018). *Das Mediensystem Deutschlands: Strukturen, Märkte, Regulierungen* (2nd ed.). Wiesbaden: Springer VS.

Begenat, M. (2016). *Öffentlichkeit – für alle? Themen und Informationsrepertoires in politischen Milieus.* Wiesbaden: Springer VS. doi: 10.1007/978-3-658-11286-8

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom.* New Haven: Yale University Press.

Benkler, Y., Roberts, H., Faris, R., Solow-Niederman, A., & Etling, B. (2015). Social mobilization and the networked public sphere: Mapping the SOPA-PIPA debate. *Political Communication, 32*(4), 594-624. doi: 0.1080/10584609.2014.986349

Bennett, W. L. (1990). Toward a theory of press-state relations in the United-States. *Journal of Communication, 40*(2), 103-125. doi: 10.1111/j.1460-2466.1990.tb02265.x

Bennett, W. L., Lang, S., & Segerberg, A. (2015). European issue publics online: The cases of climate change and fair trade. In T. Risse (Ed.), *European public spheres: Politics is back* (p. 108-137). Cambridge: Cambridge University Press.

Benoit, K., Schwarz, D., & Traber, D. (2012, June). *The sincerity of political speech in parliamentary systems: A comparison of ideal points scaling using legislative speech and votes.* Paper presented at the second annual conference of the European Political Science Association (EPSA), Berlin. Retrieved from http://www.dschwarz.ch/downloads/ EPSA_2012_BenoitSchwarzTraber.pdf?attredirects=0

Benoit, K., Watanabe, K., Nulty, P., Obeng, A., Wang, H., Lauderdale, B., & Lowe, W. (2017). quanteda: Quantitative analysis of textual data [Computer software]. Retrieved from http://quanteda.io

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM, 55*(4), 77-84. doi: 10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(4/5), 993-1022. Retrieved from http://jmlr.csail.mit.edu/papers/v3/blei03a.html

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism, 4*(1), 8-23. doi: 10.1080/21670811.2015.1096598

Boussalis, C., & Coan, T. G. (2016). Text-mining the signals of climate change doubt. *Global Environmental Change, 36*, 89-100. doi: 10.1016/j.gloenvcha.2015.12.001

Boykoff, M. T., & Boykoff, J. M. (2004). Balance as bias: Global warming and the US prestige press. *Global Environmental Change, 14*, 125-136. doi: 10.1016/j.gloenvcha.2003.10.001

Brin, S., & Page, L. (1998, April). *The anatomy of a large-scale hypertextual Web search engine.* Paper presented at the seventh International World-Wide Web Conference. Retrieved from http://ilpubs.stanford.edu:8090/361/

Broadbent, J., Sonnett, J., Botetzagias, I., Carson, M., Carvalho, A., Chien, Y.-J., … Zhengyi, S. (2016). Conflicting climate change frames in a global field of media discourse. *Socius*, *2*, 1-17. doi: 10.1177/2378023116670660

Brulle, R. (2014). Institutionalizing delay: Foundation funding and the creation of U.S. climate change counter-movement organizations. *Climatic Change*, *122*(4), 681-694. doi: 10.1007/s10584-013-1018-7

Brunnengräber, A. (2018). Klimaskeptiker im Aufwind. Wie aus einem Rand- ein breiteres Gesellschaftsphänomen wird. In O. Kühne & F. Weber (Eds.), *Baustine der Energiewende. RaumFragen: Stadt-Region-Landschaft* (p. 271-292). Wiesbaden: Springer VS.

Burck, J., Hagen, U., Marten, F., Höhne, N., & Bals, C. (2019). *Climate change performance index: results 2019.* Bonn: Germanwatch, Climate Action Network International and NewClimate Institute. Retrieved from https://www.climate-change-performance-index.org/the-climate-change-performance-index-2019

Cammaerts, B., & Van Audenhove, L. (2005). Online political debate, unbound citizenship, and the problematic nature of a transnational public sphere. *Political Communication*, *22*(2), 179-196. doi: 10.1080/10584600590933188

Capstick, S. B., & Pidgeon, N. F. (2014). What is climate change scepticism? Examination of the concept using a mixed methods study of the UK public. *Global Environmental Change*, *24*, 389-401. doi: 10.1016/j.gloenvcha.2013.08.012

Carmichael, J., Brulle, R., & Huxster, J. (2017). The great divide: Understanding the role of media and other drivers of the partisan divide in public concern over climate change in the USA. *Climatic Change*, *141*(4), 599-612. doi: 0.1007/s10584-017-1908-1

Chadwick, A. (2013). *The hybrid media system: Politics and power.* Oxford: Oxford University Press.

Chang, T.-K., Himelboim, I., & Dong, D. (2009). Open global networks, closed international flows. *International Communiction Gazette*, *71*(3), 137-159. doi: 10.1177/1748048508100910

Cobb, R., Ross, J.-K., & Ross, M. H. (1976). Agenda building as a comparative political process. *American Political Science Review*, *70*, 126-138. doi: 10.1017/S0003055400264034

Cook, J. (2016). Countering climate science denial and communicating scientific consensus. *Oxford Research Encyclopedia of Climate Science.* Retrieved from https://oxfordre.com/climatescience/view/10.1093/acrefore/9780190228620.001.0001/acrefore-9780190228620-e-314

DeepL. (2017). *Der DeepL Übersetzer im Vergleich zur Konkurrenz.* Retrieved from https://www.deepl.com/press.html

De Maeyer, J. (2012). Towards a hyperlinked society: A critical review of link studies. *New Media & Society*, *15*(5), 737-751. doi: 10.1177/1461444812462851

Denny, M. J., & Spirling, A. (2018). Text preprocessing for unsupervised learing: Why it matters, when it misleads, and what to do about it. *Political Analysis*, *26*(2), 168-189. doi: 10.1017/pan.2017.44

Deutsch, K. (1953). *Nationalism and social communication: An inquiry into the foundations of nationality.* Cambridge: University Press.

De Vries, E., Schoonevelde, M., & Schumacher, G. (2018). No longer lost in translation: Evidence that Google Translate works for comparative bag-of-words text applications. *Political Analysis*, *26*(4), 417-430. doi: 10.1017/pan.2018.26

Di Gregorio, M., Fatorelli, L., Paavola, J., Locatelli, B., Pramova, E., Nurrochmat, D. R., … Kusumadewi, S. D. (2019). Multi-level governance and power in climate change policy networks. *Global Environmental Change*, *54*, 64 - 77. doi: 10.1016/j.gloenvcha.2018.10.003

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topics modelling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics*, *41*(6), 570-606. doi: 10.1016/j.poetic.2013.08.004

Dunlap, R. E., & McCright, A. M. (2011). Organized climate change denial. In J. S. Dryzek, R. B. Norgaard, & D. Schloberg (Eds.), *The Oxford handbook of climate change* (p. 144-160). Oxford: Oxford University Press.

Dunlap, R. E., & McCright, A. M. (2015). Challenging climate change: The denial countermovement. In R. E. Dunlap & J. Brulle (Eds.), *Climate change and society. sociological perspectives* (p. 300-332). Oxford: Oxford University Press.

Dunlap, R. E., McCright, A. M., & Yarosh, J. H. (2016). The political divide on climate change: Partisan polarization widens in the U.S. *Environment: Science and Policy for Sustainable Development*, *58*(5), 4-23. doi: 10.1080/00139157.2016.1208995

Eder, K. (2000). Zur Transformation nationalstaatlicher Öffentlichkeit in Europa: Von der Sprachgemeinschaft zur issuespezifischen Kommunikationsgemeinschaft. *Berliner Journal für Soziologie*, *10*, 167-284. doi: 10.1007/BF03204349

Eder, K., & Kantner, C. (2000). Transnationale Resonanzstrukturen in Europa: Eine Kritik der Rede vom Öffentlichkeitsdefizit. In M. Bach (Ed.), *Die Europäisierung nationaler Gesellschaften* (p. 306-331). Wiesbaden: Westdeutscher Verlag.

Eide, E., & Kunelius, R. (2010). Domesticating global moments: A transnational study on the coverage of the Bali and Copenhagen climate summits. In E. Eide & R. Kunelius (Eds.), *Global climate—local journalisms: A transnational study of how media make sense of climate summits* (p. 11-50). Bochum: Projektverlag.

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. doi: 10.1080/17524032.2014.983536

Elgin, D. J. (2015). Utilizing hyperlink network analysis to examine climate change supporters and opponents. *Review of Policy Research*, *32*(2), 226-245. doi: 10.1111/ropr.12118

Ellinas, A. A. (2010). *The media and the far right in Western Europe.* Cambridge: Cambridge University Press.

Engels, A., Hüther, O., Schäfer, M. S., & Held, H. (2013). Public climate-change skepticism, energy preferences and political participation. *Global Environmental Change*, *23*, 1018-1027. doi: 10.1016/j.gloenvcha.2013.05.008

Esser, F. (2013). The emerging paradigm of comparative communication enquiry: Advancing cross-national research in times of globalization. *International Journal of Communication*, *7*, 113-128. Retrieved from https://ijoc.org/index.php/ijoc/article/view/1968/843

Esser, F., & Hanitzsch, T. (2012). On the why and how of comparative inquiry in communication studies. In F. Esser & Hanitzsch (Eds.), *Handbook of comparative communication research* (p. 3-22). London: Routledge.

Etling, B., Kelly, J., Faris, R., & Palfrey, J. (2010). Mapping the Arabic blogosphere: Politics and dissent online. *New Media & Society*, *12*(8), 1225-1243. doi: 10.1177/1461444810385096

Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., & Lin, C.-J. (2008). LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, *9*, 1871-1874.

Farrell, J. (2016a). Corporate funding and ideological polarization about climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(1), 92-97. doi: 10.1073/pnas.1509433112

Farrell, J. (2016b). Network structure and influence of the climate change counter-movement. *Nature Climate Change*, *6*, 370-374. doi: 10.1038/nclimate2875

Feldman, L., Maibach, E. W., Roser-Renouf, C., & Leiserowitz, A. (2012). Climate on cable: The nature and impact of global warming coverage on Fox News, CNN, and MSNBS. *The International Journal of Press/Politics*, *17*(1), 3-31. doi: 10.1177/1940161211425410

Finkel, J. R., Grenage, T., & Manning, C. (2005, June). Incorporating non-local information into information extraction systems by Gibbs samplings. In K. Knight, H. T. Ng, & K. Oflazer (Eds.), *Proceedings of the 43nd annual meeting of the association for computational linguistics* (p. 363-370). Ann Arbor: Stroudsburg: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/P05-1045.pdf

Fisher, D. R., Waggle, J., & Leifeld, P. (2013). Where does political polarization come from? Locationg polarization within the U.S. climage change debate. *American Behavioral Scientist*, *57*(1), 70-92. doi: 10.1177/0002764212463360

Fraser, N. (1990). Rethinking the public sphere: A contribution to the critique of actually existing democracy. *Social Text*, *25/26*, 56-80. doi: 10.2307/466240

Fraser, N. (2007). Transnationalizing the public sphere: On the legitimacy and efficacy of public opinion in a post-Westphalian world. *Theory, Culture & Society*, *24*(4), 7-30. doi: 10.1177/0263276407080090

Friedland, L. A., & Rojas, H. (2006). The networked public sphere. *Javnost - The Public*, *13*(4), 5-26. doi: 10.1080/13183222.2006.11008922

Fu, J. S., & Shumate, M. (2017). News media, social medai, and hyperlink networks: An examination of integrated media effects. *The Information Society*, *33*(2), 53-63. doi: 10.1080/01972243.2016.1271379

Galtung, J., & Ruge, M. H. (1965). The structure of foreign news. the presentation of the Congo, Cuba and Cyprus srises in four Norwegian newspapers. *Journal of Peace Research*, *2*, 64-91. doi: 10.1177/002234336500200104

Gamson, W. A., & Modigliani, A. (1987). The changing culture of affirmative action. In
R. G. Braungart & M. M. Braungart (Eds.), *Research in political sociology* (p. 137-177).
Greenwich: JAI Press.

Gamson, W. A., & Wolfsfeld, G. (1993). Movements and media as interacting systems. *The Annals of the American Academy of Political and Social Science*, *528*(1), 114-125.

Gavin, N. T., & Mashall, T. (2011). Mediated climate change in Britain: Scepticism on the web and on television around Copenhagen. *Global Environmental Change*, *21*, 1035-1044. doi:
10.1016/j.gloenvcha.2011.03.007

Gerhards, J. (1993). Westeuropäische Integration und die Schwierigkeiten der Entstehung einer europäischen Öffentlichkeit. *Zeitschrift für Soziologie*, *22*(2), 96-110.

Gerhards, J., & Neidhardt, F. (1990). *Strukturen und Funktionen moderner Öffentlichkeit: Fragestellungen und Ansätze* (Discussion Paper No. FS III 90-101). Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Gerhards, J., & Rössel, J. (1999). Zur Transnationalisierung der Gesellschaft der Bundesrepublik: Entwicklungen, Ursachen und mögliche Folgen für die europäische Integration. *Zeitschrift für Soziologie*, *25*(5), 325-344.

Gerlitz, C., & Helmond, A. (2013). The like economy: Social buttons and the data-intensive web. *New Media & Society*, *15*(8), 1348-1365. doi: 10.1177/1461444812472322

González-Bailón, S. (2009). Opening the black box of link formation: Social factors underlying the structure of the Web. *Social Networks*, *31*(4), 271-280. doi: 10.1016/j.socnet.2009.07.003

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. doi:
10.1093/pan/mps028

Grundmann, R., & Scott, M. (2014). Disputed climate science in the media: Do countries matter?
Public Understanding of Science, *23*(2), 220-235. doi: 10.1177/0963662512467732

Habermas, J. (1962). *Strukturwandel der Öffentlichkeit. Untersuchungen zu einer Kategorie der bürglerlichen Gesellschaft.* Neuwied, Berlin: Luchterhand.

Habermas, J. (1998). Die postnationale Konstellation und die Zukunft der Demokratie. In *Die postnationale Konstellation: Politische Essays* (p. 91-169). Frankfurt am Main: Suhrkamp.

Hagen, L. M. (1993). Opportune witnesses. An analysis of balance in the selection of sources und arguments in the leading german newspapers' coverage of the census issue. *European Journal of Communication*, *8*, 317-343. doi: 10.1177/0267323193008003004

Halvais, A. (2000). National borders on the World Wide Web. *New Media & Society*, *2*(1), 7-28.
doi: 10.1177/14614440022225689

Hampshire, S., & Porta Salvia, C. (2010). Translation and the Internet: Evaluating the quality of free online machine translators. *Quaderns: Revista de Traducció*, *17*, 197-209. Retrieved from https://www.raco.cat/index.php/QuadernsTraduccio/article/view/194256

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*(3859), 1243-1248. doi:
10.1126/science.162.3859.1243

Heft, A., Mayerhöffer, E., Reinhardt, S., & Knüpfer, C. (2019). Beyond Breitbart: Comparing right-wing digital news infrastructures in six western democracies. *Policy & Internet*, *12*(1). doi: 10.1002/poi3.219

Hepp, A., Elsler, M., Lingenberg, S., Mollen, A., Möller, J., & Offerhaus, A. (2016). *The communicative construction of Europe: Cultures of political discourse, public sphere and the Euro crisis*. Houndmills, Basingstoke: Palgrave Macmillan. doi: 10.1057/9781137453136

Himelboim, I. (2010). The international network structure of news media: An analysis of hyperlinks usage in news Web sites. *Journal of Broadcasting & Electronic Media*, *54*(3), 373-390. doi: 10.1080/08838151.2010.499050

Häussler, T., Adam, S., Schmid-Petri, H., & Reber, U. (2017). How political conflict shapes online spaces. a comparison of climate change hyperlink networks in the U.S. and Germany. *International Journal of Communication*, *11*, 3096–3117. Retrieved from http://ijoc.org/index.php/ijoc/article/view/5644

Häussler, T., Schmid-Petri, H., Adam, S., Reber, U., & Arlt, D. (2016). The climate of debate: How institutional factors shape legislative discourses on climate change. A comparative framing perspective. *Studies in Communication Sciences*, *16*(1), 94-102. doi: 10.1016/j.scoms.2016.04.002

Iacobuta, G., Dubash, N. K., Upadhyaya, P., Deribe, M., & Höhne, N. (2018). National climate chnage mitigation legislation, strategy and targets: A global update. *Climate Policy*, *18*(9), 1114-1132. doi: 10.1080/14693062.2018.1489772

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.* Genève: IPCC. Retrieved from http://ar5-syr.ipcc.ch/

Ivanova, A. (2017). *Transnationalisierung von Öffentlichkeit. eine länderübergreifende langzeitanalyse der klimaberichterstattung in leitmedien.* Wiesbaden: Springer Fachmedien. doi: 10.1007/978-3-658-18356-1

Ivanova, A., Schmidt, A., & Schäfer, M. S. (2014). Global climate change, global public sphere? Media attention for climate change in 27 countries. In M. José Canel & K. Voltmer (Eds.), *Comparing political communication across time and space: New studies in an emerging field* (p. 210-227). New York: Palgrave Macmillan.

Jacobi, C., van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, *4*(1), 89-106. doi: 10.1080/21670811.2015.1093271

Janc, K. (2015). Geography of hyperlinks: Spatial dimensions of local government websites. *European Planning Studies*, *23*(5), 1019-1037. doi: 10.1080/09654313.2014.889090

Jang, S. M., & Hart, P. S. (2015). Polarized frames on "climate change" and "global warming" across countries and states: Evidence from Twitter big data. *Global Environmental Change*, *32*, 11-17. doi: 10.1016/j.gloenvcha.2015.02.010

Joachims, T. (1998, April). Text categorization with Support Vector Machines: Learning with many relevant features. In C. Nédellec & C. Rouveirol (Eds.), *Proceedings of the 10th*

European conference on machine learning (p. 137-142). Chemnitz: Berlin, Heidelberg: Springer. doi: 10.1007/BFb0026683

Jordan, A., van Asselt, H., Berkhout, F., Huitema, D., & Rayner, T. (2012). Understanding the paradoxes of multilevel governing: Climate change policy in the European Union. *Global Environmental Politics*, *12*(2), 43-66. doi: 10.1162/glep_a_00108

Jänicke, M. (2017). The multi-level system of global climate governance – the model and its current state. *Environmental Policy and Governance*, *27*(2), 108-121. doi: 10.1002/eet.1747

Kaiser, J., & Rhomberg, M. (2015). Questioning the doubt: Climate skepticism in German newspaper reporting on COP17. *Environmental Communication*, *10*(5), 556-574. doi: 10.1080/17524032.2015.1050435

Kantner, C. (2004). *Kein modernes Babel: Kommunikative Voraussetzungen europaischer Offentlichkeit.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Kepplinger, H. M. (2008). News values. In W. Donsbach (Ed.), *The international encyclopedia of communication* (p. 3281-3286). Oxford: Blackwell.

Kirilenko, A. P., & Stepchenkova, S. O. (2014). Public microblogging on climate change: One year of Twitter worldwide. *Global Environmental Change*, *26*, 171-187. doi: 10.1016/j.gloenvcha.2014.02.008

Koopmans, R., & Erbe, J. (2004). Towards a European public sphere? Vertical and horizontal dimensions of Europeanized political communication. *Innovation*, *17*(2), 97-118. doi: 10.1080/1351161042000238643

Koopmans, R., & Statham, P. (2010). Theoretical framework, research design, and methods. In R. Koopmans & P. Statham (Eds.), *The making of a European public sphere: Media discourse and political contention* (p. 31-59). Cambridge: Cambridge University Press.

Koopmans, R., & Zimmermann, A. (2010). Transnational political communication on the internet: Search engine results and hyperlink networks. In R. Koopmans & P. Statham (Eds.), *The making of a European public sphere: Media discourse and political contention* (p. 171-194). Cambridge: Cambridge University Press.

Kosaka, Y., & Xie, S.-P. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, *501*, 403-407. doi: 10.1038/nclimate2875

Krackhardt, D., & Stern, R. N. (1988). Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly*, *51*(2), 123-140.

Kriesi, H. (2004). Strategic political communication: Mobilizing public opinion in "audience democracies". In F. Esser & B. Pfetsch (Eds.), *Comparing political communication: Theories, cases, and challenges* (p. 184-212). Cambridge: Campridge University Press.

Lewandowski, S., Oreskes, N., Risbey, J. S., Newell, B. R., & Smithson, M. (2015). Seepage: Climate change denial and its effect on the scientific community. *Global Environmental Change*, *33*, 1-13. doi: 10.1016/j.gloenvcha.2015.02.013

Livingstone, S. (2003). On the challenges of cross-national comparative media research. *European Journal of Communication*, *18*(4), 477-500. doi: 10.1177/0267323103184003

Livingstone, S. (2012). Challenges to comparative research in a globalizing media landscape. In F. Esser & T. Hanitzsch (Eds.), *Handbook of comparative communication research* (p. 415-429). London: Routledge.

Lotz, S., & van Rensburg, A. (2014). Translation technology explored: Has a three-year maturation period done Google Translate any good? *Stellenbosch Papers in Linguistics Plus*, *43*, 235-259. doi: 10.5842/43-0-205

Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, *23*(2), 254-277. doi: 10.1093/pan/mpu019

Lucas, C., & Tingley, D. (2014). translateR: Bindings for the Google and Microsoft translation APIs [Computer software]. Retrieved from https://CRAN.R-project.org/package=translateR

Luhmann, N. (1995). *Die Realität der Massenmedien*. Opladen: Westdeutscher Verlag.

Lörcher, I., & Taddicken, M. (2017). Discussing climate change online. topics and perceptions in online cimate change communication in different online public arenas. *Journal of Science Communication*, *16*(2). doi: 10.22323/2.16020203

Lüter, A. (2004). Politische Profilbildung jenseits der Parteien? Redaktionelle Linien in Kommentaren deutscher Qualitätszeitungen. In *Die Stimme der Medien* (p. 167-195). Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage.

Maier, D., Waldherr, A., Miltner, P., Jähnichen, P., & Pfetsch, B. (2018). Exploring issues in a networked public sphere: Combining hyperlink network analysis and topic modeling. *Social Science Computer Review*, *36*(1), 3-20. doi: 10.1177/0894439317690337

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Heyer, G., … Adam, S. (2018). Applying LDA topic modeling in communication research: Towards a valid and reliable methodology. *Communication Methods and Measures*, *12*(2-3), 93-118. doi: 10.1080/19312458.2018.1430754

Marres, N. (2006). Net-work is format work: Issue networks and the sites of civil society politics. In J. Dean, J. W. Anderson, & G. Lovik (Eds.), *Reformatting politics: Information technology and global civil society* (p. 3-18). New York: Routledge.

Marres, N. (2015). Why map issues? on controversy analysis as a digital method. *Sience, Technology, & Human Values*, *40*(5), 655-686. doi: 10.1177/0162243915574602

Marres, N., & Rogers, R. (2005). Recipe for tracing the fate of issues and their publics on the Web. In B. Latour & P. Weibel (Eds.), *Making things public: Atmospheres of democracy* (p. 922-935). Karlsruhe, Cambridge: ZKM Center for Art and Media, The MIT Press.

Maurer, M., & Reinemann, C. (2006). *Medieninhalte. Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.

McCright, A. M., Dunlap, R. E., & Marquart-Pyatt, T., Sandra. (2015). Political ideology and views about climate change in the european union. *Environmental Politics*, *25*(2), 338-358. doi: 10.1080/09644016.2015.1090371

Metag, J., Füchslin, T., & Schäfer, M. S. (2017). Global warming's five Germanys: a typology of Germans' views on climate change and patterns of media use and information. *Public Understanding of Science*, *26*(4), 434-451. doi: 10.1177/0963662515592558

Mimno, D., & Blei, D. (2011). Bayesian checking for topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing.* Retrieved from http://www.cs.columbia.edu/~blei/papers/MimnoBlei2011.pdf

Mimno, D., Wallach, H., Naradowsky, J., Smith, D., & McCallum, A. (2009). Polylingual topic models. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2* (p. 880-889). Retrieved from http://dl.acm.org/citation.cfm?id=1699571.1699627

Niekler, A., & Jähnichen, P. (2012). Matching results of latent Dirichlet allocation for rext. In N. Rußwinkel, U. Drewitz, & H. van Rijn (Eds.), *Proceedings of the 11th International Conference on Cognitive Modeling* (p. 317-322). Berlin: Universitätsverlag der TU Berlin. Retrieved from http://asv.informatik.uni-leipzig.de/publication/file/210/nieklerjaehnichenICCM2012.pdf

Olausson, U. (2009). Global warming—global responsibility? Media frames of collective action and scientific certainty. *Public Understanding of Science*, *18*, 421-436. doi: 10.1177/0963662507081242

Oreskes, N. (2004). The scientific consensus on climate change. *Science*, *306*(5702), 1686. doi: 10.1126/science.1103618

Painter, J., & Ashe, T. (2012). Cross-national comparison of the presence of climate scepticism in the print media in six countries, 2007-10. *Environmental Research Letters*, *7*(4), 1-8. doi: 10.1088/1748-9326/7/4/044005

Painter, J., & Gavin, N. (2016). Climate skepticism in British newspapers, 2007-2077. *Environmental Communication*, *10*(4), 432-452. doi: 10.1080/17524032.2014.995193

Park, H. W. (2003). Hyperlink network analysis: A new method for the study of social structure on the Web. *Connections*, *25*(1), 49-61.

Park, H. W., Barnett, G. A., & Chung, C. J. (2011). Structural changes in the 2003-2009 global hyperlink networks. *Global Networks*, *11*(4), 522-542. doi: 10.1111/j.1471-0374.2011.00336.x

Pennings, P. (2011). Assessing the 'gold standard' of party policy placements: Is computerized replication possible? *Electoral Studies*, *30*(3), 561-570. doi: 10.1016/j.electstud.2011.05.001

Peters, B. (1994). Der Sinn von Öffentlichkeit. In F. Neidhardt (Ed.), *Öffentlichkeit, öffentliche Meinung, soziale Bewegungen* (p. 43-76). Opladen: Westdeutscher Verlag.

Pew Research Center. (2018). *Fact sheet: News media and political attitudes in Germany.* Retrieved from https://www.pewresearch.org/global/fact-sheet/news-media-and-political-attitudes-in-germany/

Pfetsch, B., Adam, S., & Bennett, W. L. (2013). The critical linkage between online and offline media - an approach to researching the conditions of issue spill-over. *Javnost - The Public*, *20*(3), 5-18. doi: 10.1080/13183222.2013.11009118

Price, V., & Tewksbury, D. (1997). News values and public opinion: A theoretical account of media priming and framing. In G. A. Barett & F. J. Boster (Eds.), *Progress in communication sciences: Advances in persuasion* (p. 173-212). Greewich: Ablex.

Rahmstorf, S. (2004). *The climate sceptics.* Potsdam: Potsdam Institute for Climate Impact Research. Retrieved from https://www.pik-potsdam.de/news/public-events/archiv/alter-net/former-ss/2006/programme/28-08.2006/rahmstorf/literature/rahmstorf_climate_sceptics_2004.pdf

Reber, U. (2019). Overcoming language barriers: Assessing the potential of machine translation and topic modeling for the comparative analysis of multilingual text corpora. *Communication Methods and Measures*, *13*(2), 102-125. doi: 10.1080/19312458.2018.1555798

Reber, U. (2020). Global climate change or national climate changes? an analysis of the performance of online issue publics in integrating global issues. *Environmental Communication, Advance online publication*, 1-16. doi: 10.1080/17524032.2020.1812685

Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in social sciences. *Journal of the American Statistical Association*, *111*(515), 988-1003. doi: 10.1080/01621459.2016.1141684

Roberts, M. E., Stewart, B. M., & Tingley, D. (2017). stm: R package for Structural Topic Models [Computer software]. Retrieved from http://www.structuraltopicmodel.com

Roberts, M. E., Stewart, B. M., Tingley, D., & Airoldi, E. M. (2013). *The Structural Topic Model and applied social science.* Paper presented at the Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation. Retrieved from http://scholar.princeton.edu/files/bstewart/files/stmnips2013.pdf

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Kushner Gadarian, S., … Rand, D. G. (2014). Structural Topic Models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064-1082. doi: 10.1111/ajps.12103

Rogers, R. (2013). Mapping public Web space with the Issuecrawler. In B. Reber & C. Brossaud (Eds.), *Digital cognitive technologies: Epistemology and the knowledge economy* (p. 89-99). Hoboken: Wiley.

Rogers, R., & Ben-David, A. (2008). The Palestinian-Israeli peace process and transnational issue networks: The complicated place of the Israeli NGO. *New Media & Society*, *10*(3), 497-528. doi: 0.1177/1461444807085321

Rogers, R., & Marres, N. (2000). Landscaping climate change: A mapping technique for understanding sience and technology debates on the World Wide Web. *Public Understanding of Sience*, *9*(2). doi: 10.1088/0963-6625/9/2/304

Scharkow, M. (2013). Thematic content analysis using supervised machine learning: An empirical evaluation using German online news. *Quality & Quantity*, *47*(2), 761-773. doi: 10.1007/s11135-011-9545-7

Schmid-Petri, H. (2017). Do conservative media provide a forum for skeptical voices? The link between ideology and the coverage of climate change in British, German, and Swiss

newspapers. *Environmental Communication*, *11*(4), 554-567. doi: 10.1080/17524032.2017.1280518

Schmid-Petri, H., Adam, S., Reber, U., Häussler, T., Maier, D., Miltner, P., … Waldherr, A. (2018). Homophily and prestige: An assessment of their relative strenght to explain link formation in the online climate change debate. *Social Networks*, *55*, 47-54. doi: 10.1016/j.socnet.2018.05.001

Schmid-Petri, H., Adam, S., Schmucki, I., & Häussler, T. (2017). A changing climate of skepticism: The factors shaping climate change coverage in the US press. *Public Understanding of Science*, *26*(4), 498-513. doi: 10.1177/0963662515612276

Schmid-Petri, H., & Arlt, D. (2016). Constructing an illusion of scientific uncertainty? Framing climate change in German and British print media. *Communications*, *41*(3), 265-289. doi: 10.1515/commun-2016-0011

Schmid-Petri, H., Häussler, T., Adam, S., Waldherr, A., Miltner, P., Pfetsch, B., & Bennett, L. (2013). *Codebook for the analysis of frames in online documents and newspaper articles. Issue climate change.* Retrieved from https://www.ikmb.unibe.ch/ueber_uns/personen/prof_dr_adam_silke/e222290/files222291/Codebook_ClimateChange_SchmidPetrietal_2013_ger.pdf

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. doi: 10.1016/j.gloenvcha.2013.07.020

Schulz, W. (1976). *Die Konstruktion von Realität in den Nachrichtenmedien. Analyse der aktuellen Berichterstattung.* Freiburg: Alber.

Schwarz-Friesel, M. (2013). "Hydra, Krake, Krebsgeschwür, Sumpf, Killer-GmbH, Franchise-Unternehmen und Nebelwolke". Perspektivierung und Evaluierung von islamistischem Terrorismus durch Metaphern im deutschen Pressediskurs nach 9/11. In M. Schwarz-Friesel & J.-H. Kromminga (Eds.), *Metaphern der Gewalt. Konzeptualisierungen von Terrorismus in den Medien vor und nach 9/11* (p. 51-74). Rübigen: Francke.

Schäfer, M. S. (2012). Online communication on climate change and climate politics: A literature review. *WIREs Climate Change*, *3*(6), 527-543. doi: 10.1002/wcc.191

Schäfer, M. S. (2016). Climate change communication in Germany. *Oxford Research Encyclopedia of Climate Science.* Retrieved from https://oxfordre.com/climatescience/view/10.1093/acrefore/9780190228620.001.0001/acrefore-9780190228620-e-448

Schäfer, M. S., Post, S., Schwab, R., & Kleinen-von Königslöw, K. (2018). Transnationalisierte Öffentlichkeit und Klimapolitik: Akteursensembles in der Medienberichterstattung von 15 Ländern. *Publizistik*, *63*, 207–244. doi: 10.1007/s11616-018-0420-6

Segerberg, A. (2012). Digital media and civil society networks: National and transnational publics. *Statsvetenskaplig Tidskrift*, *113*(1), 115-124.

Segev, E. (2010). Mapping the international: Global and local salience and news-links between countries in popular news sites worldwide. *International Journal of Internet Science*, *5*(1), 48-71. Retrieved from http://www.ijis.net/ijis5_1/ijis5_1_segev_pre.html

Segev, E., & Hills, T. (2014). When news and memory come apart: A cross-national comparison of countries' mentions. *International Communication Gazette*, *76*(1), 67-85. doi: 10.1177/1748048513504159

Shapiro, M. A., & Park, H. W. (2015). More than entertainment: Youtube and public responses to the science of global warming and climate change. *Social Science Information*, *54*(1), 115-145. doi: 10.1177/0539018414554730

Sharman, A. (2014). Mapping the climate sceptical blogosphere. *Global Environmental Change*, *26*, 159-170. doi: 10.1016/j.gloenvcha.2014.03.003

Shoemaker, P. J. (1996). Hardwired for news: Using biological and cultural evolution to explain the surveillance function. *Journal of Communication*, *46*(3), 32-47. doi: 10.1111/j.1460-2466.1996.tb01487.x

Shumate, M., & Dewitt, L. (2008). The north/south divide in NGO hyperlink networks. *Journal of Computer-Mediated Communication*, *13*, 405-428. doi: 10.1111/j.1083-6101.2008.00402.x

Simonis, G. (2017). Das Mehrebenensystem der Klimagovernance: Verflochtene Handlungsebenen. In G. Simonis (Ed.), *Handbuch globale Klimapolitik* (p. 212-259). Leiden: Ferdinand Schöningh.

Staggenborg, S. (1993). Critical events and the mobilization of the pro-choice movement. *Research in Political Sociology*, *6*(1), 319-345.

Stein, L. (2009). Social movement web use in theory and practice: A content analysis of US movement websites. *New Media & Society*, *11*(5), 749-771. doi: 10.1177/1461444809105350

Taddy, M. A. (2012). On estimation and selection for topic models. In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics.* Retrieved from http://proceedings.mlr.press/v22/taddy12/taddy12.pdf

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks*, *34*(1), 73-81.

Tesler, M. (2017). Elite domination of public doubts about climate change (not evolution). *Political Communication*, *35*(2), 306-326. doi: 10.1080/10584609.2017.1380092

Tranter, B., & Booth, K. (2015). Scepticism in a changing climate: A cross-national study. *Global Environmental Change*, *33*, 154-164. doi: 10.1016/j.gloenvcha.2015.05.003

UNFCCC. (1992, May). *United Nations Framework Convention on Climate Change.* Retrieved from http://unfccc.int/files/essential_background/background_publications_htmlpdf/application/pdf/conveng.pdf

UNFCCC. (2009, December). *Copenhagen Accord.* Retrieved from https://unfccc.int/resource/docs/2009/cop15/eng/l07.pdf

UNFCCC. (2015, December). *Paris Agreement.* Retrieved from https://unfccc.int/sites/default/files/english_paris_agreement.pdf

Van Rensburg, W. (2015). Climate change scepticism: A conceptual re-evaluation. *SAGE Open*, *5*(2), 1-13. doi: 10.1177/2158244015579723

Volkmer, I. (2019). The transnationalization of public spheres and global policy. In D. Stone & K. Moloney (Eds.), *The Oxford handbook of global policy and transnational administration.* Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780198758648.013.43

Vu, H. T., Liu, Y., & Tran, D. V. (2019). Nationalizing a globel phenomenon: A study of how the press in 45 countries and territories portrays climate change. *Global Environmental Change*, *58*, 1-9. doi: 10.1016/j.gloenvcha.2019.101942

Waldherr, A., Maier, D., Miltner, P., & Günther, E. (2017). Big data, big noise: The challenge of finding issue networks on the web. *Social Science Computer Review*, *35*(4), 427-443. doi: 10.1177/0894439316643050

Wallach, H., Murray, I., Salakhutdinov, R., & Mimno, D. (2009). Evaluation methods for topic models. In *Proceedings of the 26th annual international conference on machine learning* (p. 1105-1112). Retrieved from http://dirichlet.net/pdf/wallach09evaluation.pdf

Wessler, H., Peters, B., Brüggemann, M., Kleinen-von Königslöw, K., & Sifft, S. (2008a). Segmented Europeanization. In *Transnationalization of public spheres* (p. 40-54). Basingstoke: Palgrave Macmillan.

Wessler, H., Peters, B., Brüggemann, M., Kleinen-von Königslöw, K., & Sifft, S. (2008b). *Transnationalization of public spheres.* Basingstoke: Palgrace Macmillan.

Wessler, H., & Rinke, E. M. (2013). Öffentlichkeit. In S. Mau & N. M. Schöneck (Eds.), *Handwörterbuch zur Gesellschaft Deutschlands* (p. 637-650). Wiesbaden: Springer VS.

Wessler, H., Wozniak, A., Hofer, L., & Lück, J. (2016). Global mutimodal news frames on climate change: A comparison of five democracies around the world. *The International Journal of Press/Politics*, *21*(4), 423-445. doi: 0.1177/1940161216661848

Wolfsfeld, G. (1997). *Media and political conflict. news from the Middle East.* Cambridge: Cambridge University Press.

Yang, A. (2012). When transnational civil network meets local context: An exploratory hyperlink network analysis of northern/southern NGOs' virtual network in china. *Journal of International and Intercultural Communication*, *5*(3), 40-60. doi: 10.1080/17513057.2012.719632

Zhou, Y., Cristea, A. I., & Roberts, Z. (2015). Is Wikipedia really neutral? A sentiment perspective study of war-related Wikipedia articles since 1945. In *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation* (p. 160-168). Retrieved from http://www.aclweb.org/anthology/Y15-1019

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Koautorenschaften sowie alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemass Artikel 36 Absatz 1 Buchstabe o des Gesetzes vom 5. September 1996 über die Universität zum Entzug des aufgrund dieser Arbeit verliehenen Titels berechtigt ist.

Bern, 14. Juli 2020

Ueli Reber