Political news coverage of massmedia

...

Franziska Löw*

November 2020

Introduction

In democracies, the media fulfill fundamental functions: They should inform the people, contribute to the formation of opinion through criticism and discussion and thus enable participation. In recent decades, however, concern has grown about the role of the media in politics in general and in election campaigns in particular. They are criticized for influencing election results through their reporting and for helping populist parties in particular to flourish. After the 2017 federal elections in Germany, for example, the media were accused of contributing to the success of the right-wing populist AfD by increasingly including the party's content and using the same language in their articles as the AfD. Representatives of these media houses strongly opposed this accusation. The purpose of this study is to examine whether there is evidence to support the accusation that media reports are biased, especially during election campaigns.

For advertising-financed media the battle for the attention of the recipients is at the center of economic decisions. Online media in particular, which offer their content to a large extent free of charge and generate their revenue through advertising space, compete for the scarce resource of attention. Consumers a non-monetary price providing their attention, which the media platform bundles and sells on to advertising customers. This business model corresponds to that of a platform market, in which media companies act as platforms that connect the market of advertising with the reader market to exploit the indirect network effects between them (Dewenter and Rösch 2014). A profit-maximizing publisher therefore directs its economic decisions according to what will attract the most attention.

This conclusion, derived from the economic theory of platform markets, corresponds to the notion of media logic, a central concept in the field of media and communication studies (Takens et al. 2013). The debate about media logic is embedded in the broader discussion about the interaction between the press, politics and the public. The underlying thesis is that the content of political news is the product of news values and narrative techniques that media use to attract audiences (Strömbäck 2008). According to Takens et al. (2013), three content attributes highly correspond with news values and influence how journalists interpret political events: 1) personalized content, i.e., the focus on individual politicians; 2) the framing of politics as a contest and 3) negative coverage. Similarly Blassnig et al. (2019) states that media primarily focus on news factors, i.e. the factors that turn an event into news worth reporting like conflict, drama, negativity, surprise or proximity. Likewise populist messages often co-occur with negative, emotionalized, or dramatized communication style, thus utilizing similar mechanisms as the media logic, respectively the attention economy. In fact, Blassnig et al. (2019) shows that populist key messages by political and media actors in news articles provoke more reader comments under these articles. Media competing for the attention of readers therefore have an incentive to pick up on the key messages of these parties.

Especially during election campaigns political parties want the media agenda to be congruent with their own agenda to define the issue-based criteria on which they will be evaluated by voters (Eberl, Boomgaarden,

*Institut für Industrieökonomik, Helmut Schmidt Universität. Email: .

and Wagner 2017). Parties instrumentalize their public relations in order to highlight issues that they are perceived to be competent on, that they "own" and that are important to their voters (Kepplinger and Maurer 2004).

But does increased reporting also lead to rising survey results? Especially if this reporting is largely negative. In political science, several studies have examined at least the first aspect of this question (see for example Druckman and Parkin (2005), Eberl (2018)). In general, it is assumed that smaller, non-established parties in particular benefit from placing their topics in the media in order to get them into the voters' heads. Here, the tendency of the reporting is irrelevant but rather the quantity is decisive.

However, the causal relationship between reporting and voter preferences is not the subject of this study. Rather, it is intended to investigate whether differences exist in media coverage of different parties before and after the 2017 federal elections in Germany. In order to answer these and other media-related questions in the political context, quantifying the content of media is a prerequisite. One of the key challenges is to determine the features that are used to describe media content (audio, video, text). Studies that rely on quantifying media content for their analyses use, for example, visibility (how often political actors appear in the media) or tonality (how they are evaluated). Other studies examine the topics discussed or the language used in the media, in order to identify whether political actors are able to place their own policy positions in the media. Leading studies from economic literature, for example, examine how often a newspaper quotes the same think tanks (Groseclose and Milyo 2005, @lott_is_2014) or uses the same language (Gentzkow and Shapiro 2004) as members of Congress. Following this approach, the present paper compares topics discussed in media outlets with topics addressed in the press releases of the parties in the German "Bundestag", to measure the content similarity between online news and parties press releases.[1] To discover the latent topics in the corpus of text data, the structural topic model (STM) developed by M. E. Roberts, Stewart, and Airoldi (2016) is applied. This probabilistic text model results in a probability distribution for each document across all topics, which is then aggregated to calculate the degree of difference between the news articles of different media providers and the press releases of the parties using a linear regression model.

Literature review

Newspaper articles and their metadata, such as publisher and publication date have been subject of investigation in economic research.

Background information

The political situation in Germany (June 2017 - March 2018)

The articles analyzed in this paper cover a period from June 1, 2017 to March 1, 2018 and thus cover both the most important election campaign topics for the Bundestag elections on September 24, 2017 and the process of forming a government that lasted until February 2018. After four years in a grand coalition with the Social Democrats (SPD), German Chancellor Angela Merkel, member of the conservative party CDU/CSU (also known as Union), ran for re-election. The SPD nominated Martin Schulz as their candidate.

On the right side of the political spectrum, AfD (alternative for Germany) managed to be elected to the German Bundestag for the first time in 2017. The political debate about the high refugee numbers of the past years brought a political upswing to the AfD, which used the dissatisfaction of parts of the population to raise its own profile. In the course of the reporting on the federal elections, leading party members of the AfD as well as party supporters repeatedly accused the mass media of reporting unilaterally and intentionally presenting the AfD badly.

After the election, the formation of a government was difficult due to the large number of parties elected to the Bundestag and the considerable loss of votes by the major parties CDU/CSU and SPD. Since all parties

[1] For the sake of simplicity, both news articles and press releases will be referred to as documents in the following.

rejected a coalition with the AfD, numerically only two coalitions with an absolute parliamentary majority were possible: a grand coalition ("GroKo" - from the German word Große Koalition) of CDU/CSU and SPD, and a Jamaica coalition (coalition of CDU/CSU, FDP (economic liberal party) and B90/Die Grünen (Bündnis 90/Die Grünen, green party)). The grand coalition was initially rejected by the SPD. The four-week exploratory talks on the possible formation of a Jamaica coalition officially failed on November 19, 2017 after the FDP announced its withdrawal from the negotiations. FDP party leader Christian Lindner said that there had been no trust between the parties during the negotiations. The main points of contention were climate and refugee policy. CDU and CSU regretted this result, while B90/Die Grünen sharply criticized the liberals' withdrawal. The then Green leader Cem Özdemir accused the FDP of lacking the will to reach an agreement.

After the failure of the Jamaica coalition talks, a possible re-election or a minority government as alternatives were discussed in the media before the SPD decided to hold coalition talks with the CDU/CSU. This led to great resistance from the party base, which called for a party-internal referendum on a grand coalition. After the party members voted in favor of the grand coalition, a government was formed 171 days after the federal elections.

Figure 1 shows that support for the two major popular parties has been declining in recent months since August 2017, with the CDU/CSU again showing positive survey results since November 2017.[2] However,the poll results of the SPD have been falling since March 2017. At the same time, the AfD in particular has been recording increasingly positive survey results since June 2017.

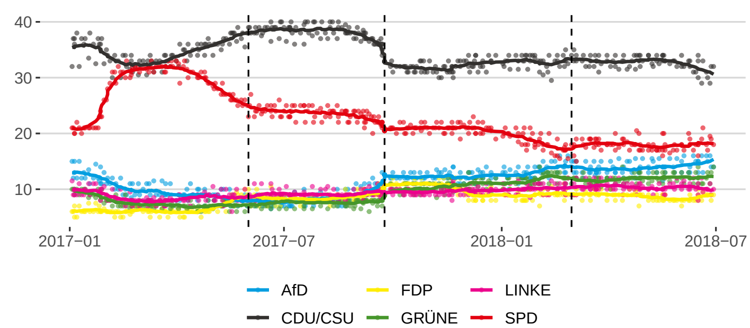

Figure 1: Election polls

German online news market

The analysis performed in this paper is based on the news articles of the following news websites: Bild.de, DIE WELT, FOCUS ONLINE, Handelsblatt.com, SPIEGEL ONLINE, stern.de, ZEIT ONLINE. As can be seen from Figure 2(a), expect for Handelsbaltt.com (53), these media outlets are among the top 30 German online news providers in the period under review in terms of visits.[3]

The main source of income for these privately managed media houses is digital advertising, even though paid content is playing an increasingly important role. However, according to a survey on digital news by the Reuters Institute (Newman et al. 2018) only 8% of respondents pay for online news. The online survey for German data was undertaken between 19th - 22nd January 2018 by the Hans Bredow Institute[4] with a total sample size of 2038 adults (aged 18+) who access news once a month or more. Among other

[2] For each party the survey results of the seven major institutes are considered. To calculate a smooth line for each party on each day, the moving average within 15 days (7 before the day, 7 after the day, and the day itself) is estimated. The data source is https://www.wahlrecht.de/.

[3] The term visit is used to describe the call to a website by a visitor. The visit begins as soon as a user generates a page impression (PI) within an offer and each additional PI, which the user generates within the offer, belongs to this visit.

[4] https://www.hans-bredow-institut.de/de/projekte/reuters-institute-digital-news-survey

questions, participants were asked which news sources they use to access news online.[5] The results displayed in Figure 2(b) indicate that the media used for the analysis play a relevant role in their consumption.

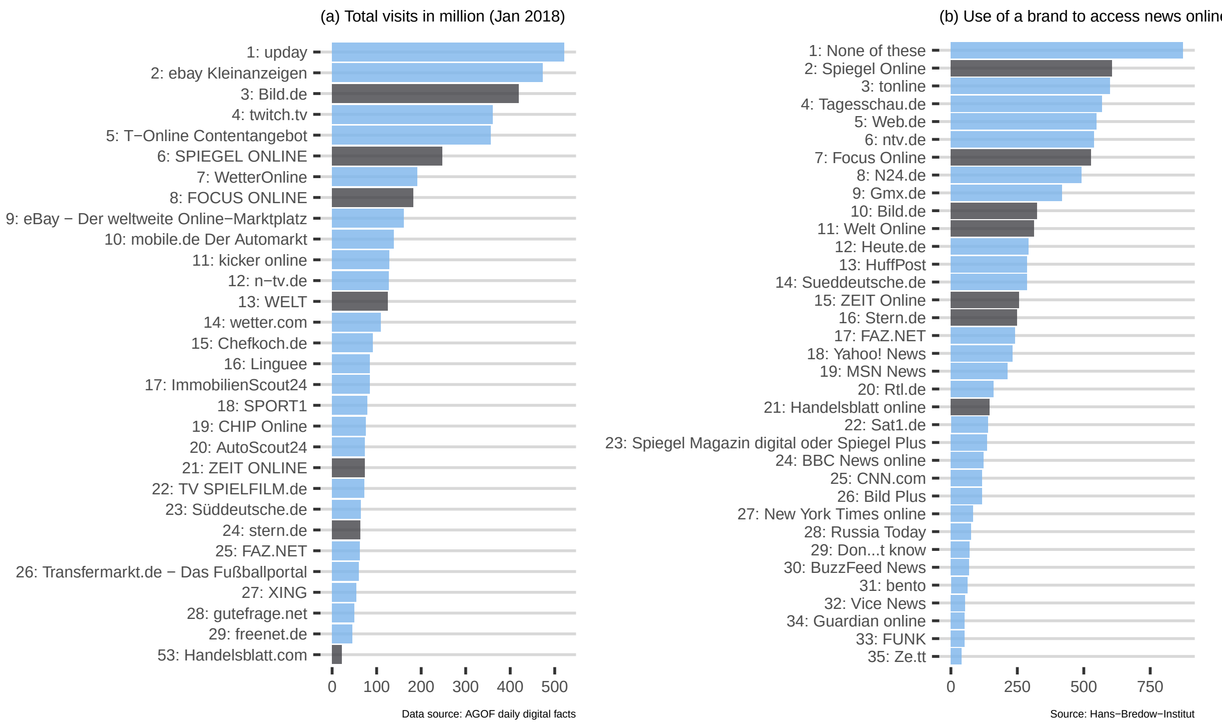

Figure 2: Selected german news brands

Data

I conduct the estimation on a sample of 16,473 online news articles from seven German news providers[6] about domestic politics and press releases of the seven parties that have been in the Bundestag since the 2017 federal elections[7]. Both news articles and press releases are dated from June 1, 2017 to March 1, 2018.

News articles scraped from the Webhose.io API.[8] In order to consider only news about national politics, the articles were filtered based on their URL.

Figure 3 shows the distribution of the number of articles by date and media outlet. There is a high peak around the federal elections on September, 24th and another one shortly after the failure of the Jamaica coalition talks on November, 19th (indicated by the red dotted lines).[9] Furthermore, Figure 3 shows that *DIE WELT* published the most articles on domestic policy, followed by *stern.de*, *Handelsblatt* and *FOCUS ONLINE*.

The press releases were scraped from the public websites of the political parties and parliamentary groups using an automated script written in *Python*.[10]

[5] The exact question was: "Which of the following brands have you used to access news online in the last week (via websites, apps, social media, and other forms of Internet access)? Please select all that apply."

[6] Bild.de, DIE WELT, FOCUS ONLINE, SPIEGEL ONLINE, stern.de, ZEIT ONLINE, Handelsblatt

[7] CDU, SPD, B90/Grüne, FDP, AfD, Die Linke

[8] For more information see https://docs.webhose.io/reference#about-webhose. The scraping code was written in Python and can be made available on request.

[9] The peak in July especially for *stern.de* is due to increased reporting about the G20 summit in Hamburg.

[10] The scraping code was written in Python and can be made available on request.

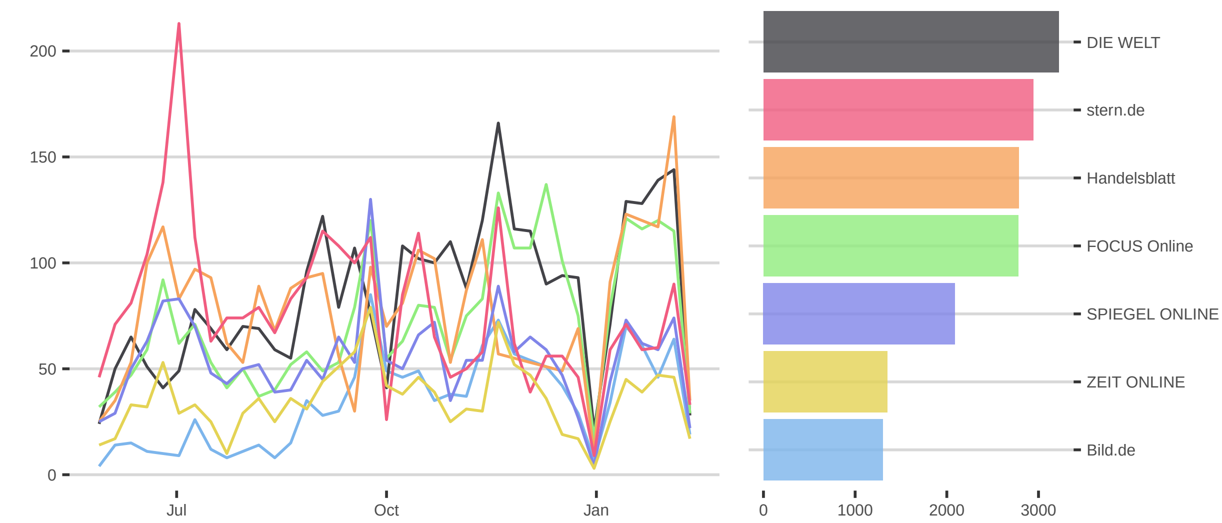

Figure 3: Distribution of news articles

Table 1: Online sources for press releases

	Party	Parliamentary Group
CDU	cdu.de	presseportal.de
SPD	spd.de	spdfraktion.de
FDP	fdp.de	fdpbt.de
B90/Die Grünen	gruene.de	gruene-bundestag.de
DIE LINKE	die-linke.de	die-linke.de/start/presse/aus-dem-bundestag
AfD	afd.de	afdbundestag.de

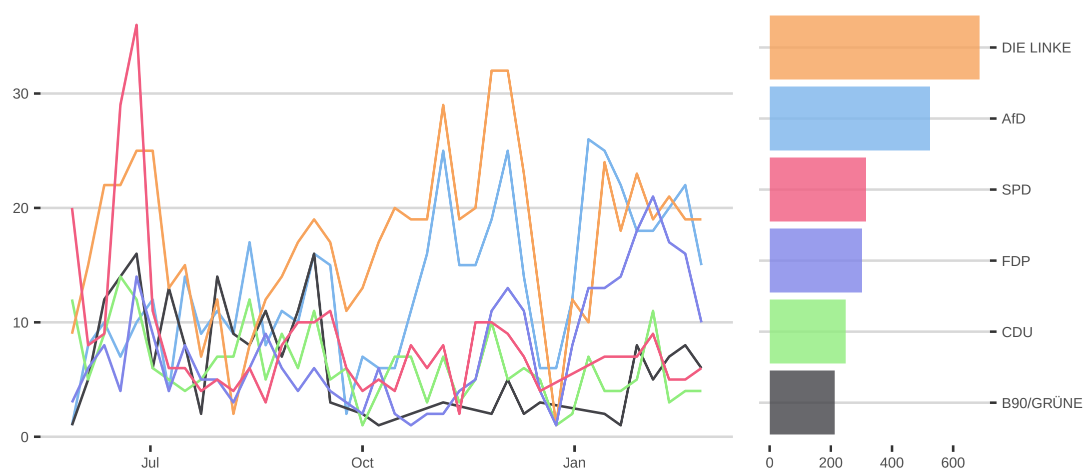

Figure 4: Distribution of press releases

Looking at the boxplots of text length (Figure 5), it becomes evident that:[11]

- mean and median text length of news articles is higher than in press releases
- Handelsblatt published new articles with the highest median text length (488) followed by ZEIT ONLINE (459), however DIE WELT has the article with the highest word count (14.507).
- press releases of CDU have the highest median (256), but also the highest standard dev. (106). press releases of FDP have the lowest median (144).

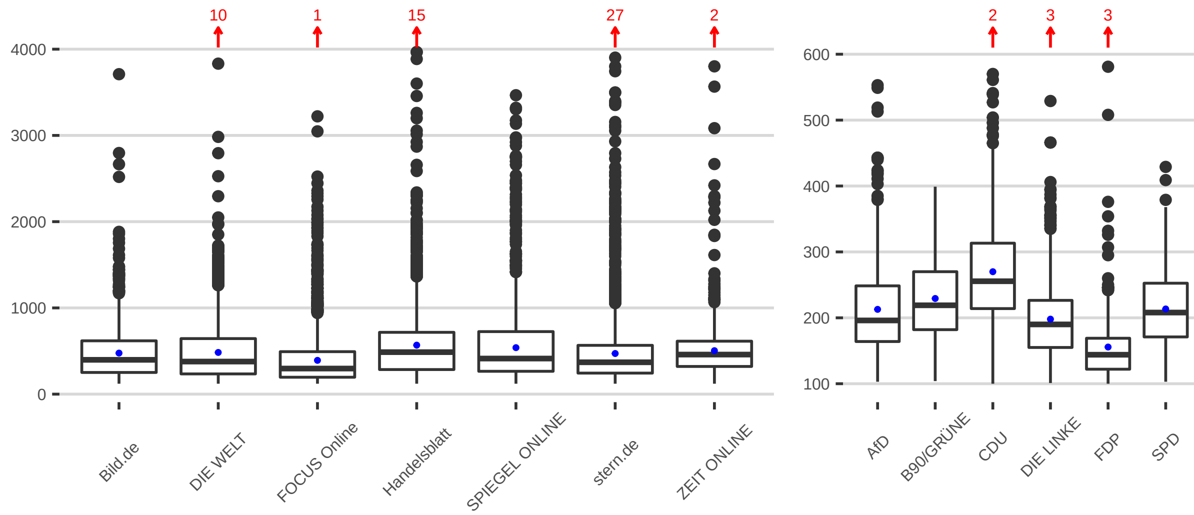

Figure 5: Text length

To remove distorting words, the pre-defined stop word list from the Snowball project[12] is used together with a customized, domain-specific list of stop-words. Additionally punctuation character (e.g. ., „ !, ?, etc.) and all numbers are removed from the data. A next step to reduce the dimensionality of text data is to apply an adequate stemming technique. Stemming is a process by which different morphological variants of a word are traced back to their common root. For example, "voting" and "vote" would be treated as two instances of the same token after the stemming process. There are many different techniques for the stemming process. I apply the widely used Porter-Stemmer algorithm, which is based on a set of shortening rules that are applied to a word until it has a minimum number of syllables.[13]

As an example, the following word clouds represent the most frequent words of the pre-processed articles for Bild.de and press releases of AfD. It becomes evident that these are texts discussing domestic policy issues. The SPD in particular seems to be highly frequent for *Bild.de*.

The next step is to divide the entire dataset into individual documents and to represent these documents as a finite list of unique terms. In this setting, each news article and each press release represents a document d, whereby each of these documents can be assigned to a news website or a party. The sum of all documents forms what is called the corpus. For each document $d \in \{1, ..., D\}$ the number of occurrences of term v in document d is computed, in order to obtain the count $x_{d,v}$, where each unique term in the corpus is indexed by some $v \in \{1, ..., V\}$ and where V is the number of unique terms. The D x V matrix $\boldsymbol{X}$ of all such counts is called the document-term matrix. Each row in this matrix represents a document, where each entry in this row counts the occurrences of a unique term in that document. This representation is often referred to as the bag of words model (Gentzkow, Kelly, and Taddy 2017), since the order in which words are used within a document is disregarded.

Estimate topic similarity of documents

A structural topic model to identify the latent topics

To discover the latent topics in the corpus of press releases and news articles, a structural topic modeling (STM) developed by (M. E. Roberts, Stewart, and Airoldi 2016) is applied. In general, topic models formalize

[12] http://snowball.tartarus.org/algorithms/german/stop.txt
[13] https://tartarus.org/martin/PorterStemmer/

Figure 6: Wordcloud before pre-processing

the idea that documents are formed by hidden variables (topics) that generate correlations among observed terms. They belong to the group of unsupervised generative models, meaning that the true attributes (topics) cannot be observed. The STM developed by (M. E. Roberts, Stewart, and Airoldi 2016) is a recent extension of the standard topic modelling technique, labeled as latent Dirichlet allocation (LDA), which refers to the Bayesian model in (Blei, Ng, and Jordan 2003) that treats each word in a topic and each topic in a document as generated from a Dirichlet - distributed prior.[14].

One crucial assumption to be made for such models is the number of topics (K) that occur over the entire corpus. The underlying idea for these models suggests that each individual topic k potentially contains all of the unique terms within the vocabulary V with different probability. Therefore, each topic k can be represented as a probability vector ϕ_k over all unique terms V. Simultaneously, each individual document d in the corpus can be represented as a probability distribution θ_d over the K topics.

The difference between the widely used LDA and the STM approaches lies in how the posterior distributions (θ and ϕ) are determined. LDA assumes that θ Dirichlet(α) and ϕ Dirichlet(β), where α and β are fitted with the model. While for STM, the prior distributions for θ and ϕ depend on document-level covariates (e.g. the author or date of a document). For this purpose, the the STM specifies two design matrices of covariates (X and Z), where each line defines a vector of covariates for a specific document. In X, the covariates for topic prevalence are given, so that the probability of a topic for each document varies according to X, rather than resulting from a single common prior. The same applies to Z, in which the covariates for the word distribution within a topic are specified. The underlying data generating process to generate each individual word $w_{d,n}$ in a document d for the n^{th} word-position can be described as follows:

- for each document i, draw its distribution of topics θ_d depending on the metadata included in the model defined in X;
- for each topic k, draw its distribution of words ϕ_k depending on the metadata included in the model defined in Z;
- for each word n, draw its topic z_n based on θ_i;
- for each word word n, draw the term distribution for the selected topic $\phi_{z_{d,n}}$.

STM assumes a fixed user-specified number of topics. There is not a "right" answer to the number of topics that are appropriate for a given corpus (Grimmer and Stewart 2013). (M. Roberts, Stewart, and Tingley 2016) propose to measure topic quality through a combination of semantic coherence and exclusivity of words to topics. Semantic coherence is a criterion developed by (Mimno et al. 2011) and is closely related to pointwise mutual information (Newman et al. 2010): it is maximized when the most probable words in a given topic frequently co-occur together.

[14]See also Griffiths and Steyvers (2002), Griffiths and Steyvers (2004) and Hofmann (1999)

Using the function *searchK* from the *stm* package several automated tests are performed to help choose the number of topics including the average exclusivity and semantic coherence as well as the held out likelihood (Wallach, Mimno, and McCallum 2009) and the residuals (Taddy 2012).

Document Cosine similarity

Cosine similarity is built on the geometric definition of the dot product of two vectors. It is a measure for the distance between two vectors and is defined between zero and one; values towards 1 indicate similarity.

$$\text{cosine similarity} = \cos(\theta) = \frac{a * b}{||a||||b||}$$

As topic proportions per document are vectors of the same length, the cosine similarity allows a comparison of the topic distribution between two documents.[15]

Linear regression model

OLS dummy regression

$$\text{CosineSimilarity}_i = \beta_0 + \beta_n Dn_j + \epsilon_i$$

where the $D2_j, D3_j, ..., Dn_j$ represent dummy variables for a political party j

Regression discontinuity model

The idea of regression regression discontinuity design is to use observations with a W_i close to c for the estimation of β_1. β_1 is the average treatment effect for observations with $W_i = c$ which is assumed to be a good approximation to the overall treatment effect. In other words, β_1 gives us the average change of news coverage of media outlet i after the election day. Since interaction terms $T_i Dn_j$ are included, we can estimate the treatment effect for each party.

Calculate a regression discontinuity model for each newspaper i.

$$\text{CosineSimilarity}_i = \beta_0 + \beta_1 T_i + \beta_2 W_{centered} + \beta_n Dn_j + \beta_{n+n} T_i Dn_j + \epsilon_i$$

where $D2_j, D3_j, ..., Dn_j$ represent dummy variables for a political party j and $W_{centered} = $ date - election day.

$$T_i = 1 \text{ if date } >= \text{election date} T_i = 0 \text{ if date } < \text{election date}$$

so that the receipt of treatment T_i is determined by the threshold c (election day) of the continuous variable W_i (date), the so called running variable.

Without interaction terms

$$\text{CosineSimilarity}_i = \beta_0 + \beta_1 T_i + \beta_n Dn_j + \epsilon_i$$

[15]For applications of cosine similarity to compare of topic model outcomes see e.g. Rehs (2020) and Ramage, Dumais, and Liebling (2010)

With interaction terms

$$\text{CosineSimilarity}_i = \beta_0 + \beta_1 T_i + \beta_2 Dn_j + \beta_3 T_i Dn_j + \epsilon_i$$

The interaction term $T_i * Dn$ means that the slope can vary on either side of the treatment threshold for each party.

- The coefficient β_1 is how the intercept jumps (the RDD effect)
- β_3 is how the slope changes for each party

Annex

Table 2: Summary statistics of text length

source	n	mean	sd	median	min	max
AfD	523	212.83	72.16	196	103	553
B90/GRÜNE	211	229.32	63.37	219	104	399
Bild.de	1303	476.07	318.28	398	121	3710
CDU	248	274.54	106.08	256	100	1030
DIE LINKE	686	200.47	71.78	190	101	1048
DIE WELT	3222	509.57	612.06	380	121	14507
FDP	301	161.9	83.78	144	100	999
FOCUS Online	2780	393.89	317.05	297.5	121	5647
Handelsblatt	2785	589.51	495.82	488	121	6899
SPD	315	213.41	56.16	208	103	429
SPIEGEL ONLINE	2089	539.09	415.05	413	121	3466
stern.de	2943	514.66	616.55	373	121	9287
ZEIT ONLINE	1351	513.75	387.14	459	121	8015

References

Bholat, David M., Stephen Hansen, Pedro M. Santos, and Cheryl Schonhardt-Bailey. 2015. "Text Mining for Central Banks." *SSRN Electronic Journal*, June. http://www.academia.edu/13430482/Text_mining_for_central_banks.

Blassnig, Sina, Sven Engesser, Nicole Ernst, and Frank Esser. 2019. "Hitting a Nerve: Populist News Articles Lead to More Frequent and More Populist Reader Comments." *Political Communication*, August, 1–23. https://doi.org/10.1080/10584609.2019.1637980.

Blei, David M., Andrew Y Ng, and Michael I Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3 (January): 993–1022.

Dewenter, Ralf, and Jürgen Rösch. 2014. *Einführung in die neue Ökonomie der Medienmärkte: Eine wettbewerbsökonomische Betrachtung aus Sicht der Theorie der zweiseitigen Märkte*. Springer-Verlag.

Druckman, James N., and Michael Parkin. 2005. "The Impact of Media Bias: How Editorial Slant Affects Voters." *The Journal of Politics* 67 (4): 1030–49. https://doi.org/10.1111/j.1468-2508.2005.00349.x.

Eberl, Jakob-Moritz. 2018. "Lying Press: Three Levels of Perceived Media Bias and Their Relationship with Political Preferences." *Communications*, March. https://doi.org/10.1515/commun-2018-0002.

Eberl, Jakob-Moritz, Hajo G. Boomgaarden, and Markus Wagner. 2017. "One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences." *Communication Research* 44 (8): 1125–48. https://doi.org/10.1177/0093650215614364.

Gentzkow, Matthew A., and Jesse M. Shapiro. 2004. "Media, Education and Anti-Americanism in the Muslim World." *Journal of Economic Perspectives* 18 (3): 117–33. https://doi.org/10.1257/0895330042162313.

Gentzkow, Matthew, Bryan T. Kelly, and Matt Taddy. 2017. "Text as Data." Working Paper 23276. National Bureau of Economic Research. https://doi.org/10.3386/w23276.

Griffiths, Thomas L., and Mark Steyvers. 2002. "A Probabilistic Approach to Semantic Representation." *Proceedings of the Annual Meeting of the Cognitive Science Society* 24 (24). https://escholarship.org/uc/item/44x9v7m7.

———. 2004. "Finding Scientific Topics." *Proceedings of the National Academy of Sciences* 101 (suppl 1): 5228–35. https://doi.org/10.1073/pnas.0307752101.

Grimmer, Justin, and Brandon Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21: 267–97.

Groseclose, Tim, and Jeffrey Milyo. 2005. "A Measure of Media Bias." *The Quarterly Journal of Economics* 120 (4): 1191–1237. https://www.jstor.org/stable/25098770.

Hofmann, Thomas. 1999. "Probabilistic Latent Semantic Indexing." In *Proceedings of the 22Nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 50–57. SIGIR '99. New York, NY, USA: ACM. https://doi.org/10.1145/312624.312649.

Kepplinger, Hans Mathias, and Marcus Maurer. 2004. "Der Einfluss Der Pressemitteilungen Der Bundesparteien Auf Die Berichterstattung Im Bundestagswahlkampf 2002." In *Quo Vadis Public Relations? Auf Dem Weg Zum Kommunikationsmanagement: Bestandsaufnahmen Und Entwicklungen*, edited by Juliana Raupp and Joachim Klewes, 113–24. Wiesbaden: VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-83381-5_9.

Lott, John R., and Kevin A. Hassett. 2014. "Is Newspaper Coverage of Economic Events Politically Biased?" *Public Choice* 160 (1): 65–108. https://doi.org/10.1007/s11127-014-0171-5.

Mimno, David, Hanna M. Wallach, Edmund Talley, Miriam Leenders, and Andrew McCallum. 2011. "Optimizing Semantic Coherence in Topic Models." In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 262–72. EMNLP '11. Stroudsburg, PA, USA: Association for Computational Linguistics. http://dl.acm.org/citation.cfm?id=2145432.2145462.

Newman, David, Jey Han Lau, Karl Grieser, and Timothy Baldwin. 2010. "Automatic Evaluation of Topic Coherence." In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 100–108. HLT '10. Stroudsburg, PA, USA: Association for Computational Linguistics. http://dl.acm.org/citation.cfm?id=1857999.1858011.

Newman, Nic, Richard Fletcher, Antonis Kalogeropoulos, David Levy, and Rasmus Kleis Nielsen. 2018. "Reuters Institute Digital News Report 2018." Reuters Institute for the Study of Journalism. http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf?x89475.

Ramage, Daniel, Susan Dumais, and Daniel Liebling. 2010. *Characterizing Microblogs with Topic Models.*

Rehs, Andreas. 2020. "A Structural Topic Model Approach to Scientific Reorientation of Economics and Chemistry After German Reunification." *Scientometrics* 125 (2): 1229–51. https://doi.org/10.1007/s11192-020-03640-0.

Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoldi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111 (515): 988–1003. https://doi.org/10.1080/01621459.2016.1141684.

Roberts, Margaret, Brandon Stewart, and Dustin Tingley. 2016. "Stm: R Package for Structural Topic Models." *Journal of Statistical Software* forthcoming (December).

Strömbäck, Jesper. 2008. "Four Phases of Mediatization: An Analysis of the Mediatization of Politics." *The International Journal of Press/Politics* 13 (3): 228–46. https://doi.org/10.1177/1940161208319097.

Taddy, Matt. 2012. "On Estimation and Selection for Topic Models." In *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*.

Takens, Janet, Wouter Atteveldt, Anita van Hoof, and Jan Kleinnijenhuis. 2013. "Media Logic in Election Campaign Coverage." *European Journal of Communication* 28 (June): 277–93. https://doi.org/10.1177/0267323113478522.

Wallach, Hanna M., David M. Mimno, and Andrew McCallum. 2009. "Rethinking LDA: Why Priors Matter." In *Advances in Neural Information Processing Systems 22*, edited by Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, 1973–81. Curran Associates, Inc. http://papers.nips.cc/paper/3854-rethinking-lda-why-priors-matter.pdf.