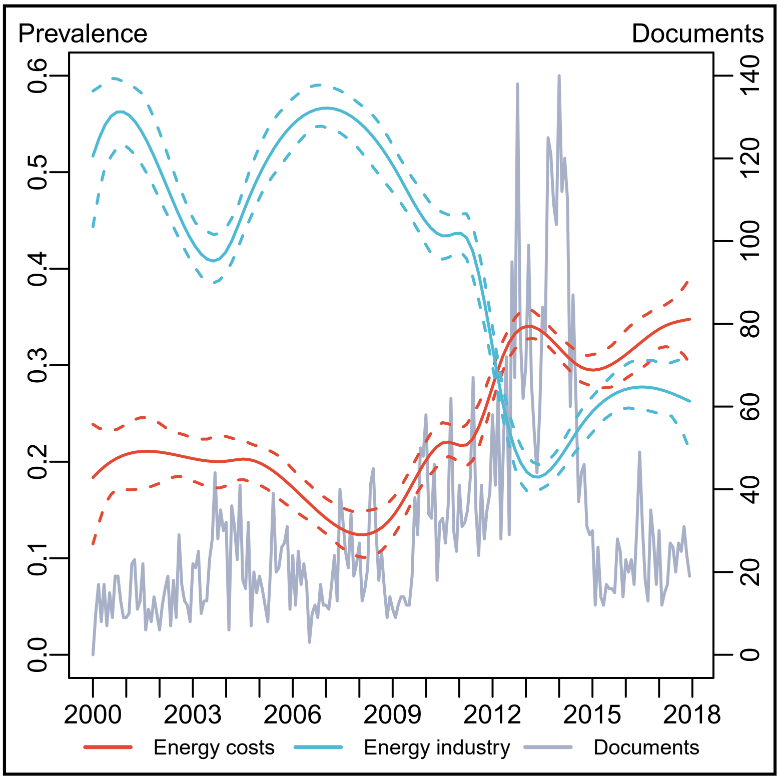
Article

Patterns

Topic Modeling Uncovers Shifts in Media Framing of the German Renewable Energy Act

Graphical Abstract



Authors

Joris Dehler-Holland,
Kira Schumacher, Wolf Fichtner

Correspondence

joris.dehler-holland@kit.edu

In Brief

A structural topic model is developed to assess the temporal dynamics of topic prevalence and sentiment in newspaper coverage of the German Renewable Energy Act. The results show that coverage followed a pattern similar to issue-attention cycles. Newspapers predominantly reported on the renewable energy industry until, in 2012, framing changed, and from then on, costs dominated the agenda. The shift in framing can affect political leverage in reaching more ambitious renewable energy targets.

Highlights

- We assess coverage of the German Renewable Energy Act in newspapers over 18 years
- Change-point analysis enables structural topic modeling to capture temporal dynamics
- We introduce the notion of topic sentiment to assess the emotional content of topics
- Positive accounts of the renewable industry shift to costs imposed on society

Dehler-Holland et al., 2021, Patterns 2, 100169
January 8, 2021 © 2020 The Author(s).
https://doi.org/10.1016/j.patter.2020.100169

Patterns

CellPress
OPEN ACCESS

Article

Topic Modeling Uncovers Shifts in Media Framing of the German Renewable Energy Act

Joris Dehler-Holland,[1,2,3,*] Kira Schumacher,[1,2] and Wolf Fichtner[1,2]
[1]Institute for Industrial Production, Karlsruhe Institute of Technology, Hertzstrasse 16, 76187 Karlsruhe, Germany
[2]French-German Institute for Environmental Research, Karlsruhe Institute of Technology, Hertzstrasse 16, 76187 Karlsruhe, Germany
[3]Lead Contact
*Correspondence: joris.dehler-holland@kit.edu
https://doi.org/10.1016/j.patter.2020.100169

THE BIGGER PICTURE Worldwide, policymakers push for a faster adoption of renewable energy technologies to mitigate climate change. Although policies that support the adoption of new technologies often have positive effects on innovation and job creation in an industry, they also involve costs borne by society. Media representations often have effects on public opinion on a policy. To understand how media reports on the German Renewable Energy Act developed over time, we developed advanced text mining models. We find that media coverage has shifted from positive accounts of the renewable energy industry toward the costs that the Renewable Energy Act imposes on society. If such patterns generalize, then public support and long-term renewable goals might be endangered. We propose that policies could be designed so that new innovative technologies, such as batteries or power-to-gas, and the optimism created by new technologies rub off onto "old" renewables to maintain broad public support.

Development/Pre-production: Data science output has been rolled out/validated across multiple domains/problems

SUMMARY

Renewable energy policies have been recognized as a cornerstone in the transition toward low-emission energy systems. Media reports are an important variable in the policy-making process, interrelating politicians and the public. To understand the changes in media framing of a pioneering renewable energy support act, we collected 6,645 articles from five Germany-wide newspapers between 2000 and 2017 on the German Renewable Energy Act. We developed a structural topic model based on a change-point analysis to assess the temporal patterns of newspaper coverage. We introduced the notion of topic sentiment to elucidate the emotional content of topics. The results show that after its enactment, optimism about renewable energies dominated the media agenda. After 2012, however, the Renewable Energy Act was more associated with its costs. Such shifts in renewable energy policy framing may limit political leverage to reach ambitious climate and energy targets.

INTRODUCTION

In light of climate change, the need to curb greenhouse gas emissions has been widely recognized, and political consensus has been reached that strong measures have to be taken.[1] However, there is also consent that the achievement of the ambitioned emission targets requires fundamental shifts in existing industries, user preferences, and markets.[2–4] The change processes required and the connected challenges in the energy sector have been commonly termed as energy transitions.[2,5]

Within energy transitions, renewable energy technologies play a pivotal role in creating an energy system with low emissions.[6] By 2016, 126 countries worldwide had implemented some form of renewable electricity policy.[7] Renewable energy technologies have achieved substantial cost reductions, yet financial risks still curb a faster development in many countries, and support policies are needed to lessen such risks.[7,8] Fast technology development, on the other hand, calls for policymakers to adapt to changing situations.[9]

The German Renewable Energy Act (Erneuerbare-Energien-Gesetz, or EEG) is one of the first policies supporting the market uptake of renewable energy technologies, and the German energy transition has attracted wide interest internationally.[10] The German EEG was enacted in 2000 and relies on a feed-in tariff or feed-in premium scheme that distributes the support for renewable electricity to many consumers through a surcharge on the electricity bill (EEG surcharge). The first significant amendments in 2004 prolonged and raised support for solar energy. A series of amendments between 2009 and 2012, among other changes, adjusted the legislation to rapidly falling solar module prices and rapid innovation in the solar sector. Amendments in 2014 and 2016 gradually introduced renewable energy tenders. In 2011, the nuclear incidents in Fukushima, Japan, caused a paradigm shift in German energy policy to phase out nuclear power by 2022. Along with the rapid expansion of wind, solar, and biomass energy, the EEG surcharge on electricity prices rose to 6.88 c€/kWh. However, various energy-intensive industries are eligible for exemptions from the EEG surcharge.

The intense dynamics of the EEG have attracted considerable scientific interest. From the scientific literature on the German energy transition policy,[11–23] we emphasize some conclusions important for our findings: researchers agree that the German energy regime has shifted from a fossil-dominated system toward a renewable system.[11,20,21] This shift has mostly been incremental, based on small policy steps rather than a radical shift in policy,[22,23] but the nuclear phase-out in 2011 represents a landmark.[11] Over time, the political discourse has de-radicalized and shifted to mainstream economical thinking,[16] and different cost narratives have accompanied the implementation of the EEG.[12] The importance of the "energy trilemma" policy goals of low environmental impacts, low energy costs, and energy security are rather stable in the parliamentary discourse. However, a decisive role in the discourse is played by a fourth goal, namely the performance of the energy industry.[21] Furthermore, policymakers' opinions[24] and public attitudes[25] toward the energy transition are strongly guided by the importance that individuals attribute to particular policy goals.

Within the political process, the media plays an important but ambiguous role. The indexing hypothesis states that media outlets closely follow governmental debates.[26] However, research also indicates that critical coverage of debates must not be based on corresponding critical elite political discourse.[27,28] On the other hand, issues perpetuated by the media may or may not enter the political agenda, depending on various circumstances.[29] McCombs and Shaw[30] have shown that the media agenda also sets public agendas by making specific issues more salient than others. On a more detailed level, attribute agenda setting posits that the media also sets the agenda regarding certain aspects of an issue.[31] Similarly, Entman argues that by selecting aspects of reality and making them more salient, framing contributes to how the media promotes certain problems, interpretations, or solutions.[32,33] On the other hand, public opinion is assumed to affect political decisions as well.[34,35] However, the multitude of existing studies has shown that all such effects are contingent.[27–29] Thus, media agenda and framing must not be congruent with elite political debates or public opinion and constitute a significant research gap, contributing to research on policy and political communication.

Some studies have analyzed newspaper coverage to understand aspects of German energy policy. Antal and Karhunmaa analyze how the German energy transition is reflected in international (non-German) newspapers and show that it is perceived differently depending on national contexts.[10] Schmid et al. showcase how advocacy coalitions in energy policy change over time and use newspaper articles to assess actors' positions and networks in the discourse.[36] Newspaper coverage that addresses specific renewable technologies has been scrutinized as well.[15,37] Rochyadi-Reetz and colleagues[38] provide an overview of international framing studies on renewable energy technologies and a comparative analysis in 11 countries. A content analysis of German newspaper coverage focusing on the evolution of coverage and the media framing of policy goals over time regarding the German EEG has not been conducted.

Due to the various amendments that the EEG has undergone, we expect that media coverage will experience shifts in the salience of policy goals and, therefore, shifts in the media framing of the EEG. In addition, technologies regularly face high expectations that often are disappointed afterward.[39,40] As the EEG supports different technologies, we are interested in how those technologies individually contribute to the media perception of renewable energies in the political context. Furthermore, as the indexing hypothesis suggests, media coverage often follows elite political discourse. We are thus interested in the representation of the political discourse and its salience in the media. Summarizing, three guiding questions structure this paper:

(1) How does the salience of the EEG and of policy goals in media representation change over time?
(2) Which topics contribute to the salience of the different policy goals?
(3) How are political debates and political activity linked to media coverage?

We show that the German EEG's media discourse follows a pattern similar to an issue-attention cycle:[41,42] the media discourse of the German EEG shifts from technology and industry optimism to emphasizing costs of the policy. This finding is surprising, as the policy discourse literature identified a regime shift from a fossil-dominated energy system to a renewable system in parallel to our findings.[11] To avoid potential bias from single sources,[43,44] we included the five largest national German newspapers in our sample, spanning the period from 2000 to 2017. The sample of 6,645 articles makes manual content analysis costly.[45]

For the analysis of large corpora, unsupervised topic models such as latent Dirichlet allocation have successfully provided insights into the contents of texts.[46,47] More recent developments of topic models have proposed to include metadata in topic models as covariates. The structural topic model[48–50] (STM) extends topic modeling by assuming a relationship between covariates and corpus content. Those structural features make STM a suitable candidate for our endeavor. We exploit the assumption that textual contents also depend on external factors in that we develop a detailed model of the temporal dynamics of newspaper coverage of the German EEG. STM has been successfully applied to test hypotheses on covariate-content relationships,[48,51–53] but rarely used as a device for detailed time-series analysis. We apply a topic modeling pipeline,[54] including preprocessing, lemmatization, corpus reduction, structural topic modeling, and, finally, careful validation. We make two

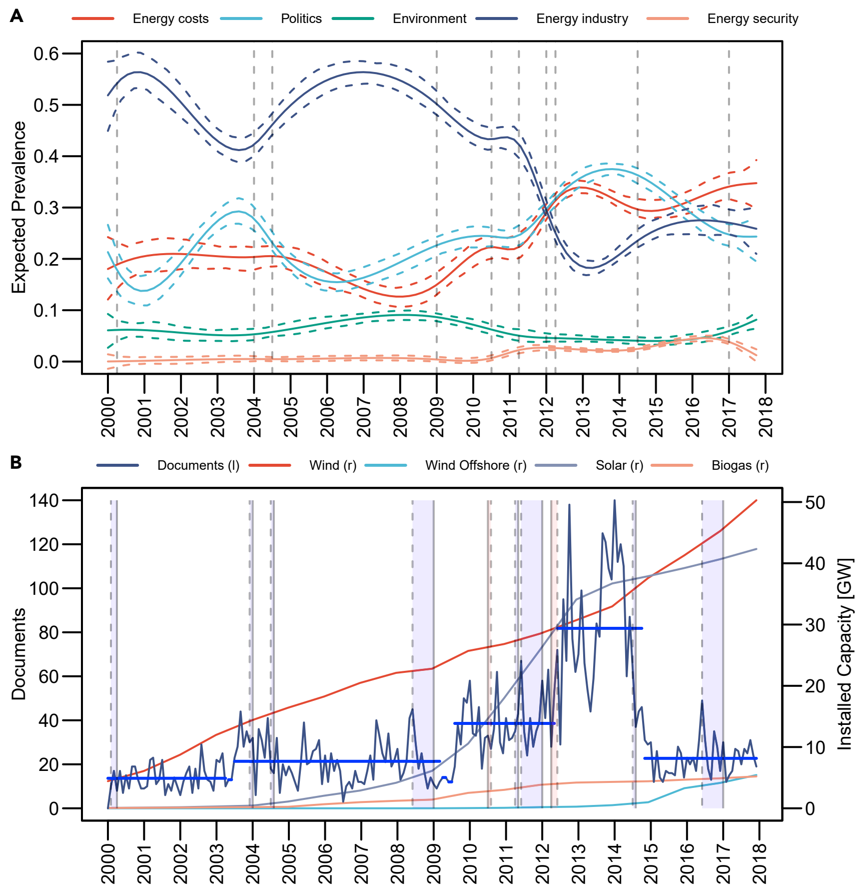


Figure 1. Prevalence of Policy Goals and Politics and Salience of the EEG over Time

(A) Evolution of policy goal coverage. The graph shows how the renewable energy industry loses prevalence over time, in contrast to the increasing prevalence of costs associated with the EEG. Curves are natural spline models as described under Experimental Procedures. Dashed curves indicate the 95% quantile. Model estimates can be found in Table S3.

(B) Left axis (l): change-point analysis of the number of documents per month covering the German Renewable Energy Act (EEG) from 2000 until 2017. Breaks in the horizontal line indicate change points, i.e., a change in mean and variance. Vertical lines indicate a change in legislation. Dotted lines are the point in time when that legislation passed Parliament. Surfaces mark the span between decision and entry to force of the legislation. Red surfaces indicate that the policy change was made retroactively (notably in 2010 and 2012). Right axis (r): installed renewable energy capacity is plotted.[60]

methodological contributions that integrate the needs of temporal content analysis into topic models and are of interest to data scientists pursuing similar objectives.

First, choices on the temporal model must be made. We choose a natural spline model to assess the temporal dynamics of newspaper coverage. Inspired by interrupted time-series analysis,[55] we propose to set spline knots based on a change-point analysis[56,57] of overall article counts to sufficiently introduce external information into the content analysis. In our case, a detailed analysis reveals that such decisions can also be made based on domain expertise: the identified change points coincide with amendments of the underlying policy.

Second, the emotional content of texts is essential to understanding media coverage and its potential effects. A classical tool to assess emotional content is sentiment analysis.[58,59] Most sentiment analysis techniques are based on assessing single words, sentences, or documents and describe sentiment as a polarity score between positive and negative.[58] However, our interest lies particularly in assessing sentiment associated with the topics that we identified with our STM. To that end, we introduce the notion of *topic sentiment*, which neatly builds upon STM's description of topics as a distribution of words. Using sentiment lexicons, one can calculate the expected sentiment of a topic, conditional to, of course, a certain lexicon. The concept of topic sentiment captures the overall qualitative impression of the emotional content of topics well.

RESULTS

The presentation of the results is structured along with the three guiding questions: we first analyze how the media representation of the EEG changes over time concerning the four policy goals of limiting the environmental burden of energy supply, energy security, limiting energy costs, and energy industry prosperity. In addition, we separate political activity as an individual category and link it to the changes in attention to policy goals. We assess newspaper article counts across the entire period from 2000 to 2017 to understand the EEG's overall media attention. The remainder of the Results section analyzes the four policy goals with attention to politics in detail, e.g., by resolving them by the different renewable energy technologies or cost drivers.

Some preliminary remarks on our methodological approach might help in understanding the results (further details under Experimental Procedures). Structural topic modeling assumes that a pre-specified number of semantically interpretable themes or topics defines each text's content and the whole collection. Each document comprises different topics to varying shares. We identify 49 topics that we assigned content-wise to the four policy goals and political activity. We assess the share of specific topics in the whole dataset (*prevalence*) in dependence on time. The time-series results of our STM are used to analyze the change in prevalence of factors that contribute to the media representation of the four main policy goals and politics. By conducting a sentiment analysis, we assign a sentiment score $-1 \ll ts_t \ll 1$ to each topic t to understand its emotional content. Furthermore, we assess a choice of articles by close reading to understand the contents of each topic and to define labels for each topic.

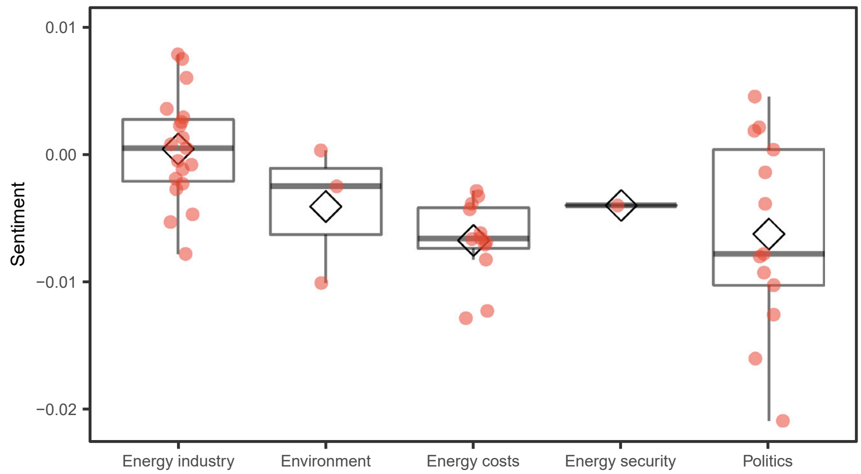


Figure 2. Topic Sentiment per Policy Goal

We find that the industry goal is covered most positively, while costs are discussed in negative contexts. Pink dots represent the topic sentiment. Whiskers display 1.5× the interquartile range. The lower and upper hinges correspond to the first and third quartiles, respectively. Transparent diamonds display the mean sentiment per category; the centerline is the median.

Coverage Shifts from Technology Optimism to a Cost Discourse

From 2000 to 2011, the EEG is debated most frequently in the context of the energy industry. In 2004, we notice a dent in industry attention, just before the 2004 amendments of the EEG (Figure 1A). Political activity gains a larger share of coverage during that time. Whereas interest in the EEG's industry effect peaks in 2007, coverage of topics related to energy costs is lowest. Renewable energy technologies are increasingly installed. From 2007 onward, costs gain prevalence, while the renewable industry loses attention until, in 2011, costs and political debates exceed industrial optimism and subsequently dominate the media agenda. Notably, the switch in attention to policy goals entails an increased coverage of political activities in 2013 and 2014 (Figure 1A). In addition, the number of articles that mention the EEG almost doubles from 2011 to 2012, just when cost issues become the most important policy goal on the media agenda (Figure 1B).

Figure 1A shows that the prevalence of the four policy goals varies greatly. Figure 1B demonstrates that the intensity of attention awarded to the EEG differs substantially, peaking between 2012 and 2014, directly after the change in framing. In addition to the intensity of coverage, the positive or negative connotation of newspaper coverage yields important information. We thus conducted a sentiment analysis to assess the emotional connotations of the topic's vocabulary distributions.

The sentiment analysis shows that, in general, the EEG tends to be presented in a negative tone, with the majority of topics having negative topic sentiment. This result is not particularly surprising, given the general negativity trend in news coverage attributable to negativity biases in human cognition.[61] Notably, energy industry topics are covered in a relatively positive tone (Figure 2): topics that relate to industry goals have a more positive sentiment than sustainability or cost topics. The lowest topic sentiment is attributed to topics that cover conventional energy production (Table 1). The qualitative assessment affirms that renewable energy technologies are particularly associated with positive accounts of job creation, industry leadership, and innovativeness. All energy cost topics are discussed in a negative tone. They discuss the EEG surcharge that must be borne by consumers, its effects on power prices or its distribution, and the overall costs of the energy transition (Table 1). Moreover, the analysis reveals that topics relating to the political process are rather diverse. They range from the overall minimum in the debate on cutbacks of solar feed-in remuneration (topic 42, Table 1) to topics that show a positive connotation.

The shift in attention to policy goals and topic sentiment shows how the media framing of the EEG switched from emphasizing economic gains that renewable technologies bring to an emphasis on costs that have to be borne by society, particularly households, while parts of the industry can avoid them. To further understand this shift in the framing of the EEG, it is interesting how the different attributes of the policy goals changed. In the next sections, we will analyze the different policy goals and the apparent change in media representation in more detail.

Energy Industry Coverage Dominated by Solar

Until 2012, the energy industry and technology coverage dominated the media perception of the EEG. The EEG supports different technologies, such as solar, wind, and biomass electricity production. It is thus a natural question, which technology received the most attention over time? A closer look into the energy industry category reveals that the topic model was able to differentiate between the different renewable energy sources, but also between conventional energy carriers (Table 1). We will discuss the energy industries in some detail (Figure 3) regarding topics that mainly drove attention during specific periods. We reduce the presentation of results to the technologies that contributed most to the shift in framing in 2011.

Splitting up the energy industry coverage by technologies, we find that solar power contributes the largest share of coverage (Figure 3A). While the run-up of the EEG amendments in 2004 co-occurs with a drop in solar energy coverage in 2003, the amendments reinforced political support for solar energy, and coverage grows fast afterward. The solar industry was depicted as a crucial future industry in Germany, and job creation in the sector contributed to its positive image (topic 34). Similarly, stocks of German solar companies were booming. However, reports also reflect the uncertainties that EEG amendments triggered for the companies. Most notably, reports show that solar stocks plummeted before the amendments in 2004, reflecting significant uncertainties whether the support would be maintained (topic 39).

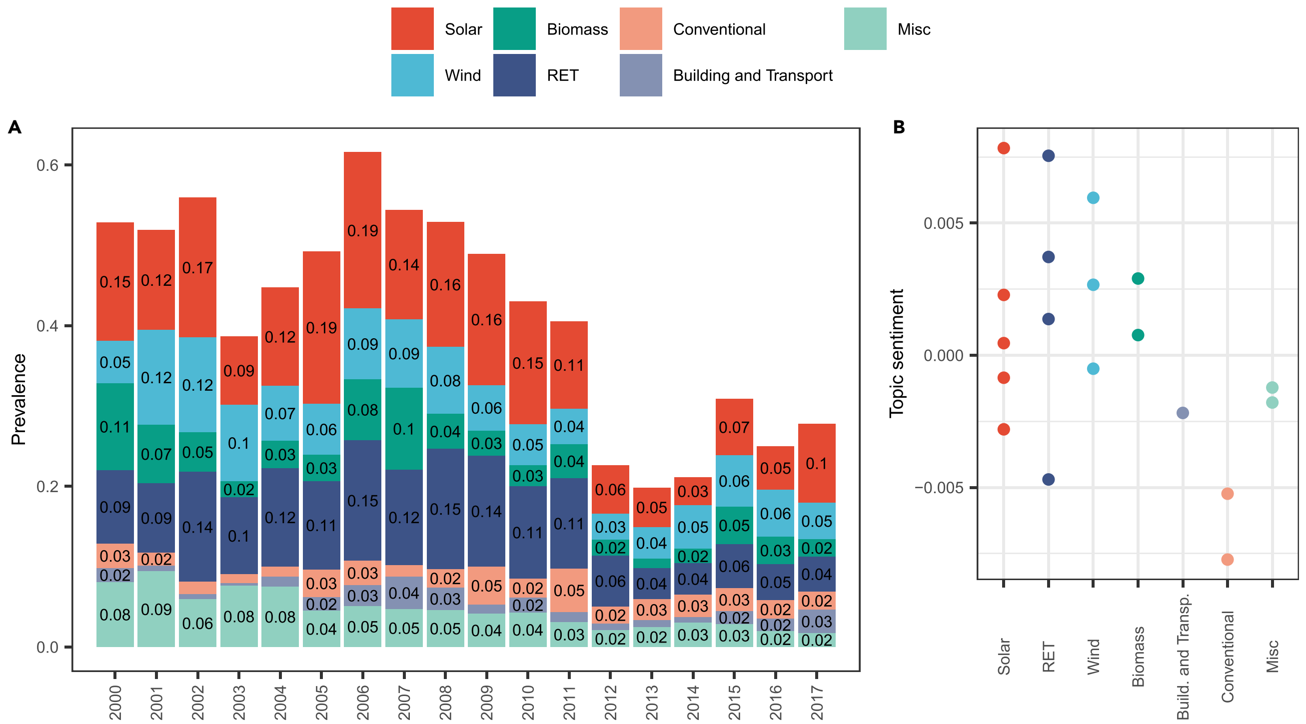


Figure 3. Prevalence and Topic Sentiment of Energy Industry Topics

(A) Split of energy industry category. Solar energy and renewable energy technologies (RET) contribute most to the coverage of the energy industry.

(B) Topic sentiment of the different technologies.

From 2004 onward, newspapers increasingly point to the growing international competition in the solar panel market, particularly from Asia (topic 33). Fears rise that German companies are losing their competitive advantages. Finally, the solar industry coverage sharply declines after 2011 (Figure 3A), when the German solar industry drastically loses market shares, and panel producers go bankrupt.[62] After 2012, solar energy production regains attention (Figure 3A), when it becomes clear that solar rooftop self-consumption, possibly in combination with batteries, becomes a profitable use case[63] (topic 48). The topic sentiment shows enthusiasm for solar energy, but also negative values when solar stocks are discussed (Figure 3B and Table 1).

Another sector, which attracts attention in the early years of the EEG, is the wind energy sector. Wind installations grow steadily (Figure 1A), and the industry is euphoric on its rapid development. Pilot plants are constructed as test cases for large new turbines (topic 8). Topic sentiment analysis shows that the language used concerning wind energy topics is predominantly positive (Figure 3B). The coverage of biomass energy generation exhibits several peaks (Figure 3A). Biomass energy production aroused criticism due to land-use conflicts that are also apparent in the coverage of transport due to the deployment of biofuels.

Other than these rather prominent technologies, renewable energy technologies are often discussed jointly, without reference to a specific technology (Figure 3A). Articles emphasize the positive economic effects of renewable energy production, such as job creation or international technology leadership in the sector (topic 32). Positive aspects of renewable technologies are also reflected in the vocabulary used to discuss renewable technologies. Compared with conventional power sources, topic sentiment is higher and predominantly positive (Figure 3B). Reports on innovative development of different technologies or energy storage are discussed with reference to the EEG, contributing to the media perception of innovativeness (topic 29).

All renewable technologies exhibit a drop in attention from 2011 to 2012. Reports on the solar industry in particular contribute to the change in framing of the EEG. The decrease in attention co-occurs with the rapid increase in attention to the costs of the EEG (Figure 1). We will now turn to a more detailed discussion of energy cost coverage to understand those changes better.

Changes in Energy Cost Coverage

The introduction of the EEG in 2000 falls in a phase of the rapid restructuring of the German energy sector. In 1998, the German government liberalized the electricity market. The liberalization still plays a significant role in the EEG discourse (Figure 4A, topic 12), in which the German energy industry association argues that the state skimmed off the profits of market liberalization by increasing the EEG surcharge.[64] Over time, the influence of the liberalization of the German electricity market and the related power price development on newspaper coverage ceases. Closely related to the liberalization, the abuse of market power of the vertically integrated utilities in the new electricity market becomes a steady topic (Figure 4B, topic 26). Utilities were accused of exerting their market power by the

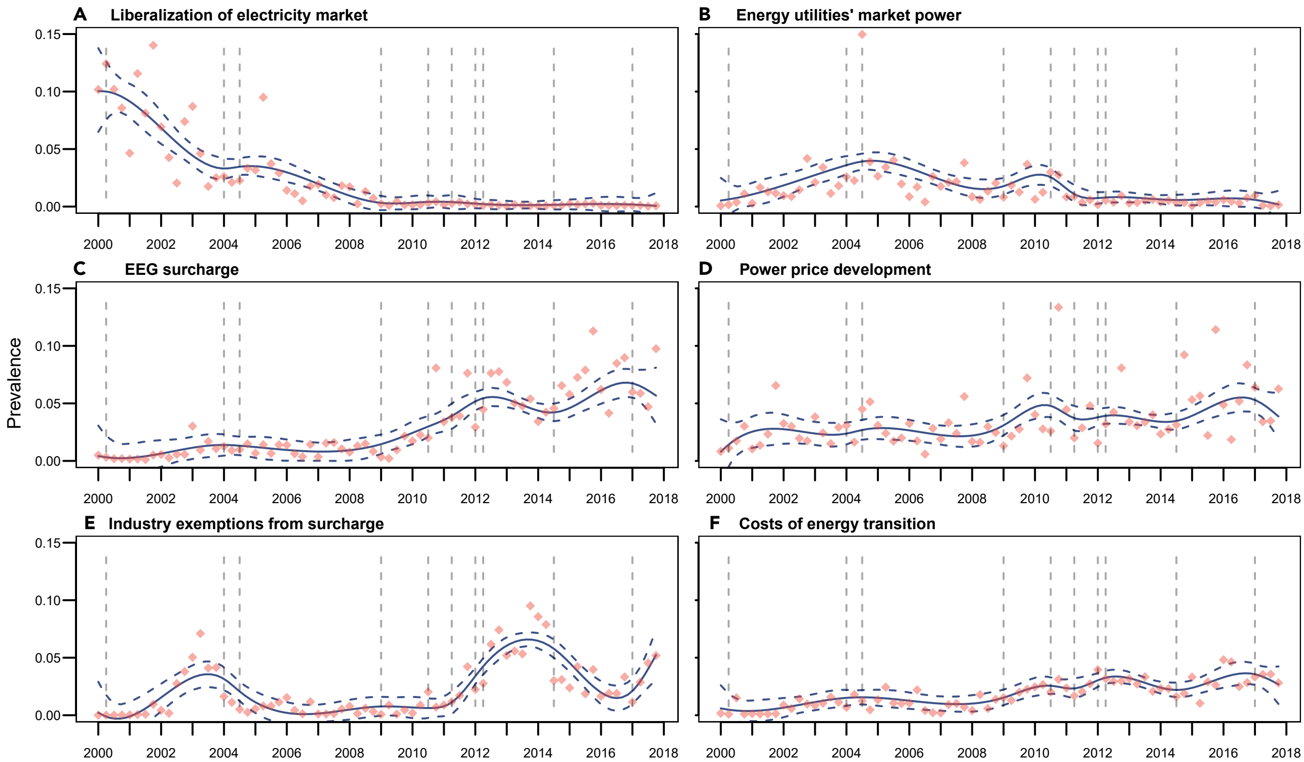


Figure 4. Prevalence of Topics Associated with Energy Costs

Prevalence is measured as the mean share of the focal topic with regard to all topics in the model. Whereas, in the early years, costs were associated with utilities' market power (B) and the EEG surcharge (C) was said to level out gains from market liberalization (A), with the increase in the EEG surcharge, it becomes a separate topic, and the power price (D) is reported increasingly. (E) Industry exemptions from surcharge; (F) costs of energy transition. Pink diamonds depict the average prevalence of the topic per quarter. The blue curve is a linear model based on natural splines to depict the trend, as described in the Experimental Procedures. Dashed blue curves indicate the 95% confidence interval, calculated by drawing Monte Carlo simulations from the topic distribution and fitting models to the simulations.[49] Dotted gray vertical lines indicate policy amendments of the EEG, as in Figure 1. Model estimates can be found in Table S4.

imposition of grid fees and pass-through tariffs. The topic disappears from the agenda with the completion of the unbundling process.

Whereas these topics disappear over time, the EEG surcharge and the general power price development enter the media agenda with force (Figures 4C and 4D, topics 4 and 19). The reports on the EEG surcharge contribute to the shift from industry contexts, in which the EEG was discussed previously, to the context of consumer prices and costs. This increase in coverage aligns with an increase in the household surcharge from 1.3 c€/kWh in 2009 to 5.28 c€/kWh in 2013. The exemption of energy-intense industries and the distribution of the EEG surcharge are discussed relative to amendments of the EEG (Figure 4E, topic 23). The discourse on the surcharge distribution to the industry also displays a strongly negative sentiment score (−0.013); in addition, overall costs of the energy transition increase in prevalence: expenditure for renewables within the EEG, but also their grid integration (Figure 4F, topic 40), is criticized. These topical changes contribute to the shift in framing of the EEG. From 2011 onward, the costs of the EEG, with particular emphasis on end consumer costs and their distribution, dominate the media agenda, replacing more positive accounts of the renewable energy industry.

Environmental Goals and Energy Security Face Lowest Coverage

Despite the demonstrated attention on cost considerations, industry matters, and political action, we turn the focus of this last section to environmental goals and energy security.

The policy goal of limiting the environmental burden of electricity supply is reflected in three topics. The EEG is discussed in the framework of climate change mitigation (topic 45). More systematic issues of transitions to sustainability in energy, transport, and building sectors are discussed, wherein the fragmented German policy is also criticized (topic 47). From the introduction of the European Emissions Trading Scheme (ETS) in 2005, the interaction of EEG and ETS is prevalent, and the EEG is said to counteract international agreements (topic 21).

Energy security is mainly reflected in the debate of missing transmission lines from northern Germany, where much of the wind energy capacities are installed, to the south, where consumption is higher due to industrial centers (topic 9). As a side topic, generation adequacy also enters the general debate on the market design (topic 13), albeit the focus here is on integrating renewables into markets and their cost efficiency.

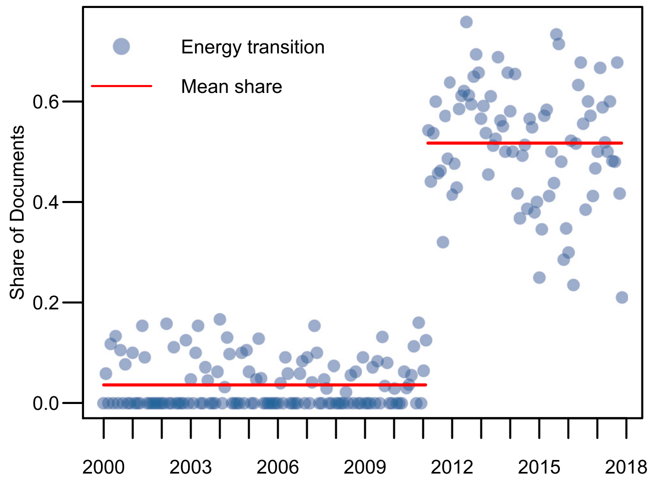


Figure 5. Usage of the Term "Energiewende" (Energy Transition)
In red, the mean share of documents per month that use the term before and after March 2011 is depicted.

Representation of Political Activity in Coverage

While the energy industry and energy costs receive a high level of attention, politics takes action to mitigate pressures that arise with regard to those policy goals. We therefore briefly investigate how political action is reflected in newspaper coverage. The analysis of topics devoted to the political process reveals that four major amendment debates can be distinguished and that the policy goals of energy cost control and industry prosperity are reflected in the coverage of political action. After the introduction of the EEG, Energy Concept 2004 (topic 18) was the first significant amendment to it. It prolonged and raised solar power support, introduced exemptions for energy-intense industries, and introduced renewable energy industry development as an explicit policy goal. The amendments between 2009 and 2012 (topic 42) struggled with the rapid deployment of solar power and reduced remuneration of new installations while raising renewable energy targets. In 2012 and 2013, an intense debate broke out on how to cap the electricity price development (topic 17) that was perceived to be driven by the EEG surcharge. Measures to cap solar installations were introduced. The amendments in 2014 and 2016 (topic 11) introduced a capacity auctioning scheme to increase competition for lower feed-in remuneration. Figure 1 shows how the four debates contribute to the coverage of political action of the EEG in the respective years. The high salience of the EEG between 2012 and 2014 is strongly related to the latter two debates (Figure 1B; topics 17 and 11) directly after the change in framing.

Along with the shift in policy goal attention we described above, the nuclear accidents in Fukushima, Japan, introduced changes to German energy policy and the decision for a rapid nuclear phase-out in 2011.[65] Directly after the nuclear accidents, the term "Energiewende" (energy transition) suddenly gained prominence (Figure 5). On March 29, 2011, 3 weeks after the nuclear accidents, Chancellor Angela Merkel demanded a faster shift to renewable energies and introduced the term Energiewende.[66] In June 2011, the government proposed the nuclear phase-out and framed the Energiewende as a Herculean task that would also bring tremendous opportunities for the future.[67] Since then, the term is commonly used in relation to the EEG. Although the term suggests a more substantial commitment to renewable energies, in fact, the renewable energy targets that already were made in the energy concept of 2010 have not been changed. In addition, the introduction of market measures was already decided in the energy concept of 2010. Our results show that newspapers quickly adopted and reproduced the new framing of German energy policy.

DISCUSSION

Structural Topic Modeling for Temporal Content Analysis

Our study found that advanced text modeling can yield valuable insights into news coverage of political instruments. In particular, our approach is able to assess temporal patterns of coverage in detail. The ability to analyze larger corpora has the advantage that multiple sources can be included in the sample, avoiding possible bias by a limitation of scope, and makes content available for rigorous statistical analysis. Comparing our findings with manual assessments of the political evolution of the German EEG show that those patterns can also be found in the policy literature.[11,12,23]

Our research is valuable for researchers and data scientists who pursue quantitative methods in content analysis. Our findings highlight that by combining change-point analysis and STM, we can define topic models sensitive to topic changes in highly dynamic settings. Furthermore, the notion of topic sentiment integrates well with topic modeling, as it uses the representation of topics as distributions of words. In manual content analysis, it is often highly labor intensive to assess the emotional content of topics by close reading.

However, an advanced data science approach cannot wholly avoid manual assessment.[68] Domain knowledge helps one to aggregate results to categories compatible with policy analysis and interpret the findings. An avenue for future research on topic sentiment may include n-grams into the analysis, as topic modeling does not harvest information encoded by the word order in a text but relies on the bag-of-words assumption. The field of natural language processing progresses quickly, and promising models are being developed. A further limitation arises from the fact that we sampled only the five most important national newspapers. The German media system is highly diverse, with the majority of print media being local newspapers. However, the newspapers we analyzed belong to the ones most cited in other media sources, and thus serve as a reference for other journalists. We thus assume that framing of the EEG in local coverage will not diverge drastically from our findings. In the following, we will discuss our results from a domain-specific standpoint.

The Attention Cycle of the EEG

Reaching ambitious climate goals requires substantial political action on introducing technologies with lower greenhouse gas emissions. Media content analysis can provide essential insights into how political action and policy change are framed by the media and communicated to the public. In that regard, our results show that the media representation of the German EEG has witnessed a frameshift from positive accounts of the renewable energy industry toward the costs that the EEG imposes on society

Table 1. Topics Assigned to Policy Goals and Political Action

Policy Goal/Category	No.	Topic	Topic Sentiment	Prevalence (%)
Energy industry[a]	5	conventional power plant profitability (C)	−0.0053	1.12%
	7	organic matter for energy production (B)	0.0029	1.48%
	8	wind power installations (W)	0.0026	2.02%
	10	business reports (M)	−0.0019	2.25%
	16	buildings and transport (BT)	−0.0023	1.37%
	20	bioenergy (B)	0.0008	1.84%
	22	SolarWorld (S)	−0.0008	1.00%
	29	innovative electricity technologies (RET)	0.0013	1.53%
	32	renewable energy shares and targets (RET)	0.0036	3.59%
	33	competitiveness of German solar industry (S)	0.0005	2.38%
	34	solar industry boom (S)	0.0079	2.00%
	36	wind energy market (W)	0.0060	2.04%
	37	offshore wind parks (W)	−0.0005	1.73%
	38	nuclear energy (C)	−0.0078	1.52%
	39	solar stocks (S)	−0.0027	1.89%
	41	siting of industry and energy plants (M)	−0.0011	1.57%
	43	international activity of energy industry (RET)	0.0075	1.27%
	46	investments in renewable energy projects (RET)	−0.0047	1.71%
	48	rooftop solar business models (S)	0.0023	2.31%
Energy cost	3	marketing of clean power	−0.0033	2.66%
	4	power price development	−0.0043	3.80%
	12	liberalization of electricity market	−0.0065	1.25%
	13	market integration	−0.0061	2.06%
	14	choice of electricity provider	−0.0039	1.61%
	19	EEG surcharge	−0.0028	3.88%
	23	industry exemptions from surcharge	−0.0128	3.47%
	26	energy utilities' market power	−0.0069	1.20%
	27	public charges and taxes	−0.0123	1.63%
	35	industries losing EEG-privileges	−0.0082	1.31%
	40	costs of energy transition	−0.0071	2.18%
	44	periodic reports on fiscal and financial regulation changes	−0.0066	0.85%
Politics	1	coordination of energy transition	0.0019	2.24%
	6	EU commission state-aid cases	−0.0126	2.47%
	11	EEG amendments 2014 + 2016	−0.0078	2.60%
	15	politics of the SPD and CDU/CSU political parties	0.0004	2.02%
	17	electricity price cap	−0.0093	2.35%
	18	energy concept 2004	−0.0102	1.87%
	24	political power structures	0.0022	1.43%
	25	election campaigns	−0.0039	1.80%
	28	legislative process	−0.0080	3.00%
	30	profiles of politicians and entrepreneurs	0.0046	2.05%
	31	EEG remuneration	−0.0014	3.33%
	42	EEG 2009–12 reforms–solar remuneration	−0.0209	2.51%
	49	complaints of interest groups	−0.0160	0.55%

(*Continued on next page*)

Table 1. *Continued*

Policy Goal/Category	No.	Topic	Topic Sentiment	Prevalence (%)
Environment	21	emission trading system	−0.0101	1.87%
	45	climate change mitigation	−0.0025	1.31%
	47	sustainability transition	0.0003	2.35%
Energy security	9	grid extension	−0.0040	1.72%
Common speech	2	common speech	−0.0008	4.01%

Prevalence indicates the share of a topic in the entire corpus. The expected sentiment per topic is described as a number between −1 and 1. Topic 2 (common speech) is a particular case, as the topic is defined by the style of articles. Highly associated articles report interviews or letters that are not strongly edited and contain common speech. Within categories, topics are ordered by topic number that is assigned arbitrarily.

[a]For further analysis, energy industry topics have been attributed to different technologies (Figure 3): B, biomass; BT, building and transport; C, conventional; RET, renewable energy technologies; S, solar; W, wind; M, miscellaneous.

in 2011. Over time, attention to the industry declined, while that to the prevalence of costs increased. The decrease in interest in the energy industry can be attributed to the demise of German photovoltaics (PV) producers and increasing international competition. At the same time, other renewable technologies also lost attention. On the other hand, the EEG surcharge doubled in only a few years and let power prices rise. This can explain the increasing prevalence of costs during the same period. The shift of goal prevalence also entailed a fierce debate on how to limit the additional costs borne mostly by end consumers (topic 17).

Shifts of attention toward public issues along with changes in the framing from enthusiasm about new solutions toward the realization of costs have been acknowledged in the study of political communication repeatedly and popularly generalized as the *issue-attention cycle*.[41,42] However, the finding that public policies as the "object" of media attention may face similar patterns is new. The existing results on issue attention, together with technology hype[39,40] implying that new technologies often induce high expectations that are disappointed after a while, suggest that these patterns might generalize to other technology policy instruments applied worldwide to foster technological change. For sustainability transitions with their long time horizons of several decades,[5] such a change in framing may have severe consequences. Scholars have suggested that technological development opens up windows of opportunities for ambitious climate and energy policies, whereas usually, new technologies or policies are framed as expensive.[69] Our results indicate that such windows may also close and that policies in the long term may again be threatened by cost concerns, limiting political leverage. As media accounts of a policy are arguably only one part of the picture, and causal inference to political decisions or public opinion is difficult, we compare our findings with those of policy scholars and surveys of public opinion on German energy policy.

According to Schmidt et al.,[21] in the German parliament from 1998 to 2002, conservative parties argued against renewable technologies, referring to the cost of energy, while center-left parties who held the government at that time argued in favor of the policy based on positive industry effects.[21] Our research shows that during that time, the energy industry was most prevalent in media accounts. During the period between 2009 and 2013, the governing conservative parties also picked up the argument of positive effects for the industry.[21] However, the majority of conservative arguments were still negative and referred to the costs of the policy, while opposition parties argued in favor of renewable energy. Also, Lauber and Jacobsson[12] and Hoppmann et al.[23] observed that from 2009 onward, cost concerns increased in parliamentary and political debates, while benefits of the policy were more prevalent before that, in line with the change in media framing we observed. Our analysis of the coverage of specific political debates showed that the media closely monitor them, and framings of core policy concepts introduced by the government, such as the Energiewende, are adopted and reproduced quickly. From the comparison of policy studies and our results, we conclude that media framing of the EEG broadly mirrored the arguments of the governing parties, evidencing support for the indexing hypothesis in this case. Opposition arguments seem to be reflected less often in media coverage. However, we cannot conclude whether one side caused the other or the coincidence is a process of mutual reinforcement.

The public opinion of the German energy transition has also repeatedly been surveyed. A series of surveys from 2013, 2015, and 2017 shows that electricity costs have gained in importance over the years: in 2017, it was the most crucial aspect of the energy transition, while it was considered least important in 2013.[70] Yearly surveys from 2017 to 2019 show that this trend continued.[71] This is remarkable, as the EEG surcharge rose from 1.3 c€/kWh in 2009 to 5.28 c€/kWh in 2013, while its increase slowed down considerably afterward. Although the rapid increase has gone hand in hand with the prevalence of costs in newspapers, public opinion seems to lag behind both developments and even worsens while the surcharge is stable. The EEG's framing in media and politics as costly for end consumers has thus preceded the actual turn in public opinion that the energy transition is too expensive. In line with the agenda-setting hypothesis, one may speculate that the change in framing and the subsequent increased salience of the EEG and the surcharge between 2012 and 2014 (Figure 1) contributed to the shift of the public agenda.

The detailed assessment of industry goal topics revealed substantial differences in coverage of the different technologies. We have noticed that the solar energy industry received the most attention, even though it is argued to contribute less to the achievement of policy goals than wind energy.[72] The close attention can be attributed to at least four reasons: (1) Solar energy has a high appeal to large shares of the population, as it offers investment opportunities for household solar installations, but also in terms of solar company stocks. (2) The solar PV market was highly dynamic, and it was presumed that Germany had an advantage over international competitors, an assumption that turned out to be false.[62] (3) Compared with wind, solar PV technology development appears to be less complex.[73] (4) The rapid development in solar markets and technology induced high policy dynamics and political learning.[23] Those four factors contributed to the fact that media devoted more attention to solar PV in combination with the EEG. Interestingly, all technologies follow the same overall trend, while each technology follows discursive subcycles.

The analysis of the topics contributing to the media perception of energy costs shows that the EEG surcharge and power price development contribute most to the change in framing. It is likely that attention to both factors can be explained along the same lines as the greater attention to solar power: the EEG surcharge on the power prices directly affects the audience of newspapers. We also have shown that the framing of costs changed over time. In the early years of the EEG, EEG surcharge and grid costs were contrasted with the efficiency gains due to market liberalization, and the market power of the big utilities was blamed for being responsible for higher prices. Later, the high share of taxes and levies in the electricity price was emphasized.

Conclusions and Policy Implications

In general, our results provide insights into how frames changed over time in the media representations of an important piece of energy legislation. As the media is an important stakeholder contributing to political discourse by filtering political news for the larger audience[74] and informing policymakers,[75] policymakers should be aware of media effects when designing policies.

Our findings point to the question of how to refinance support schemes. Although surcharges might have positive effects on energy efficiency, as they increase electricity prices, on the other hand, they may increase friction with the public as they distribute costs to a high number of voters, an issue also faced by CO_2 taxes.[76] As a consequence of ongoing cost debates, Germany introduced renewable capacity auctions[19] and agreed on a gradual decrease in the EEG levy going hand in hand with the imposition of CO_2 charges on fuels from 2021 onward.[77] However, policies might face a phase of realization of costs either way.[41,42] If that is the case, then initial political support for industries must create actor and network effects strong enough to withstand upcoming societal and economic pressures.

Our results have shown that media coverage has shifted from a framing of the EEG that highlights industry spillovers to a framing that emphasizes the costs imposed on society. Our discussion showed that public opinion followed suit, thereby potentially limiting political leverage for politicians who aim to foster the future expansion of renewables, an expansion that will be needed to reduce the climate impact of energy provision. Thus, to maintain public support, media discussion should be redirected to focus on benefits for the industry instead of costs. The recently rising attention to new business models and technologies, such as rooftop solar self-consumption with batteries, electric vehicles, or hydrogen fuels, might point to a potential direction for policy making if cost narratives threaten transition policies. Geels et al.[4] argue that innovation policies (such as the EEG) can "galvanize public enthusiasm around positive visions, and build social and business coalitions that in the longer term may support stronger climate policies." As technology expectations often follow cyclic patterns,[39] these could be used to spur public enthusiasm for "conventional" renewables and policies by supporting combinations of renewable installations with storage facilities such as batteries or hydrogen production. It might be a way to kill two birds with one stone: on one hand, the innovative appeal of new technologies can rub off onto renewable technologies. On the other hand, storage technologies are desperately needed to balance intermittent resources and foster sectoral coupling.

Particularly for Germany, where the energy transition is perceived to have lost its momentum,[78] with new wind installations and auction participation decreasing in 2018 and 2019, while added capacity is needed to reach climate and energy goals, a new wave of dynamic technology development could be needed to "galvanize public enthusiasm around positive visions."

A remark concerning our assumptions can help to place our conclusions in the context of research on media effects with its various competing results on the causal relationships between the media, politics, and the public. The discussion of our results has shown that such relationships exist in this case, but by no means were we able to establish such links statistically. Our conclusions rely on the assumption that negative coverage also influences public opinion and, further, that public opinion affects political decisions in the long run. Occasionally, policymakers may act against public opinion, or public opinion might not follow the media agenda, but in general, research shows that it is reasonable to hold such assumptions.

EXPERIMENTAL PROCEDURES

To assess the legitimacy of the EEG, we analyzed 6,645 articles from five major German newspapers applying time-series change-point analysis[56] and structural topic modeling.[50] As a result, we obtain a fine-grained time series of the prevalence of topics. A qualitative content analysis of the different topics supports the time-series analysis.

Resource Availability

Lead Contact

Further information and requests for resources and materials should be directed to and will be fulfilled by the lead contact, Joris Dehler-Holland (joris.dehler-holland@kit.edu).

Materials Availability

This study did not generate new unique materials.

Data and Code Availability

All articles analyzed in this study are available through the newspaper databases. The data were used under license for the current study, and so are not publicly available for free. However, during the peer-review process, the data were available for the reviewers, according to German copyright law (§60d UrhG), from the corresponding author upon reasonable request. The results of the text modeling (STM) that constitute the base for time-series analysis are available from the corresponding author upon reasonable request.

All algorithms used in this study have been cited in this section, are publicly available, and are well reported. However, all scripts used for the analysis are available from the lead contact upon reasonable request.

Data Collection

The study is based on articles covering the period between January 2000 and December 2017 that appeared in five nationwide German newspapers. The period covers the enactment of the German EEG in 2000 as well as all of its major amendments at the time of writing this paper. The newspaper choice covers the five non-tabloid titles with the largest circulation in Germany. We did not include online versions, as readership and structures might have evolved drastically, given the time period covered. The political orientation of the newspapers covers all facets from moderately left to moderately right and includes a financial newspaper (*Handelsblatt*). By including different newspapers, we aimed at capturing as much as possible of the variance of the media agenda that might have reached the population, as opposed to concluding to a media agenda from a single source.[44]

The articles from two newspapers analyzed in this study were collected from the LexisNexis academic database (*Die Tageszeitung* and *Die Welt*), the *Handelsblatt* was collected from GBI-Genios (wiso-net.de), and the *Frankfurter Allgemeine Zeitung* and *Süddeutsche Zeitung* were retrieved from the newspapers' own databases. For all newspapers, the query "*Erneuerbare-Energien-Gesetz OR EEG OR Einspeisevergütung OR Stromeinspeisevergütung*" ("Renewable Energy Act OR EEG OR Feed-in remuneration OR power feed-in remuneration") was searched and the results were stored. The inclusion of the abbreviation EEG also conveniently captured word combinations such as EEG-*Umlage* (EEG levy) or EEG-*Vergütung* (EEG remuneration).

The results of the original search included 7,839 articles. A first analysis revealed that some articles did not cover the EEG, as EEG also is an abbreviation for *Elektroenzephalografie* (electroencephalography). The articles were filtered locally to exclude those articles. Further, the pre-analysis revealed that some articles in the databases were duplicates. Using proximity measures of text resemblance (Levenshtein distance), the database was consolidated further. In total, 6,645 articles were assessed in the final analysis.

Pre-processing

Texts of natural language contain a high number of words with different inflections. To obtain a meaningful model of the text dataset, pre-processing ensures that words containing the same information are, in fact, associated with each other.[54] Different options are available such as stop word removal, stemming, or lemmatization. Jacobi et al. argue that lemmatization (determining the canonical form of a word) tends to give better results for richly inflected languages such as German.[79] Following this argument, we applied an advanced probabilistic procedure called TreeTagger based on Markov chains, where transition probabilities are estimated based on decision trees that take into account the context of each word.[80,81] In addition to the canonical form of each word, the software also conveniently provides a part-of-speech (POS) tag indicating the function of each word within the sentence and is highly accurate for German texts.[81] The POS tags also have another advantage: to reduce the complexity of the collection of texts, many text mining approaches use lists of stop words that are to be removed. The POS tags serve a similar purpose in that we can precisely define which word classes we include in the analysis. To capture all meaningful information, we explicitly included all adjectives, adverbs, verbs, and nouns but excluded all particles. In addition, we excluded words that appeared in fewer than 10 articles, as the weight they contributed to the topics' distribution was negligible, to make the model estimation faster without losing statistical information.[54]

We used a change-point analysis to detect changes in mean and variance of the total newspaper coverage of the EEG and to separate different phases of attention.[56,57] The results provided a first overview and pointed out that there have been massive changes in coverage over the years (Figure 1B). The span lies between one article per working day up to an average of four articles per working day. A crucial step in change-point analysis is the choice of the penalty of the cost function. As is common practice, we manually chose the penalty by visual inspection of the results[57, p. 9]. The change-point analysis informs the choice of spline knots that we conducted in defining the STM covariate model structure.

Structural Topic Modeling

To assess the contents and temporal structure of the text collection, we built upon recent developments in automated content analysis or, to be more specific, topic modeling, herein referred to as an STM.[50] The STM extends the popular latent Dirichlet allocation and its successors[46,47] by including observed covariates as linear functions to the mean parameters of the assumed prior topic distributions.[50] Information on the choice of priors is available from Roberts et al.[50] The ability to include covariate information is central to the results of this study, as it provides a framework for the time-series analysis of newspaper coverage.

Structural topic modeling and latent Dirichlet allocation have been applied successfully in different fields such as political science,[45,48] innovation management,[82] or climate change perceptions.[51,83] Both models assume that a document comprises a mixture of k topics, where topic proportions in the STM can be correlated across documents, and prevalence is influenced by covariates such as time.[48] Formally, a topic consists of the distribution over all words in the vocabulary of the text collection. The distribution of topics and distribution of words within topics are estimated based on variational inference.[46,50] Both models are unsupervised; thus, given a predefined number of topics, the topics are inferred during the modeling procedure and not predefined by the analyst. This makes the models particularly suitable for exploratory research with limited *a priori* assumptions,[45] but comes with the necessity of careful validation and interpretation of the results and the difficulty of choosing a number k of topics. Comparative experiments with human coders show that topic modeling yields competitive results.[48]

Before the model can be evaluated, the functional form of covariate dependencies must be defined. As covariate, we included the quarter within which the article was published. For dynamic dependencies on time, the developers of STM propose the usage of splines in order to detect non-linear changes in the topic prevalence over time.[49,50] First, we used natural cubic splines, which means that the second derivative at the boundary knots is required to vanish; thus, the spline extends linearly outside of its domain. This makes sure that single points close to the boundaries do not have too much influence leading to erratic boundary behavior[84, p. 24]. In addition, during regression, fewer parameters have to be estimated. Second, we chose knots corresponding to the points in time when reforms of the EEG became effective that coincided with the change-point analysis conducted above. In addition, we set a knot in each phase that separated the phase into two phases of equal length. This allows for sufficient flexibility for detecting changes in prevalence. In total, nine knots and eight parameters were included in the model. The procedure was inspired by interrupted time-series analysis (e.g., McDowall et al.)[55] and is in line with general advice on the choice of knots when prior knowledge is available.[84, p.26]

After pre-processing and model configuration, analysts must make important decisions on the final model. To reduce the complexity of model choice, we decided on a fixed initialization strategy. Roberts et al.[85] showed that initializing the STM with a solution of a simplified problem using a spectral decomposition of the word co-occurrence matrix delivers favorable results. This reduced the model choice to a choice on the number of topics. Different procedures have been proposed and applied in the literature comprising the qualitative assessment of many different models with a different number of topics (e.g., Farrell)[82] or based on different statistical indicators that measure how well the topics can be interpreted by humans, such as *semantic coherence* (e.g., Mimno et al.).[85] In this study, we followed a hybrid approach.

With the above specifications of the model structure and functional form, we decided on the number of topics as follows: first, we assessed models for $k = 10, 20, \ldots, 100$ based on lists of most probable words and FREX (frequency and exclusivity) in order to get an overview of the topics that could potentially be related to the policy goals of the EEG. FREX is an indicator leveraging the exclusivity of words to a specific topic with the probability of occurrence and has been proven to yield favorable properties for providing word lists for topic interpretation.[86] The first analysis found that the number of topics should be larger than 20. Second, we estimated models with topics between 20 and 100. We evaluated the mean exclusivity based on the FREX indicator and mean semantic coherence against each other as proposed by the package authors of STM (Figure S1).[49] The analysis left us with three promising models with differing numbers of topics, k = 25, 31, and 49, that locally dominated the solutions. A closer investigation revealed that the models with k = 25 and 31 could not distinguish some issues sharply. The choice of the number of topics is also a qualitative decision on a certain perspective on the problem and thus depends on the research questions. Due to the delicacy of the topics contributing to policy goals of the EEG, we chose the model with k = 49, which allowed for a fine-grained analysis of the articles.

Model Validation and Interpretation

Two essential dimensions of the validity of topic models are semantic and predictive validity.[68] Semantic validity refers to whether topics have a coherent meaning and is considered the most important dimension of the validity of content analysis procedures. Standard procedures to assess the semantic validity of topic modeling are the assessment of word lists based on probability or FREX[87] or close reading of a subsample of texts.[45,68] For our case, predictive validity is also highly relevant, as it reflects the extent to which topic prevalence changes relate to external events. In the following, we describe the procedures we applied to assess validity.

After choosing the model and the number of topics, we validated the semantic content of the 49 topics. On one hand, we must ensure the internal semantic validity of the topics. On the other hand, it is necessary for interpretation to find meaningful labels for each category. In accordance with the literature,[45] the first two authors each read at least 10–15 of the articles most associated with each topic and reviewed topic distribution and word lists of the most probable terms and highest FREX. Each researcher documented the results as notes for each article and concluded an overall label of the topic and a short description independently. A comparison of the results showed an agreement of 80% between the two researchers. More precisely, 39 topics were interpreted identically; 6 topics were interpreted in parts differently and 4 topics substantially differently by the two researchers. The 10 topics on which there was disagreement were reconsidered jointly regarding the notes taken by the researchers to reach consensus on the topic label. The results are documented in Table 1 and, together with a short summary, in Table S1. Overall, we reached a consensus on all topics. Only a few topics contained documents that were not expected from the general tendency of the articles. For example, the sample of articles on topic 22 contained two texts that portrayed eccentric leaders of companies other than SolarWorld. The association of the articles can be explained by the many references to SolarWorld's eccentric leader.

The predictive validity of the model can be tested by comparing topic time series against real-world events.[83,88] The analyses presented in the paper and its comparison with political decision processes and outcomes show ample evidence of the external validity of the model results.[12,14,23] In the following, by way of example, we discuss the topic "liberalization of the electricity market" (Figure 4A) along with external events that are expected to affect its prevalence. In such a way, all topics can be discussed as long as they have some resemblance to external events. The German electricity market was liberalized with the enforcement of the Energy Industry Act in 1998. The graph shows that this event affects the coverage of the EEG, particularly at the beginning of our time horizon. However, while this event becomes more distant in the past, its effects on coverage also cease. The bump in the prevalence curve in 2005 can be explained by external events: In 2005, the Energy Industry Act was amended, leading to a temporal upsurge in market liberalization issues.

Research has shown that, within German energy policy, four policy goals have played a major role: environmental sustainability, limiting energy costs, energy security, and energy industry performance.[21] To understand how those policy goals are represented in media coverage, we coded topics according to the policy goal they contributed to most. This coding was done based on the more detailed descriptions provided in Table S1, rather than only on their short labels provided in Table 1. The coding procedure revealed that a set of topics was associated rather with concrete political debates between political stakeholders than directly with the policy goals. Arguably, policy goals are the main contents of those debates, but the model was able to distinguish them from the debates. We therefore further differentiated a category, "politics," from the four policy goals that represents, therefore, an indicator of the prevalence of political activity in newspaper coverage. A description of all five categories can be found in Table S2.

Sentiment Analysis

The approach we used for sentiment analysis can best be classified as an aspect-based one: based on the topic model we have developed above, we assessed the sentiment of each topic.[58] However, we were not interested in a measure that assigns a sentiment score to each document. The topic model we have developed gives rise to a simple way of defining topic sentiment based on the word distributions of each topic and a sentiment lexicon. For sentiment analysis, the lexicon is an important foundation.[58] For the German language, SentiWS provides an established lexicon for German, with more than 3,000 words and more than 16,000 inflected forms.[89] Based on this lexicon, we defined the topic sentiment ts_t as the weighted sum of all sentiment scores $s_w \in [-1, 1]$ of words w from the vocabulary V. We weighed this with the word occurrence probability $\beta_{w,t}$ per topic t estimated by STM:

$$ts_t = \sum_{w \in V} \beta_{w,t} \cdot s_w \, .$$

The resulting topic sentiment gives a well-interpretable indicator for the overall sentiment of a topic: the higher the ts_t, the more likely the usage of words with a positive connotation in the context of this topic. The results are reported in Figure 2.

It is noteworthy that the vocabulary of our corpus consists of 12,217 words, of which 1,662 (or 13.6%) are also in the sentiment lexicon. This is to be expected, as most of the words of natural language do not carry sentiment per se. The ratio implies that the actual range of topic sentiment will be much smaller than $[-1, 1]$. It is bound by the sum of the word occurrence probabilities per topic of the words from the lexicon. Furthermore, positive and negative sentiment annul each other in the definition above. Those facts contribute to the expectation that topic sentiment values will be small and centered around zero. However, we argue that this does not affect robustness for comparative usage and affects only the scale of values. This scaling effect could be offset by normalization, but we think that the definition in terms of the statistical expectation is more intuitive.

Presentation of Results

For interpreting the resulting time series of topic prevalence, we estimated linear models with the spline structure described above. Confidence intervals were obtained by drawing from the posterior distribution of topics as implemented in the *estimateEffects()* function of the STM package.[49] We aggregated the topics to four policy goals and political activity. Aggregation was conducted manually, taking into account the qualitative topic descriptions provided in Table S1. Linear models are provided in Tables S3 and S4.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.patter.2020.100169.

AUTHOR CONTRIBUTIONS

J.D.H. designed the study, analyzed and interpreted data, and wrote the draft. K.S. contributed to the analysis and interpretation of data and writing of the draft. W.F. contributed to the interpretation of results and revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 15, 2020
Revised: October 15, 2020
Accepted: November 17, 2020
Published: December 22, 2020

REFERENCES

1. UNFCCC (2015). Paris agreement. https://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf.
2. Markard, J. (2018). The next phase of the energy transition and its implications for research and policy. Nat. Energy *3*, 628–633.
3. Geels, F.W., Sovacool, B.K., Schwanen, T., and Sorrell, S. (2017). The socio-technical dynamics of low-carbon transitions. Joule *1*, 463–479.
4. Geels, F.W., Sovacool, B.K., Schwanen, T., and Sorrell, S. (2017). Sociotechnical transitions for deep decarbonization. Science *357*, 1242–1244.
5. Markard, J., Raven, R., and Truffer, B. (2012). Sustainability transitions: an emerging field of research and its prospects. Res. Pol. *41*, 955–967.
6. Williams, J.H., DeBenedictis, A., Ghanadan, R., Mahone, A., Moore, J., Morrow, W.R., Price, S., and Torn, M.S. (2012). The technology path to deep greenhouse gas emissions cuts by 2050: the pivotal role of electricity. Science *335*, 53–59.

7. IRENA, IEA, and REN21 (2018). Renewable Energy Policies in a Time of Transition (International Renewable Energy Agency).
8. Karneyeva, Y., and Wüstenhagen, R. (2017). Solar feed-in tariffs in a post-grid parity world: the role of risk, investor diversity and business models. Energy Policy *106*, 445–456.
9. Meadowcroft, J. (2009). What about the politics? Sustainable development, transition management, and long term energy transitions. Policy Sci. *42*, 323–340.
10. Antal, M., and Karhunmaa, K. (2018). The German energy transition in the British, Finnish and Hungarian news media. Nat. Energy *3*, 994–1001.
11. Strunz, S. (2014). The German energy transition as a regime shift. Ecol. Econ. *100*, 150–158.
12. Lauber, V., and Jacobsson, S. (2016). The politics and economics of constructing, contesting and restricting socio-political space for renewables – the German Renewable Energy Act. Environ. Innovation Societal Transitions *18*, 147–163.
13. Renn, O., and Marshall, J.P. (2016). Coal, nuclear and renewable energy policies in Germany: from the 1950s to the "Energiewende". Energy Policy *99*, 224–232.
14. Strunz, S., Gawel, E., and Lehmann, P. (2016). The political economy of renewable energy policies in Germany and the EU. Util. Pol. *42*, 33–41.
15. Kriechbaum, M., López Prol, J., and Posch, A. (2017). Looking back at the future. Dynamics of collective expectations about photovoltaic technology in Germany & Spain. Technol. Forecast. Soc. Change *129*, 76–87.
16. Leipprand, A., Flachsland, C., and Pahle, M. (2017). Energy transition on the rise: discourses on energy future in the German parliament. Innovation: Eur. J. Soc. Sci. Res. *30*, 283–305.
17. Schmid, E., Pechan, A., Mehnert, M., and Eisenack, K. (2017). Imagine all these futures: on heterogeneous preferences and mental models in the German energy transition. Energy Res. Soc. Sci. *27*, 45–56.
18. Krick, E. (2018). Ensuring social acceptance of the energy transition. The German government's 'consensus management' strategy. J. Environ. Plann. Pol. Manage. *20*, 64–80.
19. Leiren, M.D., and Reimer, I. (2018). Historical institutionalist perspective on the shift from feed-in tariffs towards auctioning in German renewable energy policy. Energy Res. Soc. Sci. *43*, 33–40.
20. Buschmann, P., and Oels, A. (2019). The overlooked role of discourse in breaking carbon lock-in: the case of the German energy transition. Wiley Interdiscip. Rev. Clim. Change *24*, e574.
21. Schmidt, T.S., Schmid, N., and Sewerin, S. (2019). Policy goals, partisanship and paradigmatic change in energy policy – analyzing parliamentary discourse in Germany over 30 years. Clim. Pol. *355*, 1–16.
22. Pahle, M., Burtraw, D., Flachsland, C., Kelsey, N., Biber, E., Meckling, J., Edenhofer, O., and Zysman, J. (2018). Sequencing to ratchet up climate policy stringency. Nat. Clim. Change *8*, 861–867.
23. Hoppmann, J., Huenteler, J., and Girod, B. (2014). Compulsive policy-making—the evolution of the German feed-in tariff system for solar photovoltaic power. Res. Pol. *43*, 1422–1441.
24. Joas, F., Pahle, M., Flachsland, C., and Joas, A. (2016). Which goals are driving the Energiewende? Making sense of the German energy transformation. Energy Policy *95*, 42–51.
25. Groh, E.D., and Möllendorff, C.v. (2020). What shapes the support of renewable energy expansion? Public attitudes between policy goals and risk, time, and social preferences. Energy Policy *137*, 111171.
26. Bennett, W.L. (1990). Toward a theory of press-state relations in the United States. J. Commun. *40*, 103–127.
27. Althaus, S.L. (2003). When news norms collide, follow the Lead: new evidence for press independence. Polit. Commun. *20*, 381–414.
28. Shehata, A., and Hopmann, D.N. (2012). Framing climate change. Journal. Stud. *13*, 175–192.
29. Walgrave, S., and van Aelst, P. (2006). The contingency of the mass media's political agenda setting power: toward a preliminary theory. J. Commun. *56*, 88–109.
30. McCombs, M.E., and Shaw, D.L. (1972). The agenda-setting function of mass media. Public Opin. Q. *36*, 176.
31. McCombs, M. (2005). A Look at Agenda-setting: past, present and future. Journalism Stud. *6*, 543–557.
32. Entman, R.M. (1993). Framing: toward clarification of a fractured paradigm. J. Commun. *43*, 51–58.
33. Entman, R.M. (2007). Framing bias. Media in the distribution of power. J. Commun. *57*, 163–173.
34. Page, B.I., and Shapiro, R.Y. (1983). Effects of public opinion on policy. Am. Polit. Sci. Rev. *77*, 175–190.
35. Burstein, P. (2003). The impact of public opinion on public policy: a review and an agenda. Polit. Res. Q. *56*, 29–40.
36. Schmid, N., Sewerin, S., and Schmidt, T.S. (2019). Explaining advocacy coalition change with policy feedback. Pol. Stud. J. *109*, 183.
37. Markard, J., Wirth, S., and Truffer, B. (2016). Institutional dynamics and technology legitimacy – a framework and a case study on biogas technology. Res. Pol. *45*, 330–344.
38. Rochyadi-Reetz, M., Arlt, D., Wolling, J., and Bräuer, M. (2019). Explaining the media's framing of renewable energies: an international comparison. Front. Environ. Sci. *7*, 140.
39. van Lente, H., Spitters, C., and Peine, A. (2013). Comparing technological hype cycles. Towards a theory. Technol. Forecast. Soc. Change *80*, 1615–1628.
40. Ruef, A., and Markard, J. (2010). What happens after a hype? How changing expectations affected innovation activities in the case of stationary fuel cells. Technol. Anal. Strateg. Manage. *22*, 317–338.
41. Downs, A. (1972). Up and down with ecology - the issue-attention-cycle. Publ.Interest *28*, 38–50.
42. Gupta, K., and Jenkins-Smith, H. (2015). Anthony downs, "up and down with ecology: the 'issue-attention' cycle". In The Oxford Handbook of Classics in Public Policy and Administration, M. Lodge, E.C. Page, and S.J. Balla, eds. (Oxford University Press).
43. Woolley, J.T. (2000). Using media-based data in studies of politics. Am. J. Polit. Sci. *44*, 156.
44. Atkinson, M.L., Lovett, J., and Baumgartner, F.R. (2014). Measuring the media agenda. Polit. Commun. *31*, 355–380.
45. Quinn, K.M., Monroe, B.L., Colaresi, M., Crespin, M.H., and Radev, D.R. (2010). How to analyze political attention with minimal assumptions and costs. Am. J. Polit. Sci. *54*, 209–228.
46. Blei, D., Ng, A., and Jordan, M. (2003). Latent dirichlet allocation. J. Machine Learn. Res. *2003*, 993–1022.
47. Blei, D.M., and Lafferty, J.D. (2007). A correlated topic model of Science. Ann. Appl. Stat. *1*, 17–35.
48. Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., and Rand, D.G. (2014). Structural topic models for open-ended survey responses. Am. J. Polit. Sci. *58*, 1064–1082.
49. Roberts, M., Stewart, B., and Tingley, D. (2019). Stm: anR package for structural topic models. J. Stat. Soft. *91*.
50. Roberts, M.E., Stewart, B.M., and Airoldi, E.M. (2016). A model of text for experimentation in the social sciences. J. Am. Stat. Assoc. *111*, 988–1003.
51. Tvinnereim, E., and Fløttum, K. (2015). Explaining topic prevalence in answers to open-ended survey questions about climate change. Nat. Clim. Change *5*, 744–747.
52. KIM, S.E. (2018). Media bias against foreign firms as a veiled trade barrier: evidence from Chinese newspapers. Am. Polit. Sci. Rev. *112*, 954–970.
53. Mildenberger, M., and Tingley, D. (2019). Beliefs about climate beliefs: the importance of second-order opinions for climate politics. Br. J. Polit. Sci. *49*, 1279–1307.
54. Lucas, C., Nielsen, R.A., Roberts, M.E., Stewart, B.M., Storer, A., and Tingley, D. (2015). Computer-assisted text analysis for comparative politics. Polit. Anal. *23*, 254–277.

55. McDowall, D., McCleary, R., Meidinger, E., Hay, R.A., Jr., and others. (1980). Interrupted Time Series Analysis (SAGE).
56. Killick, R., Fearnhead, P., and Eckley, I.A. (2012). Optimal detection of changepoints with a linear computational cost. J. Am. Stat. Assoc. *107*, 1590–1598.
57. Killick, R., and Eckley, I.A. (2014). changepoint: an R Package for Changepoint Analysis. J. Stat. Soft. *58*.
58. Feldman, R. (2013). Techniques and applications for sentiment analysis. Commun. ACM *56*, 82–89.
59. Pang, B., and Lee, L. (2008). Opinion mining and sentiment analysis. FNT Inf. Retriev. *2*, 1–135.
60. Soroka, S., Fournier, P., and Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. Proc. Natl. Acad. Sci. U S A *116*, 18888–18892.
61. Quitzow, R. (2015). Dynamics of a policy-driven market. The co-evolution of technological innovation systems for solar photovoltaics in China and Germany. Environ. Innovat. Societal. Transitions *17*, 126–148.
62. Dehler, J., Keles, D., Telsnig, T., Fleischer, B., Baumann, M., Fraboulet, D., Faure-Schuyer, A., and Fichtner, W. (2017). Self-consumption of electricity from renewable sources. In Europe's Energy Transition - Insights for Policy Making (Elsevier), pp. 225–236.
63. Schürmann, H.J. (2000). VDEW-chef Marquis: Staat Schöpft Preisvorteile Wieder Ab (Handelsblatt).
64. Huenteler, J., Schmidt, T.S., and Kanie, N. (2012). Japan's post-Fukushima challenge – implications from the German experience on renewable energy policy. Energy Policy *45*, 6–11.
65. Bundesregierung (2011). Energiewende mit Augenmass, https://archiv.bundesregierung.de/archiv-de/energiewende-mit-augenmass-425304.
66. Bundesregierung (2011). Regierungserklärung von Bundeskanzlerin Angela Merkel zur Energiepolitik "Der Weg zur Energie der Zukunft" (Mitschrift). https://www.bundesregierung.de/breg-de/themen/energiewende/regierungserklaerung-von-bundeskanzlerin-angela-merkel-zur-energiepolitik-der-weg-zur-energie-der-zukunft-mitschrift–1008262.
67. Grimmer, J., and Stewart, B.M. (2013). Text as data: the promise and pitfalls of automatic content analysis methods for political texts. Polit. Anal. *21*, 267–297.
68. Schmidt, T.S., and Sewerin, S. (2017). Technology as a driver of climate and energy politics. Nat. Energy *2*, 649.
69. Schumann, D., Fischer, W., and Hake, J.-F. (2017). Bewertung der Energiewende und Energiepolitik in der Bevölkerung. Energiewirtschaftliche Tagesfragen *67*, 77–81.
70. Wolf, I. (2020). Soziales Nachhaltigkeitsbarometer der Energiewende 2019. Kernaussagen und Zusammenfassung der wesentlichen Ergebnisse (Institut für transformative Nachhaltigkeitsforschung (IASS)).
71. Pegels, A., and Lütkenhorst, W. (2014). Is Germany's energy transition a case of successful green industrial policy? Contrasting wind and solar PV. Energy Policy *74*, 522–534.
72. Huenteler, J., Schmidt, T.S., Ossenbrink, J., and Hoffmann, V.H. (2016). Technology life-cycles in the energy sector — technological characteristics and the role of deployment for innovation. Technol. Forecast. Soc. Change *104*, 102–121.
73. Shoemaker, P.J., and Vos, T.P. (2009). Gatekeeping Theory (Routledge).
74. Fawzi, N. (2018). Beyond policy agenda-setting: political actors' and journalists' perceptions of news media influence across all stages of the political process. Inf. Commun. Soc. *21*, 1134–1150.
75. Hammar, H., and Jagers, S.C. (2006). Can trust in politicians explain individuals' support for climate policy? The case of CO 2 tax. Clim. Pol. *5*, 613–625.
76. Bundestag, D. (2019). Drucksache 19/13900. Klimaschutzprogramm 2030 der Bundesregierung zur Umsetzung des Klimaschutzplans 2050.
77. Kemfert, C. (2017). Germany must go back to its low-carbon future. Nature *549*, 26–27.
78. Jacobi, C., van Atteveldt, W., and Welbers, K. (2015). Quantitative analysis of large amounts of journalistic texts using topic modelling. Digital Journal. *4*, 89–106.
79. Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. In Proceedings of International Conference on New Methods in Language Processing (UK Manchester: Association for Computational Linguistics).
80. Schmid, H. (1995). Improvements in Part-of-Speech Tagging with an Application to German. In Proceedings of the ACL SIGDAT-Workshop (Dublin, Ireland: Association for Computational Linguistics).
81. Antons, D., Kleer, R., and Salge, T.O. (2016). Mapping the topic landscape of JPIM , 1984-2013. In search of hidden structures and development trajectories. J. Prod. Innov. Manag. *33*, 726–749.
82. Farrell, J. (2016). Corporate funding and ideological polarization about climate change. Proc. Natl. Acad. Sci. U S A *113*, 92–97.
83. Harrell, F.E. (2015). Regression Modeling Strategies (Springer International Publishing).
84. Roberts, M.E., Stewart, B.M., and Tingley, D. (2016). Navigating the local modes of big data: the case of topic models. In Computational Social Science. Discovery and Prediction, R.M. Alvarez, ed. (New York, NY: Cambridge University Press), pp. 51–97.
85. Mimno, D., Wallach, H.M., Talley, E., Leenders, M., and McCallum, A. (2011). Optimizing semantic coherence in topic models. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (Association for Computational Linguistic), pp. 262–272.
86. Airoldi, E.M., and Bischof, J.M. (2016). Improving and evaluating topic models and other models of text. J. Am. Stat. Assoc. *111*, 1381–1403.
87. DiMaggio, P., Nag, M., and Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. government arts funding. Poetics *41*, 570–606.
88. Remus, R., Quasthoff, U., and Heyer, G. (2010). SentiWS - A Publicly Available German-language Resource for Sentiment Analysis. In Proceedings of the International Conference on Language Resources and Evaluation, LREC.
89. AG Energiebilanzen, E.V. (2019). Auswertungstabellen zur Energiebilanz für die Bundesrepublik Deutschland 1990 bis 2018. https://ag-energiebilanzen.de/10-0-Auswertungstabellen.html.