Algorithmic Selection and Supply of Political News on Facebook

Marcel Garz, *Jönköping International Business School*

Ferenc Szucs, *Stockholm University*

January 2023

We thank Anna Kerkhof, Greg Martin, Bei Qin, José Reis, Andrey Simonov, Daniel Stone, David Strömberg, Maiting Zhuang, and seminar participants in Hohenheim, Jönköping, and Perth for helpful comments and suggestions. We gratefully acknowledge funding by Stiftelsen Inger, Arne och Astrid Oscarssons Donationsfond. Corresponding author: Ferenc Szucs, Stockholm University, Department of Economics, 106 91 Stockholm, Sweden. Email: ferenc.szucs@ne.su.se.

Abstract

Facebook has been criticized for exposing its users to low-quality and harmful information, including fake news, hate speech, and politically one-sided content. In December 2013 and again in August 2014, the platform updated its news feed algorithm to increase user exposure to quality content of news publishers, while curbing the proliferation of non-informative posts. This paper uses a sample of German newspapers to investigate the conjecture that these modifications raised the incentives to publish quality news stories on the platform, focusing on the number and diversity of news story posts about substantive political issues. Using the newspapers' print editions as a counterfactual, our results indicate an increase in the amount of substantive political news on Facebook by approximately 30%. This expansion occurred in a politically balanced way, except that the outlets disproportionately increased their Facebook coverage of the formerly underrepresented Linke (Left Party). Consequently, the within-outlet concentration of political viewpoints decreased by about one half of the standard deviation of our concentration indices.

Keywords: algorithmic curation; diversity; news quality; political knowledge; social media; voting

JEL classification: D22; D72; D83; L82; L86; M31

1. Introduction

Over the past decade, Facebook has gained a crucial role in the dissemination of politically relevant news and the formation of public opinion. The platform's news feed algorithm helps users to distinguish between important and unimportant content, but the selective exposure to information has been criticized for causing filters bubbles. That is, users are mostly exposed to belief-confirming but not counter-attitudinal information, a pattern that could be detrimental to civic discourse (e.g., Pariser, 2011; Sunstein, 2017). In addition, there are concerns that Facebook contributes to the dissemination of hate speech (Müller and Schwarz, 2020a), conspiracy theories (Stecula and Pickup, 2021), and mis- and disinformation (e.g., Allcott and Gentzkow 2017; Chiou and Tucker, 2018; Guess et al. 2018). Over the years, the company has implemented various counter measures, such as training its employees to identify posts that encourage violence (Vega, 2013) and allowing users to flag fake news (Facebook, 2015), but there is a lack of scientific evidence on the effects of those actions.

In this paper, we investigate two major changes by Facebook of its news feed algorithm that had the objective to promote the dissemination of quality news stories on the platform: In December 2013 and again in August 2014, Facebook announced that quality content from news publishers would be more often shown in users' feeds, while downgrading meme photos, clickbait, and status updates (Facebook, 2013, 2014). As discussed in Section 2, the algorithm updates could have created incentives for news publishers to increase both the quantity of their postings and the quality of the content on Facebook. A quantity effect may have resulted from increased returns per post (e.g., additional website referrals), due to a decrease in competition for news publishers from other (non-news) content creators on the platform. A quality effect could have been induced by Facebook selectively increasing publishers' returns for high-quality content more than for low-quality posts.

We investigate these conjectures using data on the Facebook postings of 37 German newspapers between January 2013 and June 2017. We evaluate both the quantity of news as well as newspapers' tendency to publish politically relevant stories. In the economics literature, the quantity of journalistic content is often used as a proxy for news quality, based on the premise that more content is better than less (e.g., Berry and Waldfogel, 2010; Cagé, 2020). Our focus on politically relevant stories is motivated by their importance for representative democracy and their role in assessing news quality (e.g., Bachmann et al., 2021).

We treat high-quality "substantive" political stories as those including expressions that are often used in Germany's political discourse. As shown in Section 3, we retrieve these expressions from the election manifestos of the country's political parties. We argue that a frequent usage of these expressions is an indication that newspapers convey information about the core topics and ideas of political parties to citizens, which allows the electorate to acquire quality knowledge about political issues.[1] However, if a newspaper uses political expressions[2] in a one-sided way (i.e., concentrated on the topics and ideas of only one party), this could be an indication of bias[3], which implies poor news quality as citizens may not be optimally informed. In contrast, using political expressions in a balanced way is likely an indication of high quality because readers may access diverse viewpoints.

Estimating the effect of Facebook's algorithm updates on newspapers' behavior is difficult because adjustments in news postings could be driven by other events taking place at the same time, such as changes in the media agenda or political landscape. To avoid omitted variable bias, we use the newspapers' print editions as a counterfactual.[4] That is, we compare our measures of supply of political news for the time before and after the algorithm updates on Facebook with changes in the outlets' print articles. We find that the overall number of posts published on the platform followed an upward trend throughout our investigation period. This trend remained unaffected by the algorithm updates though, likely because newspapers did not find it advantageous to indiscriminately increase the quantity of all content. However, our results indicate that the outlets responded to the algorithm updates by selectively expanding the amount of substantive political news. Difference-in-differences estimates indicate an increase in political posts by approximately 30%, compared to the number of political print articles.

We show that this news expansion mostly occurred in a politically balanced way. That is, to a large degree, the newspapers increased their provision of political posts randomly across parties. However, we also find that the outlets actively changed their editorial policies, as they disproportionately expanded their Facebook coverage of topics and ideas pertaining to the Linke (Left Party), which has been underrepresented on Facebook. Both effects caused the Facebook coverage to become more balanced across parties. Our baseline estimates indicate a decrease in within-outlet concentration of party-related coverage by approximately one half of the standard deviation of our concentration measures. Simulations show that about three quarters of

[1] Party manifestos typically include substantive information and policy-specific facts (e.g., in relation to crime, taxes, education, and the environment). News coverage about these issues is arguably more valuable for imparting the kind of knowledge that is beneficial for collective decision-making than coverage of non-substantive issues, such as gossip about politicians' private lives or discussions of the fashion choices of female representatives (e.g., Gilens, 2001; Dunaway, 2008; Barabas and Jerit, 2009; Garz, 2018).

[2] In the interest of keeping the language simple, we use the term "political expressions" to refer to the expressions extracted from party manifestos, and "political news" or "political posts" to refer to stories that include them.

[3] This idea is extensively discussed in the media bias literature, although most existing studies investigate bias in a dichotomous way, such as Republican vs. Democratic (e.g., Gentzkow and Shapiro, 2010; Puglisi and Snyder, 2015; Martin and Yurukoglu, 2017), left vs. right (Gans and Leigh, 2012; Szeidl and Szucs, 2021), or pro- vs. anti-government (Qin, Strömberg, and Wu, 2018; Simonov and Rao, 2018), whereas our empirical setting involves five political parties.

[4] We prefer to use the outlets' print coverage as a counterfactual, rather than their postings on another social platform (e.g., Twitter), because spillovers within the outlets' social media departments are more likely than spillovers to the main editorial offices, as we argue in Section 3.1.

the observed decrease in concentration was driven by larger "sample" sizes of daily news stories selected by the newspapers for publication on Facebook (possibly from their pool of available content produced for the print editions). As predicted by the law of large numbers, the increase in the number of political posts caused the distribution across parties to converge to the politically more balanced distribution in the print editions. The remaining quarter of the decrease in concentration of party-specific coverage can be explained by the higher number of political posts pertaining to the underrepresented Linke (Left Party).

While the focus of our paper is on the impact of the algorithm updates on news outlets' usage of Facebook, we present circumstantial evidence of potential downstream effects. Using survey data from the German Longitudinal Election Study, we find that Facebook users' interest in politics and knowledge of political candidates, parties, and coalitions improved after the algorithm updates, relative to non-users. Our estimates also suggest that self-reported participation in federal elections increased, all of which can be cautiously interpreted as an indication that Facebook's intervention may have been beneficial for society.

We contribute to multiple strands of literature. First, our study relates to research discussing the repercussions of news consumption via social platforms for the news industry. Several studies investigate whether publishers can exploit social media to generate advertising and subscription sales via website referrals (Hong, 2012; Mahmood and Sismeiro, 2017). For instance, Sismeiro and Mahmood (2018) use a natural experiment to show that exposure to news stories on Facebook increases traffic on publishers' websites, which diminishes concerns of audience-stealing by platforms. Lischka and Garz (2021) use a game theoretic approach to investigate the interplay between platforms, media outlets, and users regarding clickbait supply and consumption. Dujeancourt and Garz (2022) investigate how Twitter's introduction of algorithmic curation affected user engagement with news stories and find that likes and shares increased more for sensationalist headlines than quality content. Others examine if media companies' pressure to create revenues affects editorial decisions (Myllylahti, 2020; Peterson-Salahuddin and Diakopoulos, 2020). For example, as Cagé, Hervé, and Mayozer (2020) show, the popularity of stories on Twitter influences decisions of journalists and editors of what news stories to produce. We do not investigate whether social media affect newsroom decisions, but our results speak to publishers' selection process regarding news stories to be disseminated on Facebook (Wortelker, 2021), and we show that media outlets act in accordance with the incentives created by platforms.

Second, our results extend previous evidence on effects of online platforms' user policy on content creation. For example, Mayzlin, Dover, and Chevalier (2014) study differences in user review policy on Expedia.com and TripAvisor.com, whereas Mousavi and Zhao (2018) investigate a change in Airbnb.com's review procedures. Both studies show that the design of these rules affects the sentiment of user reviews. Sun and Zhu (2013) and Kerkhof (2019) investigate effects of changes in advertising policies of Sina.com and YouTube on the content of blog posts and videos, respectively. Huang, Hong, and Burtch (2017) find that a Facebook integration offered by Yelp.com and TripAdvisor.com increased the quantity of user reviews and led to more positive sentiments. According to Cavusoglu et al. (2016), Facebook users became more inclined to openly share content once the platform offered granular privacy controls. All previous studies have in common that they focus on the effects of platform policy on user-generated content. To our knowledge, our study is the first to provide evidence of effects on professional content creators.

Third, our study adds to research on the economics of news quality (e.g., Angelucci and Cagé, 2019; Cagé, 2020; Djourelova, Durante, and Martin, 2021). This research usually approximates quality by looking at the number of articles or pages of a newspaper, the number of employed journalists, the ratio of hard news vs. soft news, and the diversity of news categories. We follow the existing literature in that we measure news quality in terms of quantifiable characteristics but focus on a different dimension of news quality that has been neglected in economics – the quantity of political news and the concentration of political viewpoints. This aspect of news quality has been extensively discussed in communications research (e.g., Wellbrock, 2011; Humbrecht and Büchel, 2013; Bachmann et al., 2021), where most empirical applications are based on qualitive methods and content analyses by human coders. Our approach is closely related to automated, language-based methods of measuring media slant (e.g., Gentzkow and Shapiro, 2010), as we retrieve expressions from political reference texts. However, our study differs in that we consider newspapers' usage of political expressions as an indicator of how much attention and resources they devote to substantive political issues. Importantly, our approach allows us to analyze concentration of political viewpoints in a large sample of news items and at the level of individual posts.

Fourth, our results relate to research on the role of social media for political outcomes. A large body of literature provides evidence that social media penetration facilitates collective action and protest participation offline; see Zhuravskaya, Petrova, and Enikolopov (2020) for a review. Others investigate the role of social media for xenophobia and hate crimes (e.g., Bursztyn et al. 2019; Müller and Schwarz, 2020a, 2020b). More closely related are studies that examine the impact of social media on political knowledge and mobilization. Mosquera et al. (2020) present experimental evidence that users who abstain from Facebook for a week become less competent in evaluating politically biased news. Allcott et al. (2020) find that deactivating Facebook reduces news knowledge but does not affect turnout. Using randomized controlled trials, Bond et al. (2012) and Jones et al. (2017) show that exposure to political mobilization messages on Facebook has positive effects on information seeking and turnout. Fujiwara, Müller, and Schwarz (2021) investigate the impact of Twitter penetration on voting in the US and find small increases in turnout. Rotesi (2019) confirms this finding and additionally shows that knowledge about local politics decreased in areas with high Twitter penetration, likely due to crowding out of consumption of traditional news sources. While our evidence about the political effects of social media cannot be interpreted in a causal way, we shed light on a mechanism through which social media usage can affect political knowledge and participation: exposure to news about substantive political issues.

2. Competition for attention and algorithmic content selection on Facebook

Facebook has been using an algorithm to select content into users' feeds since the public launch of the platform in 2006. For commercial reasons, the exact workings of the news feed algorithm are not shared with the public. However, there is a general consensus among marketers and content creators about the most important factors influencing the chances that posts are shown to users, including previous interactions with the publishing source, the timeliness of the content, the type of media in the post, and the popularity of the content among users' Friends (e.g., DeVito, 2017).

The news feed algorithm is constantly subject to minor tweaks and occasionally modified in more substantial ways. It is not clear how reliably Facebook announces these modifications, but representatives of the company have stated that changes expected to impact the number of users exposed to content of commercial publishers by more than 10-20% (of that number) would be communicated to the public (Kafka, 2013). These communications are archived in Facebook's "newsroom" (https://about.fb.com/news/).

During the relevant time frame (Jan 2013 – Jun 2017), we identify two modifications of the algorithm that explicitly targeted journalistic content published by professional media companies. On December 2, 2013, the platform announced an update labeled "Helping You Find More News to Talk About", according to which "high quality articles about current events" would be more prominently featured in users' feeds (Facebook, 2013). In contrast, users would be less often exposed to meme photos. In the announcement, Facebook illustrated what kind of content would be prioritized by showing a healthcare-related post of the US news magazine *The Atlantic*, displayed in Facebook's *link format*.

On August 25, 2014, the platform announced that content published via its link format would be prioritized – especially in the context of news stories – while downgrading so called status updates and posts with links shared in the text caption (Facebook, 2014). The link format strongly resembles the presentation of news stories on news outlets' websites, with a headline on top, followed by a picture and the first few lines of text of the article; see Figure A1 in the Online Appendix. Hence, the algorithm update favored a format that is very compatible with the way online teasers for news stories typically work. The link format is considered user friendly because it only shows the name of the publishing outlet. In contrast, posts that share a link in the text caption are difficult to read because they show the entire URL of the linked news article. Spot checks suggest that the downgraded post formats have been often used to post live soccer scores, greetings ("Hello everybody!"), and generic messages ("New York Times updated their profile picture"), rather than quality news stories. The update also included some modifications to curb the proliferation of clickbait. However, this aspect of the modification was likely irrelevant for the newspapers in our sample, as it targeted Pages with a dedicated focus on clickbait, such as Upworthy and Buzzfeed. In fact, Lischka and Garz (2021) show that the clickbait supply of German legacy media on Facebook was negligible at the time (i.e., less than 6% of all posts) and did not change after the update.

We argue that the algorithm updates decreased competition for (quality) news stories from other content on the platform, because of increased chances that these stories are selected in users' feeds. Higher chances of selection imply greater reach. Abstracting from the possibility that newspapers pay Facebook to boost content, the old version of the news feed algorithm organically exposed a certain number of users to any given news story post, especially among followers of the publishing newspaper. With the new version, the number of users exposed to this kind of post should be higher, considering the increased probability that the news feed algorithm selects the post. As a consequence, newspapers' returns to posting a news story on Facebook should have increased due to the algorithm updates, because greater exposure typically translates into higher user engagement (e.g., likes and shares), website referrals, and eventually advertising and subscription revenues (Mahmood and Sismeiro, 2017; Sismeiro and Mahmood, 2018). We provide evidence in support of this assumption in Section 4.2.3, by showing that user engagement with news stories increased more after the algorithm updates than user engagement with similar content posted by the Facebook Pages of political parties.[5]

Thus, the algorithm updates likely increased the returns to posting news stories on Facebook. At the same time, the costs of publishing these posts remained unchanged. News outlets typically post a small fraction of the content produced for their print and online editions on Facebook. This task involves the fixed cost of employing one or more social media editors and a variable cost of adapting the content for Facebook (e.g., tweaking the headline or changing the photo so that content is better aligned with the preferences of Facebook users). Normally, newspapers do not create much additional content for Facebook, but they simply exploit the pool of available news stories (Wortelker, 2021). Our data do not allow us to determine to what degree newspapers "recycle" their print content and to what extent they create original content for Facebook.

[5] It is also possible that the algorithm updates intensified competition between providers of news content. For instance, the increase in "market size" for quality news on Facebook could have induced entry of new competitors or caused established outlets to strive for greater market shares. While we cannot rule out shifts in market shares on Facebook between the outlets in our sample, the results in Section 4.2.3 suggest that any additional competition from market entrants was not strong enough to prevent the outlets in our sample at large to benefit from the algorithm updates in the form of increased user engagement.

However, comparing the average monthly number of Facebook posts (ca. 435; see Table A2) with the number of print articles (ca. 9,153) suggests that newspapers have plenty of reserve content that could be cross posted on Facebook at a presumably low variable cost.[6]

If the algorithm updates raised newspapers' returns to posting on Facebook by a larger extent than the variable publishing cost, newspapers should respond by increasing the daily number of news story posts, until marginal costs equal marginal revenues. However, due to Facebook's covertness about the specifics about the algorithm updates, it remains unclear whether any content by news publishers experienced a decrease in competition, or whether only news stories of a certain quality were affected. In the former case, we expect an increase in the overall number of posts, whereas in the latter case outlets have an incentive to selectively expand their provision of quality news stories, such as those about substantive political issues.

Facebook justified both algorithm updates with user demand for high quality news. At the time, the platform had become an important source of news consumption for about one third of US adults (Pew, 2013), and internal user surveys suggested that prioritizing journalistic content would have a positive impact on engagement metrics (Facebook, 2013; 2014). Thus Facebook's motivation for the algorithm updates was unrelated to German politics. The platform announced further major modifications of its algorithm during the relevant time frame. For instance, in August 2013, Facebook revealed that already engaging content would be further boosted by the algorithm; in April 2014, the platform implemented measures against posts asking for likes or shares ("engagement baiting"). However, we do not investigate those modifications as they do not explicitly target journalistic content published by professional media companies. Other measures taken by Facebook that could be relevant for news outlets (e.g., in response to privacy violations, fake news, and hate speech) took place after our investigation period.

3. Data and measurement

3.1 Newspaper sample

Our analyses are based on data from Germany. This choice of country is motivated by two reasons. First, Germany's multi-party system allows us to investigate concentration of viewpoints across the political spectrum, rather than the balance between two factions as would be the case in the context of a political two-party system. Second, a concern with using the newspapers' print editions as a counterfactual is that any effects of the algorithm updates could impact the offline branch of the outlets, especially that optimizing

[6] There are limits to how much content a newspaper should optimally post on Facebook though. If the outlet posts too much or too often, the probability that the content is selected in users' feeds decreases algorithmically because Facebook wants to avoid that users feel spammed by a brand.

content for Facebook spills over from the social media department to the main editorial office. A spillover would bias the estimates towards a zero effect, due to decreasing differences between Facebook and print. This is not a likely scenario in the German context though, because the outlets in our sample all have a long tradition in print journalism and a strong ethical sense for reporting standards (Hanitzsch et al., 2011). At the time, there was widespread skepticism in the industry about disseminating news via social platforms and a consensus that journalistic norms must not be corrupted (Cision, 2013). In contrast to news industries in other countries, Facebook traffic accounted for a very small fraction of newspaper revenues in Germany (Newman et al., 2016).

Most major German newspapers obtained their Facebook Page between 2010 and 2012 (Cision, 2013). We choose January 2013 as the starting point of our estimation sample, because there are many outlet-months with very few or zero news story posts before that. However, as of January 2013, most newspapers had developed a routine of posting several times per day. The estimation sample ends in June 2017, shortly before we started collecting the Facebook data. Restricting the analyses to these start and end dates gives us long enough pre- and post-treatment periods to investigate the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Data on the news output of the outlets' print editions come from the Genios database, the most comprehensive archive of German press coverage (www.genios.de). Our sample comprises all newspapers that were consistently archived in Genios throughout our period under investigation. This criterion is met by 37 outlets, including all daily national newspapers (except for Bild[7]), the national weeklies Focus, Der Spiegel, and Die Zeit, as well as the largest regional outlets. At the time, the total number of newspapers in Germany was 361 (KEK, 2015). We consider complementing the sample with data from other news archives, such as Factiva and Nexis, but any sources included in these databases are already contained in Genios. The number of outlets in our sample is rather small – compared to the total number of newspapers – but the selection covers a large number of different publishing companies, and it includes the most-read and agenda-setting newspapers. According to their circulation, the newspapers in the sample account for approximately 40% of the German market for print news (e.g., KEK, 2015). See Table A1 in the Online Appendix for details.

We download information about all 856,532 posts by the official Facebook Pages of the outlets in the period under investigation, using the Facebook Graph Application Programming Interface (API). For each post,

[7] Bild is Germany's largest tabloid. We do not have access to data on the outlet's news output via Genios or any other newspaper database. However, our sample comprises various outlets owned by Bild's publisher, Axel Springer, including the national newspapers Die Welt and Die Welt Kompakt, as well as the local daily B.Z. Berlin.

we obtain the post message, type and format of the post, date and time of publication, the number of likes, shares, and comments, and the URL linking to external content, if applicable.[8]

3.2 News quantity

We use the Facebook data to compute the overall number of posts by outlet and month. For commercial reasons, the Genios database does not offer bulk downloads of newspaper articles.[9] We therefore use the browser interface of the database to obtain counts of articles at the outlet-month level. To assess to what degree the total number of news items reflects news quality, we compare newspapers' mean values of this metric with a ranking of journalistic quality provided by Wellbrock (2011). This ranking is based on surveying media researchers and journalists about the outlets' perceived quality in seven areas, including accuracy, comprehensiveness, diversity, independence, intelligibility, relevance, and timeliness. Based on the survey responses, the ranking computes an overall quality score that ranges from 4.67 (lowest observed quality) to 8.38 (highest observed quality) and is available for 26 newspapers in our sample.

Figure 1 shows the relationship between outlets' news quantities and their perceived quality. According to Panel A, there is a moderate positive correlation on Facebook, which implies that newspapers with a higher perceived quality tend to post more on the platform. Panel B indicates no relationship between the number of print articles and perceived quality. Hence, the overall number of news items might not be a good proxy of news quality in our context. We consider this measure regardless because it is useful to investigate the mechanisms driving our findings. In Figure 2, we plot the sample mean of the number of posts and articles over time. The figure indicates a rather constant news output of the print editions, whereas the monthly number of Facebook posting followed an almost linear upward trend throughout the sample period.

3.3 Political news items

We identify posts and articles about substantive political issues by checking whether they include expressions that are typically used by political parties, as these expressions are likely highly relevant for civic discourse, opinion formation, and collective decision-making. In contrast to topic modeling, for example, this approach allows us to measure the concentration of viewpoints in news. We consider all parties that were represented in Germany's federal parliament (Bundestag) during our period of investigation, including the Christian Democratic Union (CDU) and its Bavarian counterpart Christian Social Union (CSU), the Free

[8] The Facebook data are also used by Lischka and Garz (2021) to investigate the outlets' provision of clickbait.
[9] This limitation also restricts our options to investigate other proxies of news quality (e.g., number of words, hard vs. soft news, use of sensationalist language) because we cannot create those measures for the print editions.

Democratic Party (FDP), the Green Party (Grüne), the Left Party (Linke), and the Social Democratic Party (SPD).[10]

In Germany, each party publishes its election manifesto – sometimes also referred to as party program or party platform – a few months before an upcoming election. Election manifestos are better suited to track changes in news content over multiple years (as in our case) than more fast-paced reference texts (e.g., press releases, interviews, or parliamentary speeches), because the former do not much refer to current events but tend to describe the kind of topics and ideas that are indicative of parties' long-run ideologies. For instance, as shown in Figure A2, expressions such as "nuclear phase-out" and "climate crisis" pertain to core topics of the Green Party, whereas the market liberal FDP traditionally cares about "freedom of contract" and the "tax factor". The Social Democrats emphasize topics that are relevant for workers and families, such as "solidary retirement benefits" and "family working time", while the conservative CDU/CSU calls attention to "demographic change" and "volunteer work". In contrast, the Left Party uses expressions such as "minimum income" and "militarization" to describe long-run issues on its political agenda. Figure A3 illustrates that Facebook posts with political expressions tend to address more substantive issues than posts without those expressions.

Using the manifestos pertaining to the 2013 and 2017 national elections, we follow Garz, Sörensen, and Stone (2020) and apply the product of the term frequency and the inverse document frequency (TF-IDF) to identify expressions that often occur in one text but rarely in others (Jurafsky and Martin, 2008; Gentzkow, Kelly, and Taddy, 2019). In our context, these are terms that appear frequently and uniquely in the parties' election manifestos. On average, the manifestos contain 47,162 words and thus constitute a rich resource of text. We remove punctuation, numbers, stop words, formatting, and party references from the elections program, reduce the remaining words to their stem, and compute the TF-IDF as

$$TFIDF = \frac{f_{e,p}}{F_p} \times \log\left(\frac{P}{pf_e}\right) \tag{1}$$

where f denotes the frequency of expression e used by party p, F_p is the total number of words per party, $P = 5$ indicates the number of parties (CDU and CSU publish joint election programs), and pf counts the number of election programs including expression e. For each party, we retain those terms that fall in the

[10] We do not include the Alternative for Germany (AfD), which was established during our sample period and gained the third highest vote share in the 2017 national elections. The reason is that the party did not have a manifesto for the 2013 elections, where it failed to exceed the 5% electoral threshold, but only a 4-page flyer. The 2017 manifesto does not reflect the party's ideology at the time of Facebook's algorithm changes in 2013 and 2014, because the AfD in its current form as a far-right party started to exist only in mid-2015, when its founder left the party and was replaced by new leaders with different topics and ideas.

top 0.1% of the distribution of TF-IDF values. This is the optimal cut-off determined by Garz, Sörensen, and Stone (2020) based on various benchmarking exercises.

Next, we obtain counts of news items (i.e., Facebook posts or print articles) that contain these expressions in the post message or article text (labeled $items^{match}$). That is, on Facebook, we search for matches in the post message, whereas we examine headline and article text in the case of the print output. We do not include the (online) articles linked in the Facebook posts for the following reasons. First, most outlets in our sample use a paywall or have the access to their online content restricted in some other way. In addition, a fraction of linked articles are not online anymore, and the websites of the smaller outlets are not consistently archived by services such as https://web.archive.org/. Thus it would only be possible to analyze a highly skewed sample of online articles. It is not feasible to analyze the content of the print counterparts of the Facebook posts instead because some posts are not based on any print content, such as greetings or live soccer scores, and due to restrictions of Genios database. Second, the updates of the algorithm primarily targeted the posts, not the linked articles. Hence, we should expect for the updates to mainly affect the former. Third, and perhaps most importantly, users usually do not click on the linked content but only process the post message (e.g., Dor, 2003; Gabielkov et al., 2016). As for the print content, it is certainly possible to restrict the analyses to article headlines. Doing so leads to very similar results than evaluating the entire article (see Tables A4 and A5) but the resulting measures of news quality correlate much less with Wellbrock's (2011) benchmark index.

Figure 1, Panels C and D, show the relationship between newspapers' tendency to use political expressions in their coverage and survey-based news quality. The graphs reveal a strong positive correlation between these indices, both on Facebook and in print. That is, outlets with a high perceived journalistic quality tend to publish more political news items. As Figure 3 shows, the newspapers expanded the absolute number of political posts throughout the sample period, but especially after the second algorithm update. In relative terms, the share of political posts among all posts followed a slight downward trend (Figure A4, Panel A)[11], which implies that the overall number of posts increased faster than the number of political posts. We observe a slight increase of the share of political news items in the print editions (Panel B).

[11] Also, we can observe a slight upward shift relative to the downward trend after the second algorithmic change.

3.4 Within-outlet concentration of viewpoints

We normalize the counts of political news items ($items^{match}$) by the overall number of posts or articles ($items^{total}$), yielding shares of political news items x containing expressions associated with party p, by outlet i's channel c (Facebook or print) at time t:

$$x_{p,i,c,t} = \frac{items_{p,i,c,t}^{match}}{items_{i,c,t}^{total}} \tag{2}$$

Considering that profit-maximizing media companies tend to cater to the preferences of their consumers (Gentzkow and Shapiro, 2010; Puglisi and Snyder, 2015), newspapers should devote more coverage to topics and ideas of parties that are more popular among their readers. Testing this conjecture helps us to validate our approach. We capture reader preferences by using voting data from the 2013 and 2017 national elections, as provided by the Federal Returning Officer at the level of Germany's 299 electoral districts. Information from the German Newspaper Publishers Association allows us to determine the electoral districts in which the outlets have their main area of circulation[12], based on which we match outlets and local voting data.

Figure A5 shows the relationship between the outlets' use of political expressions and matched vote shares. Accordingly, party-specific shares of political news items are positively correlated with the parties' popularity among the outlets' readers. The strength of the relationship is quite similar when comparing Facebook and print, as well as 2013 and 2017 vote shares, with correlation coefficients between 0.299 and 0.361. Of course, the raw correlations could merely reflect overall differences in political expressions and vote shares (e.g., outlets are generally more likely to use expressions pertaining to CDU/CSU and SPD, simply because these parties have the highest average vote shares). However, as Table A3 indicates, we also find a positive and statistically significant relationship when we regress the shares of political news items on voting, while including outlet and party fixed effects. As another benchmark, we compare outlets' use of political expressions with their tendency to explicitly reference political parties (i.e., mentions of parties' names). Figure A6 illustrates that newspapers are more likely to reference a specific party, the more often they use expressions related to that party.[13] The congruence between outlets' use of party-related expressions and party

---

[12] See https://www.die-zeitungen.de/media/mediadaten/pdf-preislisten.html. Note that we cannot include the national outlets in our sample when conducting this validity check, because their circulation is not restricted to any specific electoral district(s), and we do not have other data that would allow us to determine the ideological composition of their readers.

[13] In addition, counts of Facebook posts and print articles including party mentions followed similar developments over time than the number of news items containing political expressions, with a relatively even trend in print and an upwards trajectory on Facebook; see Figure A13.

mentions alleviates concerns that our language-based approach assigns news items to the "wrong" party or mistakenly classifies non-political content as political news.

The shares x of political news items can then be used to compute the diversity of posts and articles across parties, per outlet and month, using standard indices of information concentration (e.g., McDonald and Dimmick, 2003). Our main measure is a "Gini-style" index. Omitting outlet and time indices, this measure evaluates the sum of absolute differences between all pairs p and q of party-specific shares x of political items relative to the sum of these shares:

$$Gini = \frac{\sum_{p=1}^{n} \sum_{q=1}^{n} \left| x_p - x_q \right|}{\sum_{p=1}^{n} x_p} \tag{3}$$

An alternative approach is to calculate the relative standard variation (RSD), which compares the variance in the shares of political news items with the mean of these shares μ in a given outlet-month:

$$RSD = \frac{\sum_{p=1}^{n} (x_p - \mu)^2}{\mu} \tag{4}$$

The indices are computed in a way that small values indicate a more balanced coverage, whereas large values reflect one-sided reporting; see Table A2 for summary statistics. A graphical representation of changes in the Gini concentration index can be found in Figure A7. On average, the Facebook posts were less balanced than the print articles (mean Gini scores of 2.06 and 1.32, respectively). The concentration scores followed relatively flat trends on both channels until August 2014, although the development was more volatile on Facebook. Afterwards, the trend remained unchanged for the print articles but dropped by about 1 point (or 30%) for the Facebook posts. Thus, news stories posted on the platform became more balanced at that point.

3.5 Complementary survey data

To evaluate potential downstream outcomes of Facebook's algorithm update, we complement the news data with information from the German Longitudinal Election Study (GLES). The GLES continuously collects survey data about political attitudes, knowledge, and media consumption of the electorate. Occasionally, the GLES surveys include questions about respondents' social media usage, as in the case of the 18th and 26th waves of the "Long-term Online Tracking" component of the GLES (Rattinger et al., 2014, 2015). The 18th wave (survey period: 17.09.2012 – 01.10.2012) is the closest survey that includes questions about

Facebook usage before the platform's first algorithm update took place, while the 26th wave (survey period: 21.11.2014 – 05.12.2014) is the closest survey with information on Facebook usage after the second algorithm update. The surveys were conducted as a rolling cross-section of German residents aged 18 or older who regularly use the Internet to obtain political information. We use the data to distinguish Facebook users and non-users and to construct measures of political knowledge in three categories (candidates, parties, and coalitions), interest in politics, participation in the most recent federal and state elections, and party identification, as well as a battery of demographic controls and media usage; see Table A11 for details.

4. Results

4.1 Main effects

Estimating the effects of Facebook's algorithm updates on the postings of the newspapers in our sample is challenging because the number and distribution of news stories related to different parties could be affected by shocks to the news agenda resulting from political events.[14] For that reason, we use the newspapers' print coverage as a counterfactual. We estimate the effects of the modifications of Facebook's news feed algorithm on outcome variable y of outlet i's channel c (Facebook or print) in month t, using versions of the following difference-in-differences model:

$$y_{i,c,t} = \alpha_1 + \alpha_2 After_t^{Nov2013} \times FB_c + \alpha_3 After_t^{Jul2014} \times FB_c + \alpha_4 \pi_{i,c,t} + \gamma_{i,c} + \theta_t + \varepsilon_{i,c,t} \tag{5}$$

where FB is binary variable that takes the value 1 for outlets' Facebook coverage and 0 for their print coverage. $After^{Nov2013}$ and $After^{Jul2014}$ equal 1 after November 2013 and July 2014, respectively, for the time periods following the updates of the news feed algorithm. The coefficients of interest are on the interactions between the FB and $After$ dummies, α_2 and α_3. We include an outlet-channel-specific time trend polynomial π of order three to address concerns about differential developments between Facebook and print coverage that are unrelated to the updates of the news feed algorithm. All regressions include outlet × channel fixed effects $\gamma_{i,c}$ and a full set of time dummies θ_t. Regressions that involve the concentration indices as the dependent variable can be estimated by using OLS. However, the other outcome variables are overdispersed counts (i.e., the total number of news items and the number of political news

[14] For example, the first algorithm update in December 2013 coincided with the start of the third Merkel cabinet following the 2013 national elections, which may have induced journalists and editors to revise their decisions about how much coverage to devote to the individual parties. Our difference-in-differences model absorbs any common shocks to the news agenda resulting from these and other political events.

items). In that case, we use both OLS – after taking the logarithm of these variables – and negative binomial regression. Throughout, we cluster the standard errors by 74 outlet-channel combinations.

We complement the main regressions with tests for parallel pre-trends, which can be implemented by specifying models that include interactions between the Facebook dummy FB_c and f leads and l lags of the point of time of treatment:

$$y_{i,c,t} = \sum_{j=-l}^{f} a_j \theta_{t+j} \times FB_c + \gamma_{i,c} + \theta_t + \varepsilon_{i,c,t} \tag{6}$$

where θ_t and $\gamma_{i,c}$ are again time and outlet-channel fixed effects. The a_j's are the coefficients of interest, as they capture differences in the outcome variable y between the outlets' Facebook and print channels at different points of time. Significant a_j's before the algorithm updates indicate that the pre-trends are not parallel, which prevents us from attributing any observed post-treatment differences between Facebook and print to the algorithm updates. In contrast, obtaining insignificant pre-treatment a_j's lends support to a causal interpretation of post-treatment changes.

Results of estimating Equation (5) with the overall number of news items as the dependent variable are presented in Table 1. The coefficients on the interaction terms are not statistically different from zero, which implies that the algorithm updates did not affect newspapers' general activity level on Facebook. This finding is confirmed by the corresponding test for parallel pre-trends (see Figure A8, Panel A), according to which the overall number of Facebook posts simply followed a near-linear upward trend during our investigation period, with multiple significant a_j's in the pre-treatment period.

Table 2 shows results for the provision of political news items. Estimates in Columns (1) and (2) indicate increases in the absolute number of political Facebook posts after both algorithm updates, with statistically significant coefficients after the second one. The magnitude of this effect can be assessed after exponentiating the coefficients. For example, the negative binomial estimate in Column (2) of 0.265 implies that the number of political Facebook posts increased by approximately 30.3% after this update (i.e., exp(0.265) – 1 = 0.303).[15] As Panel B in Figure A8 indicates, none of the a_j's in the period before the first algorithm update are statistically different from zero. A test of the joint significance of these coefficients does not contradict the assumption of parallel pre-trends either (joint F-statistic = 1.38; *p*-value = 0.206). However, there are multiple significant a_j's between algorithm updates, which suggests that the increase in politically

[15] We obtain similar results when we use counts of political news items based on explicit party mentions, rather than political expressions; see Table A6.

posts cannot be exclusively attributed to the second one. That is, the effects of the first update may have been delayed and/or become visible only in combination with the second update. We further investigate pre-trends by conducting a placebo test. Focusing on the pre-period (before the first algorithmic change) we repeat our most preferred specifications using multiple "fake" algorithm update events. The results, plotted in Figure A9, speak against any Facebook-specific pre-trends.

The estimates in Columns (3) and (4) of Table 2 show that there is a similar effect when we control for the total number of news items.[16] Here, the relevant point estimates indicate effects sizes of 21.7% and 20.3%, depending on the estimation method (i.e., exp(0.196) – 1 = 0.217 in the case of OLS and exp(0.185) – 1 = 0.203 in the case of the negative binomial model). Hence, we can rule out that the increase in political Facebook posts was mechanically driven by the growth of the overall number of postings. In contrast, the results in Columns (3) and (4) suggest that the share of political posts among all posts significantly increased after Facebook's second algorithm update, after accounting for underlying trends.[17] To confirm that the increased coverage of political news on Facebook relative to print media is driven by changes in the treatment (Facebook) rather than the control (print media) group, in Table A8, we estimate the effects of the algorithm updates on the two groups separately. We only find significant effects for Facebook posts, which confirms that the effects in Table 2 are driven by changes in the treatment group.

We summarize results pertaining to the concentration of political news towards individual parties in Table 3. We do not find any significant effects after the first algorithm update but all specifications indicate significant changes after the second update. The sign of the relevant coefficient is negative, which implies that the concentration of the newspapers' political Facebook coverage decreased. The coefficients do not substantially differ for the models that include (Columns 3 and 4) and do not include the total number of news items as a control (Columns 1 and 2). Hence, the result is not driven by a mechanical effect. Compared to the standard deviations of the concentration measures on Facebook, the coefficients in Columns (3) and (4) indicate treatment effects in the magnitude of 50.6% (Gini) and 43.0% (RSD).[18]

The corresponding test for parallel pre-trends does not indicate any significant a_j's during the pre-treatment period. The F-statistic on the joint significance of the pre-treatment coefficients is low (F = 0.79; *p*-value = 0.638). There are no significant effects after the December 2013 change of the news feed algorithm either, but the graph indicates multiple significantly negative coefficients following the August 2014 update (joint

[16] As a robustness check, in Table A7, we also report estimates on the share of political news items, where we find qualitatively similar results.
[17] After the algorithm updates, the mean value of political posts published by the newspapers on Facebook is 39.31, compared to 5,498.24 political print articles and 1,826.68 political print headlines. Hence, the number of political posts remains relatively small (0.71% and 2.15%, respectively, compared to the print baseline).
[18] The mean Gini score on Facebook is 1.92 after the algorithm updates, which is approximately 1.5 times higher than the degree of concentration in newspapers' print versions (mean Gini score = 1.31).

F-statistic = 8.94; *p*-value = 0.000). This pattern supports the interpretation that the lower concentration of political posts was caused by the August 2014 algorithm update, rather than some unrelated development.

Facebook postings of news outlets are on average substantially shorter than print articles, since the postings are often used as teasers designed to induce clicks. Thus it could be argued that the outlets' print versions are less than an ideal control group for the Facebook Pages. That is, print articles are more likely to contain political expressions simply because they are longer than Facebook posts. To exclude the possibility that our results are driven by this length difference, we obtain counts of print articles that include relevant expressions in their (sub-)headings while ignoring the article texts. Similarly, we compute versions of our concentration measures based on the articles (sub-)headings only. These alternative measures do not correlate as much as our baseline measures with Wellbrock's (2011) benchmark index, which is why we use them only for robustness checks. However, as Tables A4 and A5 show, regressions using the headline-only versions of political news and concentration yield very similar results than our baseline specifications. Thus we can rule out that the length differences in news items between Facebook and print drive our results. We can also rule out that the increase in political news was a mechanical result of the newspapers posting longer text messages. As Figure A10 shows, the average number of characters per post decreased after the algorithm updates.

4.2 Mechanisms

In this subsection, we investigate the reasons for the increase in political postings and the lower concentration of viewpoints after the second algorithm update. In Section 4.2.1, we evaluate the role of the extended use of the link format, whereas Section 4.2.2 discusses changes in editorial policy. Section 4.2.3 provides evidence in support of the premise that the new version of the news feed algorithm increased the returns to posting quality news stories by lowering competition from non-news content.

4.2.1 Link format and law of large numbers

Facebook's press release pertaining to the first update of the news feed algorithm in December 2013 did not explicitly state that news publishers are better off when they use the link format, but the screenshot example chosen by Facebook may have implicitly drawn publishers' attention to this format. With the second algorithm update in August 2014, Facebook explicitly announced to increase users' exposure to posts that use the link format, while downgrading status updates and posts with links shared in the text caption. Figure A1 shows examples of the different types of posts. The figure illustrates that news outlets often used status updates to interact with the audience without providing any political news; for instance, by sending greetings (e.g., “Good morning everybody!”) or posting live soccer scores (e.g., “Goal for Germany!”). Posts including a link in the text caption often contained a very short or no message text, because displaying the URL to the linked article consumes space. In contrast, posts using the link format are more likely to provide proper news stories, including politically relevant content. On the one hand, this format appears to be particularly compatible with the way news stories are typically presented online, including a headline, the beginning of the article text, and a photo. On the other hand, the link format does not show the URL to the linked article but only displays the Internet domain of the publishing source, which leaves more space for the news headline. Given the focus on news and the efficient use of text space, posts using the link format could be more likely to include political expressions than the outdated formats. This mechanism could explain the boost of political posts after the second algorithm update. It could also explain the decrease in the concentration of viewpoints, if the news expansion followed the law of large numbers, meaning that the additional expressions were randomly distributed over the different parties.

Figure A11 confirms that the news outlets raised their use of the link format after both algorithm updates, but especially after the second one. The share of posts using this format increased from around 60% to approximately 90% towards the end of the observation period. In addition, as Table A9 shows, link format posts are more likely to include political expressions, after controlling for the length of the post message. For example, according to Column (1) of the table, the likelihood that a post includes political expressions is 1.5 percentage points higher in the case of the link format than for other types of posts. Compared to the mean probability (ca. 7.5%), the estimate implies an increase of 0.015/0.075 = 20%.

We can think of the outlet’s Facebook posts on any given day as a selection of the news stories published in the print edition on that day. Comparing the total number of news items on Facebook and in print (see Table A2) suggests that this selection covers less than 5% of the available print stories. If we abstract from the possibility that social media editors intentionally provide a different selection of political news items than the main editorial office, the small size of the “sample” of news stories posted on Facebook could imply that the distribution of political expressions across parties included in these posts is a poor approximation of the actual distribution of expressions in the print version. However, the boost in the number of posts using party-specific expressions after the August 2014 algorithm update could have improved how well the distribution of these expressions on Facebook approximates this distribution in the print versions – simply due to the increased “sample size” – as predicted by the law of large numbers.

To investigate this possibility, we compare our results to a simulated counterfactual, which assumes no change in the expected ratio of expressions pertaining to the different parties. We simulate this counterfactual by the following data generating process. As detailed in the Online Appendix, we first calculate the distribution of all combinations of political expressions for each outlet by pooling all political posts during the pre-algorithm-change period. Second, for each outlet-month in our sample we simulate an artificial sample using the pre-period distribution of expressions across parties – computed in the first step – and the empirical number of political posts. Using this simulated sample of posts, we calculate the concentration of political viewpoints for each outlet-month cell. We repeat this procedure for 100 iterations and compute the average concentration index.

The results are summarized in Figure 4, which plots the empirical and the simulated Gini scores over time. The main message of the figure is that the simulations track the empirical time series very closely and they produce similar, although slightly smaller, drops in the Gini score after the August 2014 algorithm change.

To quantify how much of the change in concentration is explained by the law of large numbers, we feed the real data and the simulation into the following difference-in-differences specification:

$$d_{i,s,t}^{Gini} = \beta Emp_S \times Post_t^{Jul2014} + \mu_{i,s} + \delta_t + \varepsilon_{i,s,t} \qquad (7)$$

where $d_{i,s,t}^{Gini}$ is the simulated ($s = 1$) or empirical ($s = 0$) Gini score of outlet i at time t, Emp_S is an indicator for the empirical sample, $\mu_{i,s}$ is an outlet-sample fixed effect, and δ_t is a full set of time effects.

Results are reported in Table 4. Column (1) presents a simple difference-in-differences specification with an indicator for the empirical sample, an indicator for the time after July 2014, and their interaction. Column (2) reports a more conservative specification by including outlet-sample and time fixed effects. Column (3) also includes sample (either empirical or simulated) specific linear time trends.[19] All three specifications point to the same conclusion, that the simulated sample experiences a slightly smaller drop in the Gini score than the empirical sample. The point estimates in Column (1) suggest that the law of large numbers is responsible for about three quarters of the total drop, while the remaining one quarter is explained by changes in editorial polices (i.e., 0.131 / 0.471 ≈ 27.8%).

4.2.2 Changes in editorial policy

Table 5 presents difference-in-differences estimates on the provision of political Facebook posts and print articles by party. With these regressions, we test whether the newspapers expanded their provision of political posts evenly across parties, or whether the expansion favored some party more than others.[20] All regressions include the total number of news items as a control variable. Thus, the estimates refer to the

[19] Adding sample specific polynomial trends results in qualitatively similar results.
[20] In Table A10, we also test whether the supply of political news changed across parties depending on newspapers' dominant ideology. Here we do not find any systematic differences though.

change in posts pertaining to a party relative to the overall increase in Facebook posts during the investigation period. Again, we do not find any significant effects after the first algorithm update, but partially significant changes after the second update. The magnitude of these changes differed across parties, especially when comparing the effect size for the Linke (Left Party) with the effect sizes for the other parties. For example, according to the OLS estimates in Panel A, the coefficient pertaining to the Linke implies an increase in political posts of 69.9% (i.e., exp(0.530) – 1 = 0.699). In contrast, the effect sizes range between 11.5% (Grüne – Green Party) and 24.4% (SPD – Social Democrats) for posts pertaining to the other parties. Cross-model Wald tests indicate that the size of the effect in the Linke model is significantly larger (at the 1% level) than for all other parties.

The mean number of posts related to the different parties prior to the algorithm updates – shown at the bottom of Table 5 – indicates that the Linke was underrepresented on Facebook at that time. That is, the newspapers on average published 3.5 posts per month using expressions related to the Linke, whereas they published between 5.9 (Grüne) and 7.6 (CDU/CSU) posts related to either of the other four parties. Hence, the increase in posts pertaining to the Linke after the second algorithm reduced the degree of underrepresentation of that party and led to a more even distribution of political posts across parties, which explains part of the lower concentration of viewpoints that we find in Table 3.

Figure A12 shows the development of the Gini-style measure of concentration when we exclude each party at a time. This measure remains quite similar when we remove any party other the Linke but excluding this party results in a much flatter Gini curve. These findings suggest that the changes in outlets' editorial policies may have been mostly related to a relative gain in posts pertaining to the underrepresented Linke.

To verify that the increased relative representation of the Linke drove the changes unexplained by the law of large numbers, we compare the leave-out Gini scores of the empirical and the simulated samples. Table 6 reports specifications similar to Column (3) of Table 4 but excluding each party at a time. Table 6 shows that excluding all parties except Linke increases the unexplained drop in the Gini score. In contrast, the specification that excludes the Linke (Column 4) documents no unexplained decrease in concentration. Here the simulated Gini score fully explains the change in the empirical Gini score after the algorithm update, confirming that the increased relative representation of the Linke drove that part of the effect of the algorithm change unexplained by the law of large numbers.

Are the low levels of Facebook coverage of the Linke plausible? In many situations, social media tend to promote right-wing content more than left-content content (e.g., Schradie, 2019; Reuning et al. 2022). In our context, it seems reasonable that relatively more posts pertain to CDU/CSU and SPD because these are Germany's large catch-all parties. However, it is not immediately obvious why the number of posts related to the Linke is much lower than in the case of FDP and Grüne, since the three parties enjoy similar levels of popularity among voters.[21] Thus, from a demand-side perspective, one would expect similar degrees of coverage of these smaller parties. However, there is a supply-side explanation for the underrepresentation of the Linke. According to survey data, German journalists tend to identify themselves much more often with Grüne or FDP than Linke (Lünenborg and Berghofer, 2010). When joining Facebook at the beginning of the 2010s, there was a lot of uncertainty among German newspapers about how to use the platform (Cision, 2013), which could have enabled a situation where the personal preferences of journalists and social media editors affect the political selection of postings. The recommendations provided by Facebook when announcing the August 2014 algorithm update may have reduced this kind of supply-side influence by helping the newspapers to "educate" their staff about the importance of catering to consumer preferences, leading to a better representation of the Linke and a generally more balanced mix of postings.

4.2.3 User engagement with news story posts vs. other content

An important premise of our analysis is that Facebook's algorithm updates decreased competition for news stories from other content on the platform and raised newspapers' returns to posting news content. We cannot evaluate whether and to what degree the algorithm updates affected referrals to outlets' websites, due to the lack of data. However, it is possible to analyze user engagement with posts. As previous research shows, likes, shares, and comments of news story posts often induce page visits for the publishing outlet (Mahmood and Sismeiro, 2017; Sismeiro and Mahmood, 2018). To evaluate whether the algorithm updates increased the returns to posting news stories on Facebook, we compare user engagement with the posts of the newspapers in our sample with the posts by the official Facebook Pages of political parties. We believe that the posts of political parties are an ideal comparison group because their content also relates to political topics. However, the Facebook Pages of political parties should not have been directly affected by the algorithm updates, as these Pages do not classify as news publishers.

We present corresponding difference-in-differences estimates in Table 7. In Columns (1) to (3), we regress the monthly sum of likes, shares, and comments, respectively, over all posts that do not include any political expressions on the interaction between the $After$ dummies and an indicator that takes the value 1 for Pages of newspapers and 0 for Pages of political parties. In contrast, Columns (4) to (6) pertain to user engagement with posts that do include political expressions. The results consistently indicate that posts of newspapers received more engagement after the algorithm updates than posts of political parties. This result holds after

[21] The vote shares of the Linke in the 2013 and 2017 national elections, respectively, were 8.6% and 9.2%; of the FDP 4.8% and 10.7%; and of the Grüne 8.4% and 8.9%.

including the monthly total number of posts, which suggests that the observed differences in user engagement were not driven by changes in the amount of content. Importantly, after the second algorithm update, the increase in user engagement was larger for political posts (Columns 4 to 6) than unpolitical posts (Column 1 to 3). Hence, the estimates support our assumption that the new version of Facebook's news feed algorithm increased the returns to posting news content relative to non-news content, especially in the case of substantive political news. Event-study plots shown in Figure A14 indicate that the boost in likes and shares faded towards the end of the sample period, whereas the increase in comments persisted.

4.3 Political knowledge, preferences, and participation

We use survey data from a rolling cross-section of Internet users in Germany (cp. Section 3.5) to estimate whether measures of political knowledge, preferences, and participation z of Facebook users changed after the algorithm updates, relative to non-users:

$$z_{r,w} = \alpha_1 + \alpha_2 FB\_user_{r,w} + \alpha_3 After_w + \alpha_4 FB\_user_{r,w} \times After_w + a_5 X_{r,w} + \varepsilon_{r,w} \qquad (8)$$

where r indices survey respondents interviewed in wave w, with w = 1 denoting interviews before and w = 2 after the updates. The binary variable FB_user takes the value 1 for respondents who indicate to have used Facebook at least once in the week before their interview (N = 318), and 0 otherwise (N = 273). $After$ equals 1 for respondents interviewed in November and December 2014 (N = 325), and 0 for responses collected in September and October 2012 (N = 266). Hence, the interaction between FB_user and $After$ captures the relative change in the outcome variables for Facebook users from the time before to after the algorithm updates. The vector of control variables X includes respondents' gender, age, nationality, education, employment status, income, party identification, separate variables for the degree of TV, print, and online news consumption, as well as day-of-the-interview fixed effects.

Results are summarized in Table 8. In Columns (1) to (3), we evaluate respondents' knowledge, measured in terms of shares of "do not know" answers in question sets about political candidates, parties, and coalitions, respectively (see Table A11 for details). As the coefficient on the interaction between FB_user and $After$ shows, we observe reductions in "do not know" answers of Facebook users in all three categories. The magnitude of the reductions amounts to 0.019 / 0.073 = 0.260 standard deviations in the case of candidate-related knowledge, 0.073 / 0.173 = 0.422 standard deviations of missing party-specific knowledge, and 0.076 / 0.151 = 0.503 standard deviations of missing knowledge about Germany's political coalitions. These estimates suggest that the political knowledge of Facebook users improved more for more complex issues (i.e., coalitions) than less sophisticated knowledge (i.e., candidates). As Column (4) suggests, Facebook users also indicate to have an increased interest in politics when comparing the interviews before and after the algorithm updates. The magnitude of this change equals 0.279 / 0.803 = 0.347 standard deviations on the interest scale. In Column (5), the estimates indicate that Facebook users were relatively more likely to state that they voted in the 2013 than 2009 federal elections, which is suggestive of an increased turnout among this group. The magnitude of this change corresponds to 0.096 / 0.298 = 0.322 standard deviations. We also find a positive sign on the interaction between FB_user and $After$ when we evaluate respondents' participation in the most recent state elections, but this coefficient is not significant at conventional levels.

The survey data used in Table 8 have limitations that do not allow for a causal interpretation of the results. First, due to the rolling cross-section design of the surveys, we cannot track changes of individuals but rely on comparisons at the group level. Since the use of social media became more universal over time (Newman et al., 2016), the self-selection of individuals into the Facebook user group is different in the pre- and the post-algorithm change periods. To mitigate this selection bias, we control for a large set of observable characteristics (e.g., age and education), factors that might be important for the selection into social media usage. Second, our survey data have a lower time frequency than our content data – namely just one observation before and one after the algorithm updates – which prevents us from credibly distinguishing the effect of the algorithm updates from a secular increasing trend of political coverage on Facebook. Nevertheless, the results suggest that whatever the reason for an increased coverage, it may have affected readers' knowledge and political participation. Third, survey responses tend to be biased (e.g., because of social desirability), as the discrepancy between respondents' reported turnout (e.g., 90.2%; see bottom of Column 7) and actual turnout (around 70%) demonstrates. Fourth, interactions between social media, offline media consumption, and face-to-face conversations about political issues are complex and difficult to capture by self-reported Facebook usage in the week before the interview. Despite these limitations, the estimates offer circumstantial evidence that Facebook users' political knowledge, interest, and turnout increased after the algorithm updates.

As Table A12 indicates, we do not find any indications of changes in respondents' party identification. It could be argued that survey respondents should be more likely to identify with the Linke after the algorithm updates, given the increase in postings related to that party. An explanation for unchanged party identification patterns is that postings related to the Linke may not specifically refer to that party but left-leaning topics in general (e.g., those shared with other left-leaning parties in Germany). Another explanation would be that changes in party identification require changes in somebody's core political and ideological values. The increase in postings related to the Linke may not have been large enough to induce measurable effects on these values, compared to variables such as political knowledge and interest – which are arguably easier to manipulate.

5. Conclusion

In December 2013 and again in August 2014, Facebook announced to take action against the proliferation of low-quality and non-informative content on the platform by having its algorithm select high-quality content of news publishers more often in users' feeds. We investigate the conjecture that the algorithm updates raised the incentives for media companies to publish news stories about substantive political issues on the platform, using a sample 37 German newspapers. To avoid problems related to omitted variables, we compare the newspapers' provision of news stories on Facebook with their print editions, assuming that the latter were not affected by Facebook's interventions.

Difference-in-differences estimates do not indicate an effect of the algorithm updates on the overall number of postings, likely because the newspapers found it more beneficial to increase their provision of news story posts selectively rather than indiscriminately. In fact, we find that the newspapers expanded the number of political stories, especially after the second algorithm update. Facebook's first algorithm update was not necessarily inconsequential, but the effects may have been delayed and/or become visible only in combination with the second update. Compared to the print editions, the number of political posts increased by about 30%. For the most part, this expansion occurred in a politically balanced way. That is, the newspapers randomly increased their coverage of the core topics and ideas pertaining to different political parties. For mechanical reasons, this expansion led to a lower within-outlet concentration of political viewpoints – as predicted by the law of large numbers – by approximately one half of the standard deviation of our concentration indices. Simulations show that the distribution of political posts across parties converged to the more balanced distribution of political articles in the print versions, simply due to larger daily "samples" of Facebook postings. However, we also find a change in newspapers' editorial policies, implying that the number of posts pertaining to the core topics and ideas of the Linke (Left Party) increased more than posts related to the other parties. This party has been underrepresented on Facebook, especially prior to the algorithm updates, but the disproportionate increase in these posts led to a more even distribution across parties. We find that the law-of-large-numbers effect explains about three quarters of the overall decrease in concentration, while the better representation of the Linke accounts for the remaining quarter.

These results are not without limitations. Our analyses are based on data from Germany. It is advisable to be careful to generalize the results to countries with different political systems and other forms of organization in the news industry. For example, Facebook traffic accounts for a relatively small share of all website traffic of German media outlets (Newman et al., 2016). It is possible that the algorithm updates investigated in this paper had larger effects in countries where social media are more important for revenues than in Germany. In addition, our results only relate to political aspects of news quality. It would be useful to investigate how the algorithm updates affected other dimensions of quality – such as the publication of sensationalist content, the number of employed social media editors, or journalistic awards – but we do not have access to those data. Finally, while our measure of concentration of viewpoints captures the coverage of different political parties within a given newspaper (i.e., internal diversity), our data do not allow us to assess the distribution of viewpoints across newspapers within a user's news feed (i.e., external diversity). Hence, it remains unclear how the decrease in within-outlet concentration that we find affected the consumption of political viewpoints. Future research is necessary to evaluate to which degree the news feed algorithm exposes users to sources that offer different vs. similar political views.

Our results have important implications for media companies, despite these limitations. First, our findings show that the news industry acts according to the incentives created by Facebook. Given the atomistic competition for user attention online, ignoring the rules set by the platform would make it difficult to reach certain audiences. Thus, concerns about filter bubbles and belief polarization need to be first and foremost addressed to the platform – rather than content creators – because of its power over how news stories are selected and consumed.

Second, the algorithm updates investigated in this paper are a good example that the financial interests of social platforms are not necessarily incompatible with societal needs for civic dialogue and the information of voters. Facebook changed its algorithm for profit reasons, with the intention to attract more users and increase engagement levels on the platform. However, Facebook's push of quality news story posts was potentially also helpful from an information perspective. It induced news outlets to provide more news stories that are likely useful for voter knowledge about substantive policy issues, which may facilitate political accountability and collective decision-making. Based on the complementary survey data we analyze, we cautiously confirm this conjecture, as Facebook users appear to be relatively better informed, more interested in politics, and more likely to vote after the algorithm updates.

Third, our results are relevant to the debate about renumerating publishers for the content they post on Facebook. In February 2021, Australia passed legislation that forces large platforms to negotiate agreements with publishers about compensations for the provision of news stories (i.e., the News Media Bargaining Code). Legislators in other places, such as Canada, the EU, and the UK have indicated to move forward with similar initiatives. While this kind of regulation could level the playing field between powerful platforms and news publishers, there are concerns that large media conglomerates will be able to negotiate satisfactory agreements, while Facebook might not be interested to enter agreements with small publishers, as their content might not generate enough user engagement (e.g., Kaye, 2021; Khalil, 2021). Even though the Australian legislation stipulates compulsory arbitration in case Facebook refuses to negotiate, it is conceivable that this process leads to a reduction in news quality on the platform. If Facebook is forced to pay publishers, it will likely be very selective about what kind of content to purchase, with a focus on content that drives user engagement. Future research is necessary to evaluate whether the News Media Bargaining Code and similar legislation lead to an overall decrease in news content published on Facebook, as well as less diversity, both within and across outlets.

References

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The Welfare Effects of Social Media. *American Economic Review*, 110, 629–676.

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31, 211–236.

Angelucci, C., & Cagé, J. (2019). Newspapers in Times of Low Advertising Revenues. *American Economic Journal: Microeconomics*, 11, 319–364.

Azzimonti, M., & Fernandes, M. (2018). Social Media Networks, Fake News, and Polarization. Working Paper.

Bachmann, P., Eisenegger, M., & Ingenhoff, D. (2021). Defining and Measuring News Me-dia Quality: Comparing the Content Perspective and the Audience Perspective. *The International Journal of Press/Politics*, forthcoming.

Barabas, J., & Jerit, J. (2009). Estimating the Causal Effects of Media Coverage on Policy-Specific Knowledge. *American Journal of Political Science*, 53, 73–89.

Berry, S., & Waldfogel, J. (2010). Product Quality and Market Size. *Journal of Industrial Economics*, 58, 1–31.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-Million-Person Experiment in Social Influence and Political Mobilization. *Nature*, 489, 295–298.

Bursztyn, L., Egorov, G., Enikolopov, R., & Petrova, M. (2019). Social Media and Xenophobia: Evidence from Russia. NBER Working Paper 26567.

Cagé, J. (2020). Media Competition, Information Provision and Political Participation: Evidence from French Local Newspapers and Elections, 1944–2014. *Journal of Public Economics*, 185.

Cagé, J., Hervé, N., & Mayozer, B. (2020). Social Media and Newsroom Production Decisions. Working Paper.

Cavusoglu, H., Phan, T. Q., Cavusoglu, H., & Airoldi, E. M. (2016). Assessing the Impact of Granular Privacy Controls on Content Sharing and Disclosure on Facebook. *Information Systems Research*, 27, 848–879.

Chiou, L., & Tucker, C. (2018). Fake News and Advertising on Social Media: A Study of the Anti-Vaccination Movement. NBER Working Paper 25223.

Cision (2013). "Social Journalism" Studie 2012/2013 – Deutschland. Cision Germany White Paper.

DeVito, M. A. (2017). From Editors to Algorithms – A Values-Based Approach to Understanding Story Selection in the Facebook News Feed. *Digital Journalism*, 5, 753–773.

Djourelova, M., Durante, R., & Martin, G. (2021). The Impact of Online Competition on Local Newspapers: Evidence from the Introduction of Craigslist. Working Paper.

Dor, D. (2003). On Newspaper Headlines as Relevance Optimizers. *Journal of Pragmatics*, 35, 695–721.

Dujeancourt, E., & Garz, M. (2021). The Effects of Algorithmic Content Selection on User Engagement with News on Twitter. Working Paper.

Dunaway, J. (2008). Markets, Ownership, and the Quality of Campaign News Coverage. *Journal of Politics*, 70, 1193–1202.

Facebook (2013). Helping You Find More News to Talk About. Press release, December 2, 2013. URL: https://newsroom.fb.com/news/2013/12/news-feed-fyi-helping-you-find-more-news-to-talk-about/.

Facebook (2014). Click-baiting. Press release, August 25, 2014. URL: https://newsroom.fb.com/news/2014/08/news-feed-fyi-click-baiting/.

Facebook (2015). Showing Fewer Hoaxes. Press release, January 20, 2015. URL: https://about.fb.com/news/2015/01/news-feed-fyi-showing-fewer-hoaxes/.

Fujiwara, T., Müller, K., & Schwarz, C. (2021). The Effect of Social Media on Elections: Evidence from the United States. NBER Working Paper 28849.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social Clicks: What and Who Gets Read on Twitter? ACM SIGMETRICS / IFIP Performance 2016, Antibes Juan-les-Pins, France.

Gans, J. S., & Leigh, A. (2012). How Partisan is the Press? Multiple Measures of Media Slant. *Economic Record*, 88, 127–147.

Garz, M. (2018). Retirement, Consumption of Political Information, and Political Knowledge. *European Journal of Political Economy*, 53, 109–119.

Garz, M., Sörensen, J., & Stone, D. F. (2020). Partisan Selective Engagement: Evidence from Facebook. *Journal of Economic Behavior & Organization*, 177, 91–108.

Gentzkow, M., Kelly, B. T., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, 57, 535–574.

Gentzkow, M., & Shapiro, J. M. (2010). What Drives Media Slant? Evidence from U.S. Daily Newspapers. *Econometrica*, 78, 35–71.

Gilens, M. (2001). Political Ignorance and Collective Policy Preferences. *American Political Science Review*, 95, 379–396.

Groseclose, T., & Milyo, J. (2005). A Measure of Media Bias. *Quarterly Journal of Economics*, 120, 1191–1237.

Guess, A. M., Nyhan, B. & Reifler, J. (2020). Exposure to Untrustworthy Websites in the 2016 US Election. *Nature Human Behaviour*, 4, 472–480.

Hanitzsch, T., Hanusch, F., Mellado, C., et al. (2011). Mapping Journalism Cultures across Nations: A Comparative Study of 18 Countries. *Journalism Studies*, 12, 273–293.

Hong, S. (2012). Online News on Twitter: Newspapers' Social Media Adoption and Their Online Readership. *Information Economics and Policy*, 24, 69–74.

Huang, N., Hong, Y., & Burtch, G. (2017). Social Network Integration and User Content Generation: Evidence from Natural Experiments. *MIS Quarterly*, 41, 1035–1058.

Humbrecht, E., & Büchel, F. (2013). More of the Same or Marketplace of Opinions? A Cross-National Comparison of Diversity in Online News Reporting. *The International Journal of Press/Politics*, 18, 436–461.

Jones, J. J., Bond, R. M., Bakshy, E., Eckles, D., & Fowler, J. H. (2017). Social Influence and Political Mobilization: Further Evidence from a Randomized Experiment in the 2012 U.S. Presidential Election. *PLoS ONE*, 12, 1–9.

Jurafsky, D., & Martin, J. H. (2008). *Speech and Language Processing*. 2nd edition, Prentice Hall.

Kafka, P. (2013). Like This if You Like Pandas! Facebook Says Publishers Shouldn't Fret About News Feed Changes. All Things Digital, December 6, 2013. URL: http://allthingsd.com/20131206/like-this-if-you-like-pandas-facebook-says-publishers-shouldnt-fret-about-news-feed-changes/.

Kaye, B. (2021). Facebook Rejects Talks with Australia Publisher, May Test Online Law. Reuters, June 25, 2021.

KEK (2015). *Fünfter Medienkonzentrationsbericht*. Kommission zur Ermittlung der Konzentration im Medienbereich. Vistas, Berlin.

Kerkhof, A. (2019). Advertising and Content Differentiation: Evidence from YouTube. Working Paper.

Khalil, S. (2021). Facebook to Pay News Corp for Content in Australia. BBC News, March 16, 2021.

Lischka, J., & Garz, M. (2021). Clickbait News and Algorithmic Curation: A Game Theory Framework of the Relation Between Journalism, Users, and Platforms. *New Media & Society*, forthcoming.

Lünenborg, M., & Berghofer, S. (2010). *Politikjournalistinnen und -journalisten: Aktuelle Befunde zu Merkmalen und Einstellungen vor dem Hintergrund ökonomischer und technologischer Wandlungsprozesse im deutschen Journalismus.* Deutscher Fachjournalisten-Verband, Berlin.

Mahmood, A., & Sismeiro, C. (2017). Will They Come and Will They Stay? Online Social Networks and News Consumption on External Websites. *Journal of Interactive Marketing*, 37, 117–132.

Martin, G. J., & Yurukoglu, A. (2017). Bias in Cable News: Persuasion and Polarization. *American Economic Review*, 107, 2565–2599.

Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. *American Economic Review*, 104, 2421–2455.

McDonald, D. G., & Dimmick, J. (2003). The Conceptualization and Measurement of Diversity. *Communication Research*, 30, 60–79.

Mosquera, R., Odunowo, M., McNamara, T., Guo, X., & Petrie, R. (2020). The Economic Effects of Facebook. *Experimental Economics*, 23, 575–602.

Mousavi, R., & Zhao, K. (2018). Examining the Impacts of Airbnb's Review Policy Change on Listing Reviews. SSRN Working Paper 3206895.

Müller, K., & Schwarz, C. (2020a). Fanning the Flames of Hate: Social Media and Hate Crime. *Journal of the European Economic Association*, 19, 2131–2167.

Müller, K., & Schwarz, C. (2020b). From Hashtag to Hate Crime: Twitter and Anti-Minority Sentiment. Working Paper.

Myllylahti, M. (2020). Paying Attention to Attention: A Conceptual Framework for Studying News Reader Revenue Models Related to Platforms. *Digital Journalism*, 8, 567–575.

Newman, N., Fletcher, R., Levy, D. A. L., & Nielsen, R. K. (2016). Reuters Institute Digital News Report 2016. Reuters Institute for the Study of Journalism.

Pariser, E. (2011). *The Filter Bubble: What the Internet is Hiding from You*. Penguin, London.

Peterson-Salahuddin, C., & Diakopoulos, N. (2020). Negotiated Autonomy: The Role of Social Media Algorithms in Editorial Decision Making. *Media and Communication*, 8, 27–38.

Pew (2013). The Role of News on Facebook. Pew Research Center.

Puglisi, R., & Snyder Jr., J. M. (2015). The Balanced US Press. *Journal of the European Economic Association*, 13, 240–264.

Qin, B., Strömberg, D., & Wu, Y. (2018). Media Bias in China. *American Economic Review*, 108, 2442–2476.

Rattinger, H., Roßteutscher, S., Schmitt-Beck, R., Weßels, B., Wolf, C., Bieber, I., & Scherer, P. (2014). ZA5351 Langfrist-Online-Tracking, T18 (GLES). GESIS Datenarchiv, Cologne.

Rattinger, H., Roßteutscher, S., Schmitt-Beck, R., Weßels, B., Wolf, C., Henckel, S., Bieber, I., & Scherer, P. (2015): ZA5726 Langfrist-Online-Tracking T26 (GLES). GESIS Datenarchiv, Cologne.

Reuning, K., Whitesell, A., & Hannah, A. L. (2022). Facebook Algorithm Changes May Have Amplified Local Republican Parties. *Research & Politics*, 9(2).

Rotesi, T. (2019). The Impact of Twitter on Political Participation. Working Paper.

Schradie, J. (2019). *The Revolution That Wasn't: How Digital Activism Favors Conservatives*. Harvard University Press, Cambridge, MA.

Simonov, A., & Rao, J. (2018). Demand for Online News under Government Control: Evidence from Russia. *Journal of Political Economy*, forthcoming.

Sismeiro, C., & Mahmood, A. (2018). Competitive vs. Complementary Effects in Online Social Networks and News Consumption: A Natural Experiment. *Management Science*, 64, 5014–5037.

Stecula, D. A., & Pickup, M. (2021). Social Media, Cognitive Reflection, and Conspiracy Beliefs. *Frontiers in Political Science*, 3, 1–8.

Szeidl, A., & Szucs, F. (2021). Media Capture through Favor Exchange. *Econometrica*, 89, 281–310.

Sun, M., & Zhu, F. (2013). Ad Revenue and Content Commercialization: Evidence from Blogs. *Management Science*, 59, 2314–2331.

Sunstein, C. R. (2017). *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press, Princeton, NJ.

Vega, T. (2013). Facebook Says It Failed to Bar Posts With Hate Speech. New York Times. May 28, 2013.

Wellbrock, C. M. (2011). Die journalistische Qualität deutscher Tageszeitungen – Ein Ranking. *Medienwirtschaft*, 8, 22–31.

Wortelker, J. (2021). User Preferences and Editorial Content Selection on Facebook. Working Paper.

Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2020). Political Effects of the Internet and Social Media. *Annual Review of Economics*, 12, 415–438.

Tables and figures

Table 1: Effects of the algorithm updates on the provision of news content

	(1) Log number of news items (OLS)	(2) Number of news items (negative binomial)
Facebook × $After^{Nov2013}$	-0.009	0.005
	(0.062)	(0.054)
Facebook × $After^{Jul2014}$	0.041	0.078
	(0.084)	(0.070)
Time fixed effects	Yes	Yes
Outlet × channel fixed effects	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes
Adj. R^2	0.988	
Pseudo R^2		0.261
Observations	3965	3965

Notes: The regressions use data at the outlet-channel-month level. The column headers denote the dependent variable and estimation method. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table 2: Effects of the algorithm updates on the provision of political news

	(1) Log number of political news items (OLS)	(2) Number of political news items (negative binomial)	(3) Log number of political news items (OLS)	(4) Number of political news items (negative binomial)
Facebook × $After^{Nov2013}$	0.075	0.071	0.081	0.036
	(0.107)	(0.082)	(0.091)	(0.053)
Facebook × $After^{Jul2014}$	0.228**	0.265***	0.196***	0.185***
	(0.097)	(0.093)	(0.056)	(0.050)
Log total number of news items			0.939***	0.894***
			(0.049)	(0.055)
Time fixed effects	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes
Adj. R^2	0.986		0.990	
Pseudo R^2		0.310		0.353
Observations	3903	3903	3903	3903

Notes: The regressions use data at the outlet-channel-month level. The column headers denote the dependent variable and estimation method. Political news items are Facebook posts or print articles that include relevant expressions from parties' election manifestos. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table 3: Effects of the algorithm updates on the concentration of political news

	(1) Gini	(2) RSD	(4) Gini	(5) RSD
Facebook × $After^{Nov2013}$	-0.083	-0.006	-0.082	-0.006
	(0.232)	(0.058)	(0.233)	(0.058)
Facebook × $After^{Jul2014}$	-0.557***	-0.114**	-0.558***	-0.114**
	(0.182)	(0.049)	(0.182)	(0.049)
Total number of items			0.000	0.000
			(0.000)	(0.000)
Time fixed effects	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes
Adj. R^2	0.474	0.435	0.473	0.434
Observations	3965	3965	3965	3965
Standard deviation of dependent variable on Facebook	1.103	0.265	1.103	0.265

Notes: OLS estimates. The regressions use data at the outlet-channel-month level. RSD denotes concentration in terms of the relative standard deviation of political news. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table 4: Comparison of empirical and simulated samples

	(1)	(2) Gini score	(3)
Simulated × $After^{Jul2014}$	0.131**	0.107**	0.199***
	(0.048)	(0.047)	(0.073)
Simulated	-0.030		
	(0.042)		
$After^{Jul2014}$	-0.471***		
	(0.061)		
Outlet × sample fixed effects	No	Yes	Yes
Time effects	No	Yes	Yes
Sample specific linear time trend	No	No	Yes
Observations	3844	3844	3844

Notes: OLS estimates using data at the outlet-month level. The column header denotes the dependent variable. Bootstrapped standard errors are in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 5: Effects of the algorithm updates on the provision of political news, by party

	(1)	(2)	(3)	(4)	(5)
Panel A: OLS	Log number of posts/articles pertaining to ...				
	… CDU/CSU	… FDP	… Grüne	… Linke	… SPD
Facebook × $After^{Nov2013}$	0.039	0.124	0.110	0.109	0.053
	(0.106)	(0.111)	(0.104)	(0.085)	(0.121)
Facebook × $After^{Jul2014}$	0.169^{***}	0.132^{*}	0.109	0.530^{***}	0.218^{**}
	(0.063)	(0.072)	(0.067)	(0.072)	(0.100)
Log total number of news items	0.923^{***}	0.898^{***}	0.857^{***}	0.745^{***}	0.915^{***}
	(0.050)	(0.056)	(0.067)	(0.058)	(0.059)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.986	0.985	0.988	0.988	0.985
Observations	3965	3927	3965	3965	3945

Panel B: Negative binomial	Number of posts/articles pertaining to ...				
	… CDU/CSU	… FDP	… Grüne	… Linke	… SPD
Facebook × $After^{Nov2013}$	-0.052	0.063	0.034	0.084	-0.045
	(0.066)	(0.078)	(0.067)	(0.102)	(0.082)
Facebook × $After^{Jul2014}$	0.133^{**}	0.123^{*}	0.086	0.632^{***}	0.168^{***}
	(0.057)	(0.070)	(0.059)	(0.066)	(0.064)
Log total number of news items	0.872^{***}	0.855^{***}	0.814^{***}	0.771^{***}	0.917^{***}
	(0.095)	(0.106)	(0.103)	(0.082)	(0.084)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes	Yes
Outlet × channel specific trend polynomial	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.340	0.344	0.360	0.404	0.327
Observations	3965	3927	3965	3965	3945
Mean number of Facebook posts before Dec 2013	7.600	6.526	5.863	3.456	6.258

Notes: The regressions use data at the outlet-channel-month level. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table 6: Comparison of empirical and simulated samples, leaving out party identities

	(1)	(2)	(3)	(4)	(5)
			Gini score		
Left out ideology	CDU	FDP	Grüne	Linke	SPD
Simulated × After^{Jul2014}	0.242***	0.189***	0.179***	-0.009	0.214***
	(0.056)	(0.060)	(0.060)	(0.051)	(0.057)
Outlet × sample fixed effects	Yes	Yes	Yes	Yes	Yes
Time effects	Yes	Yes	Yes	Yes	Yes
Sample specific linear time trend	Yes	Yes	Yes	Yes	Yes
Observations	3844	3844	3844	3844	3844

Notes: OLS estimates using data at the outlet-month level. The column header denotes the dependent variable. Bootstrapped standard errors are in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 7: Effects of the algorithm updates on user engagement with Pages of newspapers vs. Pages of political parties

	Unpolitical posts			Political posts		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS	Log likes	Log Shares	Log comments	Log likes	Log Shares	Log comments
Newspaper × $\text{After}^{\text{Nov2013}}$	0.47^{***}	0.50^{**}	0.66^{***}	0.46^{**}	0.65^{**}	0.71^{***}
	(0.15)	(0.20)	(0.16)	(0.22)	(0.27)	(0.21)
Newspaper × $\text{After}^{\text{Jul2014}}$	0.13	0.22	0.51^{**}	0.30^{*}	0.48^{**}	0.73^{***}
	(0.13)	(0.19)	(0.19)	(0.16)	(0.23)	(0.21)
Log total number of posts	0.74^{***}	0.68^{***}	0.83^{***}	0.91^{***}	0.86^{***}	0.88^{***}
	(0.07)	(0.06)	(0.06)	(0.12)	(0.11)	(0.13)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Page fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Page-specific trend polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.959	0.933	0.972	0.892	0.853	0.900
Observations	2237	2237	2237	2237	2237	2237

Panel B: Negative binomial	Likes	Shares	Comments	Likes	Shares	Comments
Newspaper × $\text{After}^{\text{Nov2013}}$	0.60^{***}	0.66^{**}	0.70^{***}	0.63^{***}	0.73^{**}	0.75^{***}
	(0.17)	(0.27)	(0.17)	(0.22)	(0.32)	(0.18)
Newspaper × $\text{After}^{\text{Jul2014}}$	0.16	0.18	0.52^{***}	0.33^{**}	0.42^{*}	0.76^{***}
	(0.13)	(0.21)	(0.19)	(0.16)	(0.23)	(0.21)
Log total number of posts	0.75^{***}	0.66^{***}	0.84^{***}	0.82^{***}	0.77^{***}	0.83^{***}
	(0.07)	(0.06)	(0.05)	(0.09)	(0.09)	(0.10)
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Page fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Page-specific trend polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.148	0.141	0.190	0.131	0.129	0.153
Observations	2237	2237	2237	2237	2237	2237

Notes: The regressions use monthly data from the Facebook Pages of 37 newspapers and 5 political parties (CDU, FDP, Grüne, Linke, and SPD). The column headers denote the dependent variable. In Columns (1) to (3), the engagement metrics refer to the monthly sum of likes, shares, and comments over posts that do not contain any political expressions. In Columns (4) to (6), these metrics refer to posts that do include political expressions. Standard errors (in parentheses) are clustered by Page.

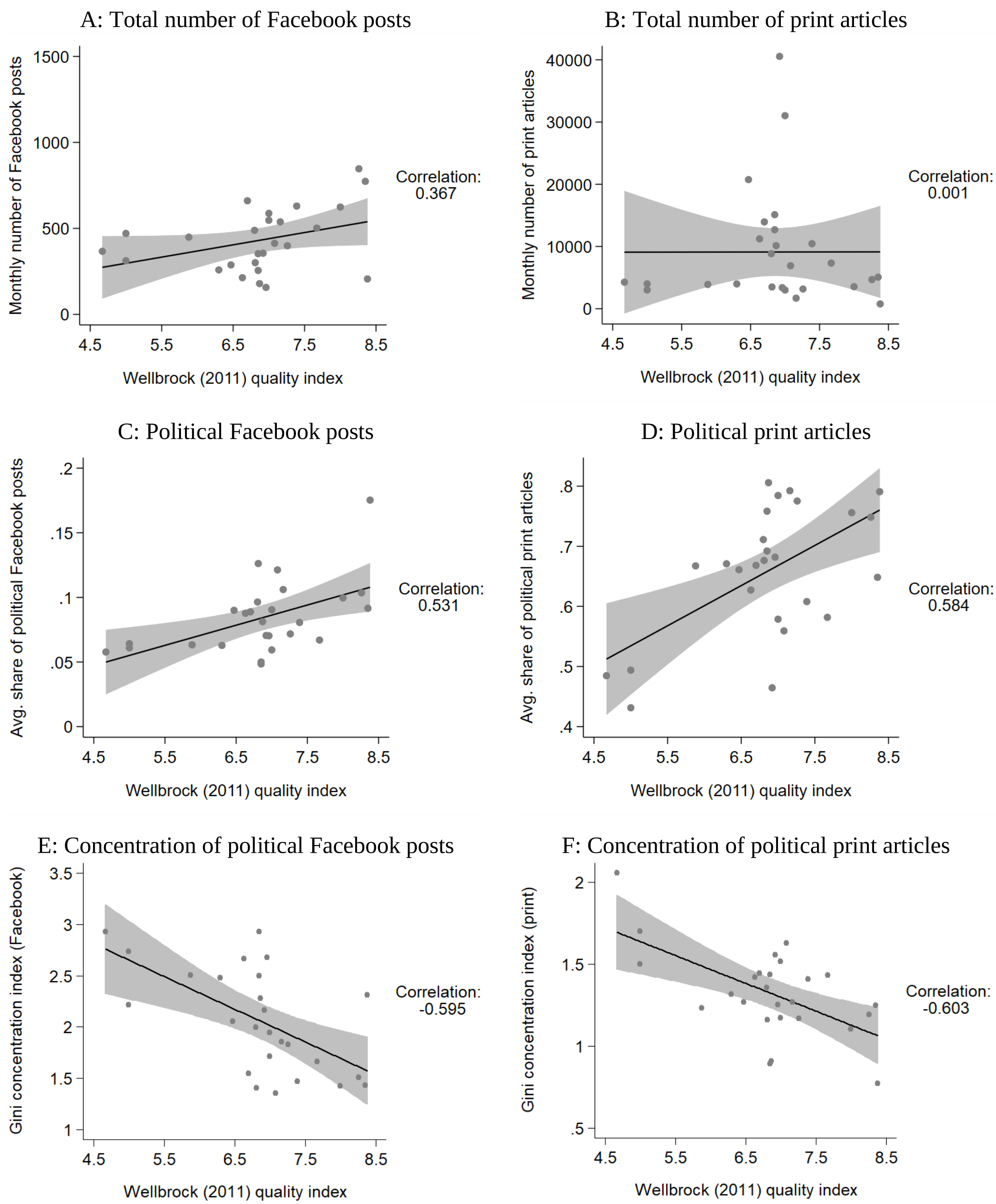
* p<0.10, ** p<0.05, *** p<0.01

Table 8: Relative changes in political knowledge, interest, and self-report turnout of Facebook users

	(1) Share of missing candidate knowledge	(2) Share of missing party knowledge	(3) Share of missing coalition knowledge	(4) Interest in politics (scale 1 to 5)	(5) Voted in federal election (yes/no)	(6) Voted in state election (yes/no)
Facebook user × After	-0.019*	-0.073**	-0.076***	0.279**	0.096**	0.059
	(0.011)	(0.030)	(0.024)	(0.123)	(0.046)	(0.060)
Facebook user	0.021***	0.049*	0.060**	-0.286***	-0.056*	-0.039
	(0.008)	(0.029)	(0.024)	(0.095)	(0.032)	(0.048)
After	0.005	-0.028	-0.013	-0.202**	-0.047	0.038
	(0.007)	(0.018)	(0.016)	(0.096)	(0.034)	(0.046)
Dependent variable: Mean	0.013	0.036	0.027	3.714	0.902	0.832
Dependent variable: SD	0.073	0.173	0.151	0.803	0.298	0.374
R^2	0.149	0.176	0.181	0.224	0.155	0.174

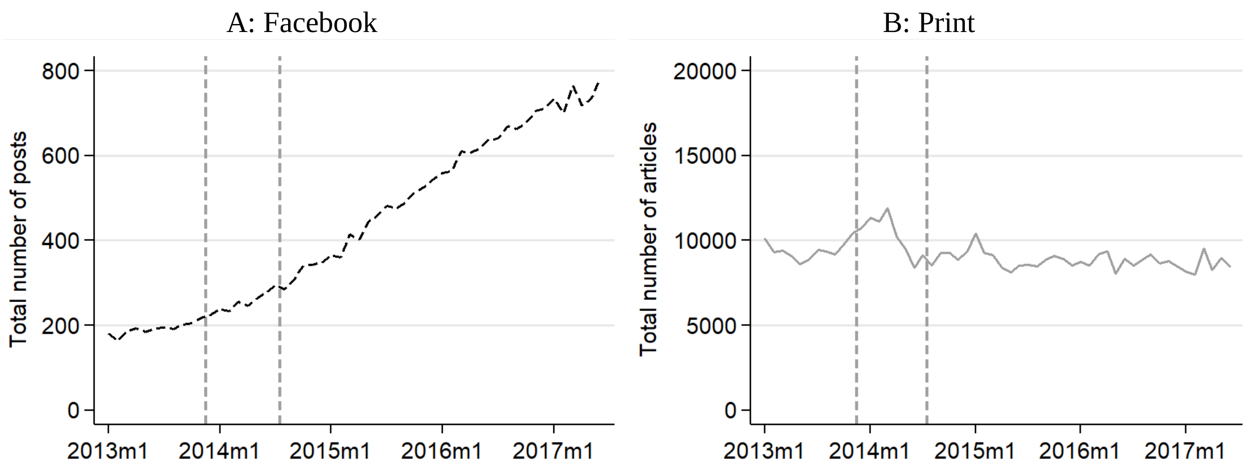
Notes: OLS estimates, using data from N = 591 survey respondents. All models include controls for gender, age, nationality, education, employment status, income, party identification, degree of TV news consumption, degree of print news consumption, degree of online news consumption, day-of-the-interview fixed effects, and a constant. Robust standard errors are in parentheses.
* p < 0.10, ** p < 0.05, *** p < 0.01

Figure 1: Supply of news stories and survey-based measures of news quality



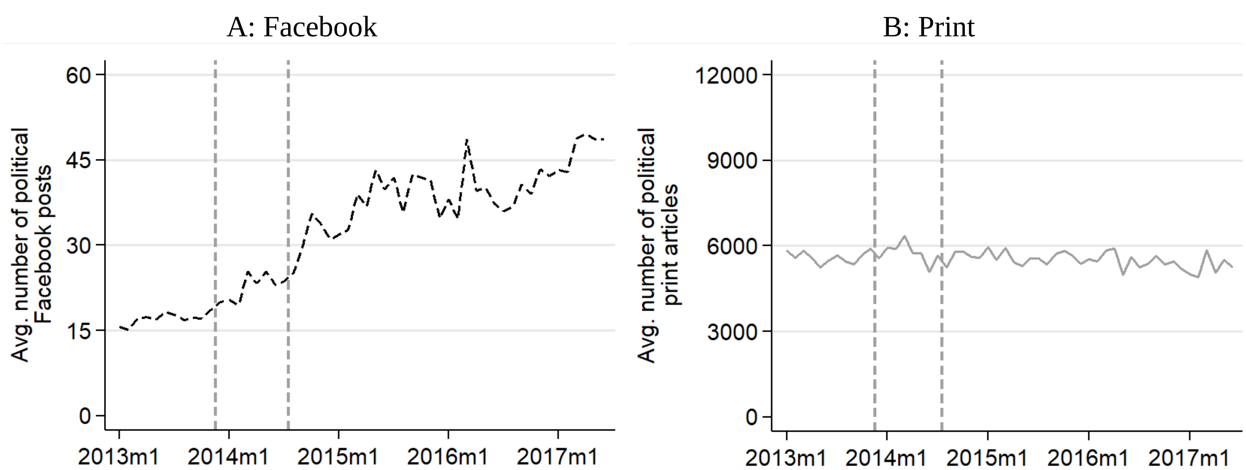
Notes: Political news items are Facebook posts or print articles that include relevant expressions from parties' election manifestos. The black solid line shows the linear fit and the shaded area denotes corresponding 95% confidence interval.

Figure 2: Overall news quantities over time



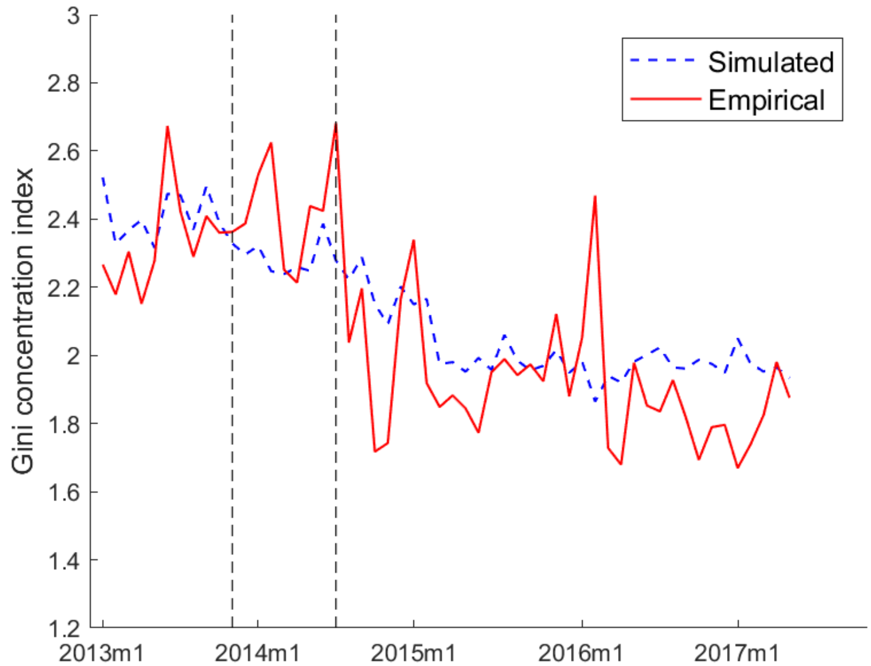
Notes: The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure 3: Absolute occurrence of political news items



Notes: Political news items are Facebook posts or print articles that include relevant expressions from parties' election manifestos. The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure 4: Ideological concentration of empirical and simulated samples over time



Notes: Lower concentration scores reflect more balanced coverage, whereas higher scores indicate more one-sided reporting. The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Online Appendix

In Section 4.2.1, we conduct a simulation to explore a possible mechanism behind the decrease in ideological concentration after the algorithm change in August 2014. The simulation investigates the hypothesis that the change was driven by the increase of postings with ideologically relevant expressions, which activated the law of large numbers and resulted in a more balanced coverage. The simulated bias of outlet i at time t is given by

$$d_{i,t,k}^{Gini} = \sum_{k=1}^{K} \frac{\sum_{p=1}^{n} \sum_{q=1}^{n} \left| x_{p,i,t,k} - x_{q,i,t,k} \right|}{\sum_{p=1}^{n} x_{p,i,t,k}}$$

where $x_{p,i,t,k} = \frac{simitems_{p,i,t,k}^{match}}{simitems_{p,i,t,k}^{total}}$ is the share of articles in outlet i at time t assigned to the ideology of party p by iteration k. In each iteration, we randomly assign an ideology to each article covering politics using the empirical distribution of ideology combinations. We calculate the nonparametric probability of all ideology combinations $r \in I^5$ (where I is the set of ideologies) for all outlets separately:

$$\hat{P}_{i,r} = \frac{\sum_{t<20} items_{r,i,t}^{match}}{\sum_{t<20} politems_{i,t}^{total}}$$

where the algorithm change happened at $t = 20$.

Table A1: List of outlets

Outlet	Facebook domain	Page likes	Circulation	Owner
Abendblatt Hamburg	abendblatt	105,294	170,579	Funke Mediengruppe
Berliner Morgenpost	morgenpost	222,188	76,798	Funke Mediengruppe
Berliner Zeitung	berlinerzeitung	182,601	95,189	DuMont Mediengruppe
B.Z. Berlin	B.Z.Berlin	116,138	99,975	Axel Springer
Express	EXPRESS.Koeln	213,365	86,061	DuMont Mediengruppe
Frankfurter Allgemeine	faz	494,080	240,551	Frankfurter Allgemeine
Focus	focus.de	710,333	438,055	Hubert Burda Media
Frankfurter Rundschau	FrankfurterRundschau	86,480	176,393	Ippen
Freie Presse	freiepresse	98,810	229,163	Chemnitzer Verlag/Druck
General-Anzeiger	gaonline	55,144	66,363	Rheinische Post Gruppe
Handelsblatt	handelsblatt	218,853	126,107	DvH Medien
Kölner Stadt-Anzeiger	ksta.fb	128,168	263,035	DuMont Mediengruppe
Lausitzer Rundschau	lausitzerrundschau	20,761	73,254	Neue Pressegesellschaft
Leipziger Volkszeitung	lvzonline	90,762	173,345	Madsack Mediengruppe
Main-Post	mainpost	38,759	115,815	Mediengruppe Pressedruck
Märkische Allgemeine	MAZonline	36,489	106,296	Madsack Mediengruppe
Mitteldeutsche Zeitung	mzwebde	97,607	170,729	DuMont Mediengruppe
MOPO	hamburgermorgenpost	150,097	71,313	DuMont Mediengruppe
Neue Westfälische	NeueWestfaelische	53,288	211,928	SPD-Medienholding
Nürnberger Zeitung	nordbayern.de	33,189	245,897	Verlag Nürnberger Presse
Nordwest-Zeitung	nwzonline	51,261	112,040	Nordwest Medien
Neue Osnabrücker	neueoz	78,946	63,647	Neue Osnabrücker Zeitung
Ostthüringer Zeitung	otz.de	33,136	242,634	Funke Mediengruppe
Passauer Neue Presse	pnp.de	88,167	160,229	Verlagsgruppe Passau
Rheinische Post	rponline	136,312	291,473	Rheinische Post Gruppe
Schwäbische Zeitung	schwaebische.de	46,645	165,373	Schwäbisch Media
Der Spiegel	DerSpiegel	438,871	765,178	Spiegel-Verlag
Süddeutsche Zeitung	ihre.sz	698,695	358,402	Südwestdt. Medien Holding
Südkurier	Suedkurier.News	28,781	117,762	Mediengruppe Pressedruck
Südwest Presse	swp.de	40,007	266,173	Neue Pressegesellschaft
Der Tagesspiegel	Tagesspiegel	134,908	109,938	DvH Medien
Die Tageszeitung	taz.kommune	271,001	50,986	Taz Verlagsgenossenschaft
Thüringische LZ	tlz.de	22,711	242,634	Funke Mediengruppe
Thüringer Allgemeine	thueringerallgemeine	63,298	242,634	Funke Mediengruppe
Die Welt	welt	933,445	171,433	Axel Springer
Die Welt Kompakt	weltkompakt	72,053		Axel Springer
Die Zeit	diezeit	428,543	498,439	DvH Medien

Notes: Page likes refer to June 2017. The circulation refers to the second quarter in 2017 and is based on data from the German audit bureau of circulation (IVW).

Table A2: Summary statistics of outlet-month data

	Mean	SD
Panel A: Facebook		
Total number of posts	435.45	359.27
Number of political posts	32.49	25.72
Concentration indices		
-Gini	2.06	1.10
-relative standard deviation (RSD)	0.44	0.26
Panel B: Print		
Total number of articles	9153.66	9653.93
Number of political articles	5559.82	4719.46
Concentration indices		
-Gini	1.32	0.47
-relative standard deviation (RSD)	0.30	0.11

Notes: N = 3,965 (2 channels, 37 outlets, and up to 54 months).

Table A3: Regression of regional publication of political news on regional voting

	(1) Facebook	(2) Facebook	(3) Print	(4) Print
Vote share 2013	0.016***		0.108***	
	(0.005)		(0.040)	
Vote share 2017		0.025***		0.136**
		(0.007)		(0.055)
Outlet fixed effects	Yes	Yes	Yes	Yes
Ideology fixed effects	Yes	Yes	Yes	Yes
Adj. R^2	0.830	0.836	0.961	0.961

Notes: OLS estimates, using 135 observations (27 outlets, 5 parties). Dependent variable: party-specific share of news items (Facebook posts or print articles) containing political expressions. The vote shares refer to the 2013 and 2017 national election results in the electoral districts where the outlets have their main area of circulation. Robust standard errors in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table A4: Effects of the algorithm updates on provision of political news (using article headlines only)

	(1) Log number of political news items (OLS)	(2) Number of political news items (negative binomial)	(3) Log number of political news items (OLS)	(4) Number of political news items (negative binomial)
Facebook × After^{Nov2013}	0.028	0.053	0.034	0.018
	(0.098)	(0.086)	(0.076)	(0.057)
Facebook × After^{Jul2014}	0.213*	0.212**	0.176**	0.128**
	(0.113)	(0.103)	(0.072)	(0.063)
Log total number of news items			1.082***	1.049***
			(0.128)	(0.097)
Time fixed effects	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes
Adj. R^2	0.962		0.976	
Pseudo R^2		0.307		0.350
Observations	3903	3903	3903	3903

Notes: The regressions use data at the outlet-channel-month level. The column headers denote the dependent variable and estimation method. Political news items are Facebook posts or print headlines that include relevant expressions from parties' election manifestos. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table A5: Effects of the algorithm updates on the concentration of political news (headline-only versions of concentration indices)

	(1) Gini	(2) RSD	(4) Gini	(5) RSD
Facebook × After^{Nov2013}	0.088	0.027	0.092	0.028
	(0.218)	(0.052)	(0.217)	(0.052)
Facebook × After^{Jul2014}	-0.385**	-0.081*	-0.388**	-0.082*
	(0.177)	(0.048)	(0.177)	(0.048)
Total number of items			0.000	0.000
			(0.000)	(0.000)
Time fixed effects	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes
Adj. R^2	0.457	0.421	0.457	0.421
Observations	3965	3965	3965	3965

Notes: OLS estimates. The regressions use data at the outlet-channel-month level. The column headers denote the dependent variable (i.e., the specific measure of concentration when analyzing (sub-)headings only). RSD refers to the relative standard deviation of political news. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table A6: Effects of the algorithm updates on the provision of political news (based on party mentions)

	(1) Log number of political news items (OLS)	(2) Number of political news items (negative binomial)	(3) Log number of political news items (OLS)	(4) Number of political news items (negative binomial)
Facebook × $\text{After}^{\text{Nov2013}}$	0.021	-0.122	0.030	-0.197**
	(0.110)	(0.114)	(0.102)	(0.098)
Facebook × $\text{After}^{\text{Jul2014}}$	0.300**	0.314**	0.258**	0.225*
	(0.142)	(0.158)	(0.115)	(0.124)
Log total number of news items			1.020***	1.130***
			(0.131)	(0.098)
Time fixed effects	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes
Adj. R^2	0.916		0.929	
Pseudo R^2		0.275		0.293
Observations	3965	3965	3965	3965

Notes: The regressions use data at the outlet-channel-month level. The column headers denote the dependent variable and estimation method. Political news items are Facebook posts or print articles that include the name of a relevant party (CDU, CSU, FDP, Grüne, Linke, or SPD). Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table A7: Effects of the algorithm updates on the share of political new items

	(1) Fractional regression (Logit)	(2) Fractional regression (Probit)	(3) Beta regression
Facebook × $\text{After}^{\text{Nov2013}}$	0.025	0.012	0.047
	(0.048)	(0.026)	(0.045)
Facebook × $\text{After}^{\text{Jul2014}}$	0.123***	0.054***	0.128***
	(0.038)	(0.020)	(0.038)
Time fixed effects	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes
Pseudo R^2	0.314	0.314	
Observations	3903	3903	3903

Notes: The regressions use data at the outlet-channel-month level. The dependent variable is share of political news items of total news items. The column headers denote the estimation method. Political news items are Facebook posts or print articles that include relevant expressions from parties' election manifestos. Standard errors are in parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table A8: Effects of the algorithm updates on the treatment and the control group separately

	(1) Treatment group	(2) Control group	(3) Treatment group	(4) Control group
	Log number of political news items (OLS)		Number of political news items (negative binomial)	
$After^{Nov2013}$	0.020	-0.055	0.042	-0.023
	(0.091)	(0.058)	(0.079)	(0.053)
$After^{Jul2014}$	0.184^{*}	0.046	0.226^{***}	-0.026
	(0.093)	(0.028)	(0.091)	(0.019)
Outlet × channel fixed effects	Yes	Yes	Yes	Yes
Outlet × channel-specific trend polynomial	Yes	Yes	Yes	Yes
Adj. R^2	0.815	0.924		
Pseudo R^2			0.209	0.197
Observations	1967	1936	1967	1936

Notes: The regressions use data at the outlet-month level. The column headers denote sample (Treatment group – Facebook posts, Control group – print articles), dependent variable, and estimation method. Political news items are Facebook posts or print articles that include relevant expressions from parties' election manifestos. Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table A9: Link format and occurrence of political expressions

	Post includes political expression(s) (yes/no)		Number of political expressions	
	(1)	(2)	(3)	(4)
	OLS	Probit	OLS	Negative binomial
Link format (yes/no)	0.015^{***}	0.116^{***}	0.020^{***}	0.293^{***}
	(0.003)	(0.022)	(0.006)	(0.059)
Adj. R^2	0.038		0.026	
Pseudo R^2		0.058		0.045
Observations	856532	856532	856532	856532

Notes: Using data at the level of the individual Facebook post, the table shows regressions of the occurrence of political expressions in these posts on a binary variable that takes the value 1 for link format posts and 0 otherwise. The column headers denote the exact definition of the dependent variable and the estimation method. All models include a constant, outlet fixed effects, hour, day, month, and year effects, the number of characters of the post message, and a trend polynomial of order 3. Standard errors (in parentheses) are clustered by outlet.
* p<0.10, ** p<0.05, *** p<0.01

Table A10: Effects of the algorithm updates on the provision of political news, by outlets' ideology

	(1)	(2)	(3)	(4)	(5)
Panel A: OLS	Log number of posts/articles pertaining to ...				
	...CDU/CSU	...FDP	...Grüne	...Linke	...SPD
Dominant outlet ideology (reference category: Facebook × After^{Jul2014} × SPD)					
Facebook × After^{Jul2014} × CDU/CSU	-0.065	-0.008	0.166	-0.078	-0.002
	(0.128)	(0.165)	(0.128)	(0.210)	(0.419)
Facebook × After^{Jul2014} × FDP	-0.083	-0.265	0.075	-0.279	-0.222
	(0.208)	(0.186)	(0.200)	(0.229)	(0.155)
Facebook × After^{Jul2014} × Grüne	-0.170	-0.089	0.071	-0.202	-0.126
	(0.161)	(0.226)	(0.219)	(0.224)	(0.172)
Facebook × After^{Jul2014} × Linke	-0.085	-0.128	0.276	0.040	-0.099
	(0.167)	(0.176)	(0.183)	(0.190)	(0.193)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes	Yes
Trend polynomial	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.986	0.985	0.988	0.988	0.985
Observations	3965	3927	3965	3965	3945
Panel B: Negative binomial	Number of posts/articles pertaining to ...				
	...CDU/CSU	...FDP	...Grüne	...Linke	...SPD
Dominant outlet ideology (reference category: Facebook × After^{Jul2014} × SPD)					
Facebook × After^{Jul2014} × CDU/CSU	0.052	0.154	0.195	-0.026	0.093
	(0.167)	(0.174)	(0.146)	(0.175)	(0.200)
Facebook × After^{Jul2014} × FDP	-0.127	-0.251	0.076	-0.443**	-0.315*
	(0.214)	(0.196)	(0.197)	(0.198)	(0.179)
Facebook × After^{Jul2014} × Grüne	-0.147	-0.043	0.031	-0.324	-0.177
	(0.184)	(0.236)	(0.219)	(0.246)	(0.202)
Facebook × After^{Jul2014} × Linke	-0.197	-0.236	0.155	-0.189	-0.260*
	(0.138)	(0.172)	(0.157)	(0.182)	(0.154)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Outlet × channel fixed effects	Yes	Yes	Yes	Yes	Yes
Trend polynomial	Yes	Yes	Yes	Yes	Yes
Pseudo R^2	0.340	0.344	0.360	0.405	0.327
Observations	3965	3927	3965	3965	3945

Notes: The regressions use data at the outlet-channel-month level. An outlet's dominant ideology derives from the party whose language the outlet uses disproportionately often in its print version. We determine the dominant language based on the residuals of a regression of the number of political articles on outlet and party fixed effects, to account for differences between outlets' overall extent of political news, as well as differences across outlets in the propensity to use language associated with individual parties. The corresponding interaction terms are binary variables (e.g., the FDP dummy takes the value 1 if most political expressions used in the outlet's print version refer to that party). Standard errors (in parentheses) are clustered at the outlet-channel level.
* p<0.10, ** p<0.05, *** p<0.01

Table A11: Main survey measures and underlying questions

Variable	Survey item(s)
Facebook user (1 if yes, 0 if no)	“Have you used Facebook during the past week?”
Share of missing candidate knowledge	Based on the number of times respondents answered the following questions with “do not know”: – “What do think of Angela Merkel?” – “What do think of Horst Seehofer?” – “What do think of Sigmar Gabriel?” – “What do think of Cem Özdemir?” – “What do think of Gregor Gysi?”
Share of missing party knowledge	Based on the number of times respondents answered the following questions with “do not know”: – “How do you rate CDU on the left-right scale?” – “How do you rate CSU on the left-right scale?” – “How do you rate FDP on the left-right scale?” – “How do you rate Grüne on the left-right scale?” – “How do you rate Linke on the left-right scale?” – “How do you rate SPD on the left-right scale?”
Share of missing coalition knowledge	Based on the number of times respondents answered the following questions with “do not know”: – “How likely you believe CDU/CSU and SPD would enter a coalition?” – “How likely you believe CDU/CSU and FDP would enter a coalition?” – “How likely you believe CDU/CSU, FDP, and Grüne would enter a coalition?” – “How likely you believe CDU/CSU and Grüne would enter a coalition?” – “How likely you believe SPD and Grüne would enter a coalition?” – “How likely you believe SPD, FDP, and Grüne would enter a coalition?” – “How likely you believe SPD, Linke, and Grüne would enter a coalition?”
Interest in politics (Scale from 1 “not at all” to 5 “very much”)	“How much are you generally interested in politics?”
Participation in federal election (1 if yes, 0 if no)	– “Did you vote in the 2009 federal elections?” (18th wave) – “Did you vote in the 2013 federal elections?” (26th wave)
Participation in state election (1 if yes, 0 if no)	– “Did you vote in the most recent state elections?”
Party identification	“Which political party do you identify most with?”

Notes: The data come from the 18th and 26th waves of GLES component “Long-Term Online Tracking”.

Table A12: Relative changes in party identification of Facebook users

	(1) CDU/CSU	(2) FDP	(3) Grüne	(4) Linke	(5) SPD
Facebook user × After	-0.051	0.019	-0.070	0.008	0.021
	(0.077)	(0.032)	(0.047)	(0.043)	(0.076)
Facebook user	0.008	-0.019	0.090**	-0.021	0.042
	(0.060)	(0.029)	(0.038)	(0.032)	(0.058)
After	0.061	-0.031	0.045	0.032	-0.031
	(0.065)	(0.032)	(0.037)	(0.036)	(0.060)
Dependent variable: Mean	0.308	0.032	0.100	0.068	0.249
Dependent variable: SD	0.462	0.177	0.300	0.251	0.433
R^2	0.118	0.056	0.104	0.089	0.038

Notes: OLS estimates, using data from N = 591 survey respondents. The dependent variable equals 1 if the respondent identifies with the party listed in the column header and 0 if not. All models include controls for gender, age, nationality, education, employment status, income, party identification, degree of Internet usage, degree of TV news consumption, degree of print news consumption, day-of-the-interview fixed effects, and a constant. Robust standard errors are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A1: Examples of Facebook posts

a) Link format

Translation: Angela Merkel criticizes Vladimir Putin harshly.

b) Status update (greetings)

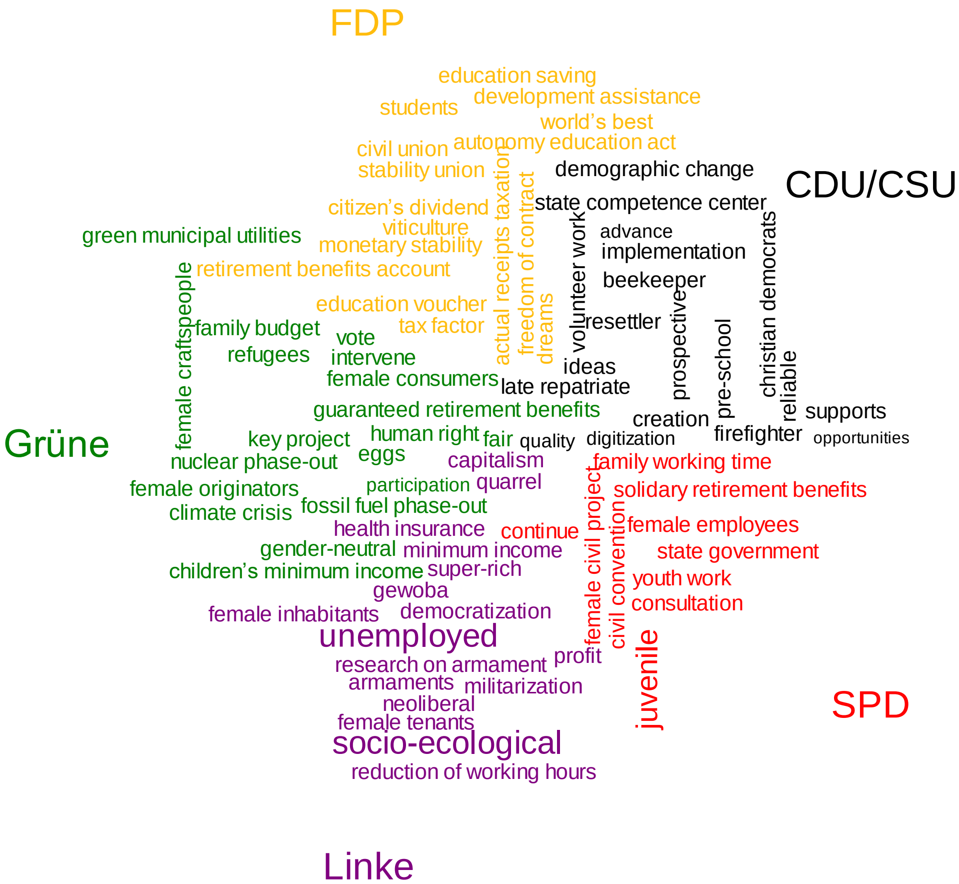
Translation: Good morning everybody!

c) Status update (live soccer scores)

Translation: Goal for Germany! Germany has the lead!

d) Link shared in text caption

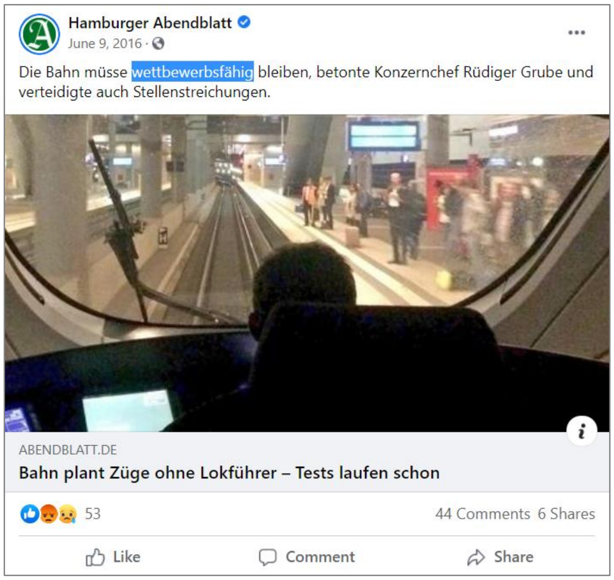
Figure A2: Expressions with the largest TF-IDF values (top 20) in parties' election programs



Notes: The figure shows English translations.

Figure A3: Examples of Facebook posts with and without political expressions

A: Expression pertaining to Conservative Party



Translation: Deutsche Bahn needs to stay <u>competitive</u>, CEO Rüdiger Grube emphasized while defending layoffs

B: Expression pertaining to Green Party

Translation: This <u>pesticide</u> is used on may fields in Germany

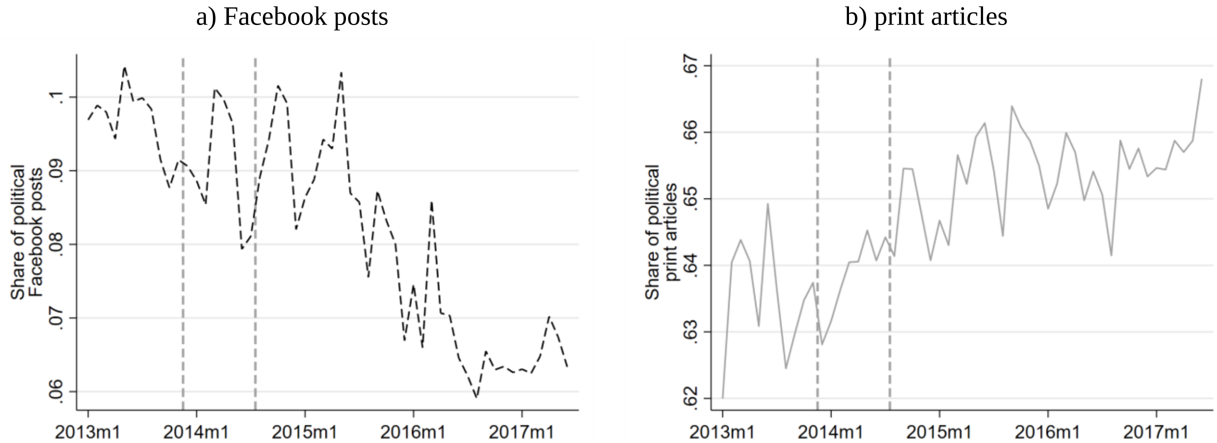
C: Expression pertaining to Social Democrats

Translation: Lousy <u>workers' conditions</u> at Vapiano

D: No political expression

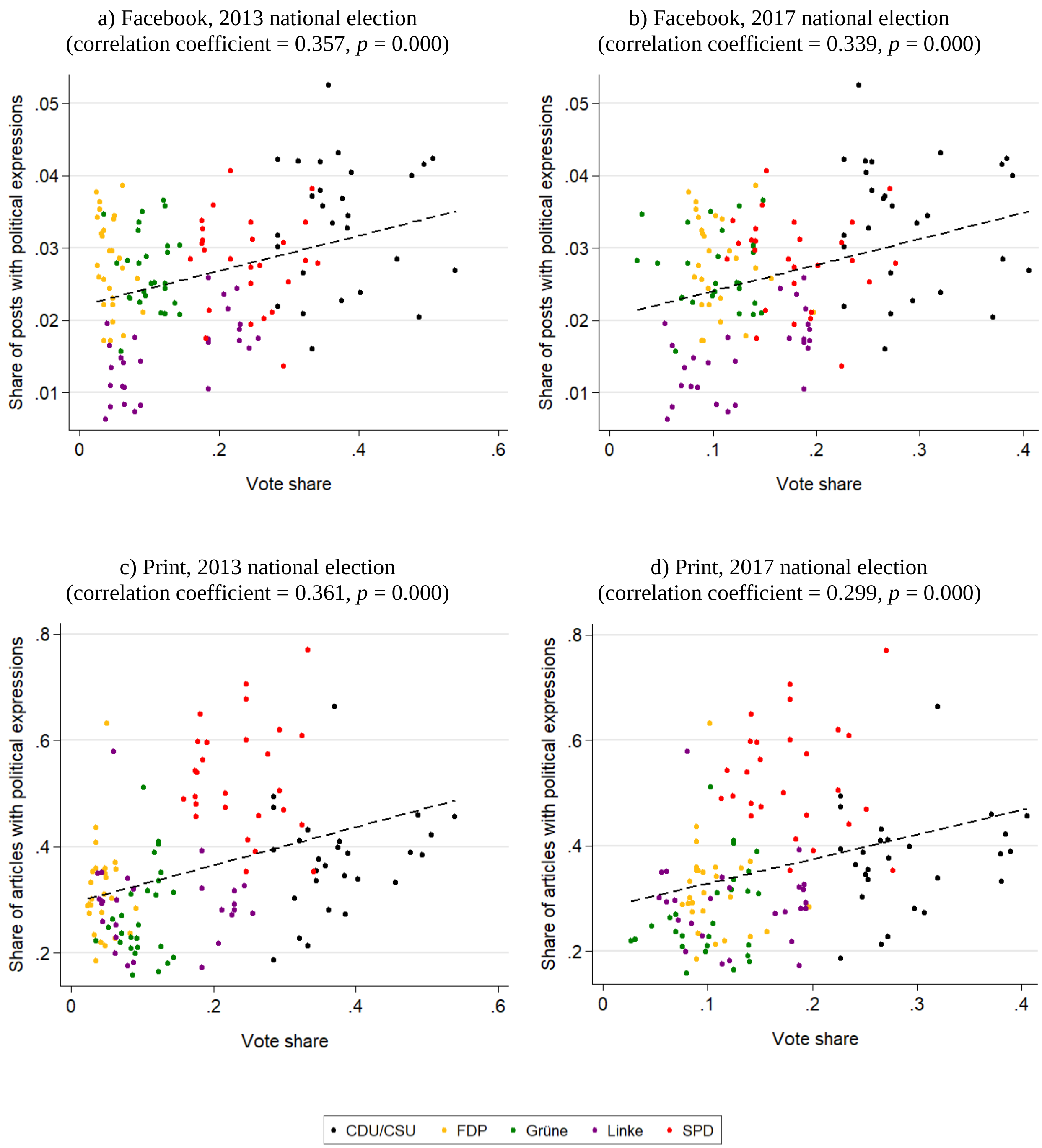
Translation: GQ's style experts took Trump to task and gave him a new look

Figure A4: Relative occurrence of political news items



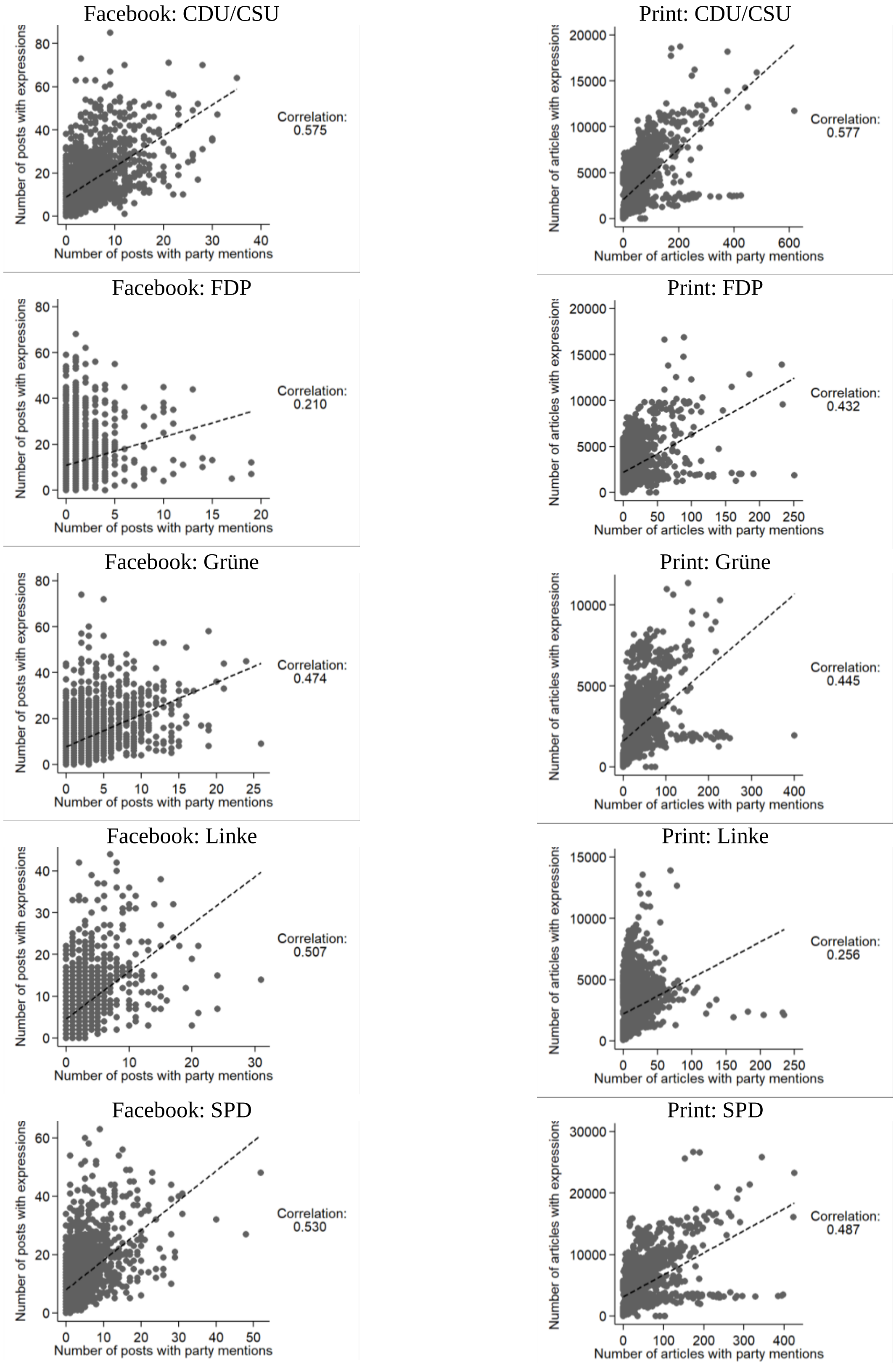
Notes: Political news items are Facebook posts or print articles that include relevant expressions from parties' election manifestos. The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure A5: Distribution of political news items across parties and regional voting



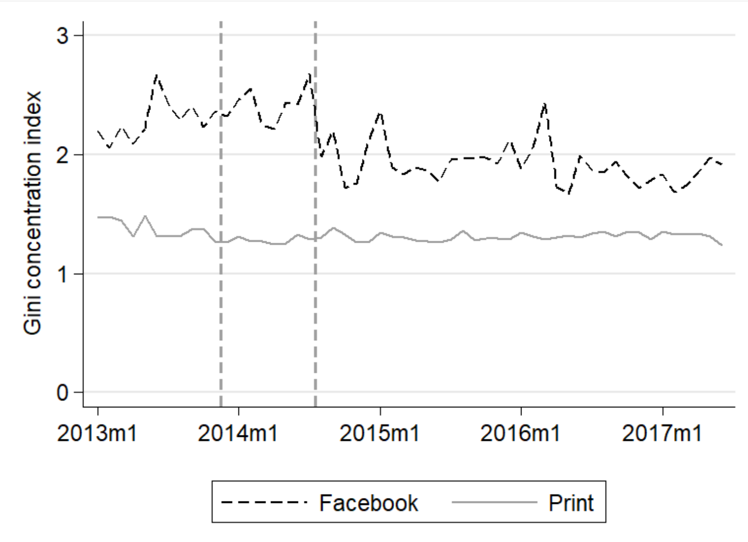
Notes: The figure compares the outlets' average shares of news related to different parties with the parties' national election vote shares in the electoral districts where the outlets have their main area of circulation. N = 135 (27 outlets, 5 parties). Each data point represents an outlet-party combination. The dashed lines show the linear fit.

Figure A6: Use of political expressions and party mentions



Notes: Posts with party mentions are those that include the search terms CDU or CSU, FDP, Grüne, Linke, or SPD. Each data point represents an outlet-channel-month combination. The dashed lines show the linear fit.

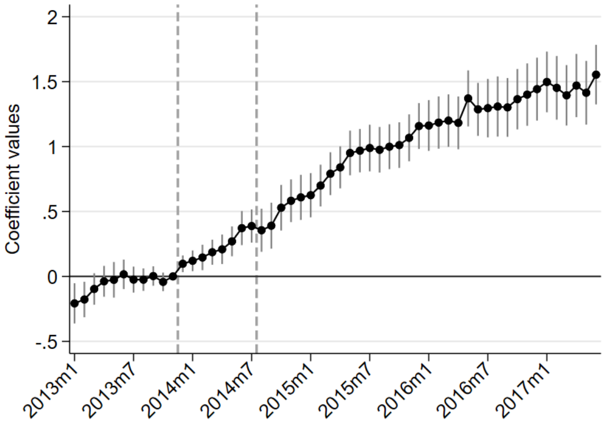
Figure A7: Concentration of political news



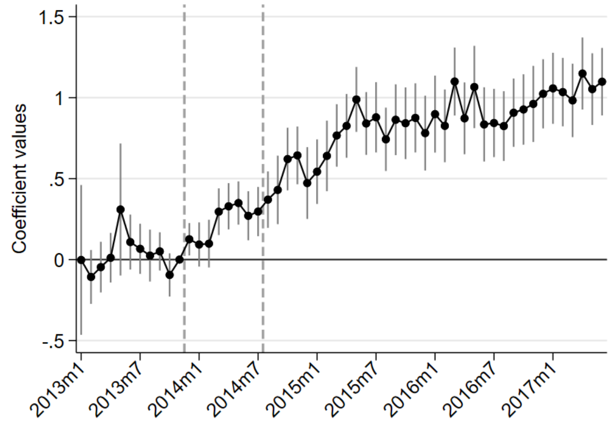
Notes: Lower concentration scores reflect more balanced coverage, whereas higher scores indicate more one-sided reporting. The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure A8: Tests for parallel pre-trends (main variables)

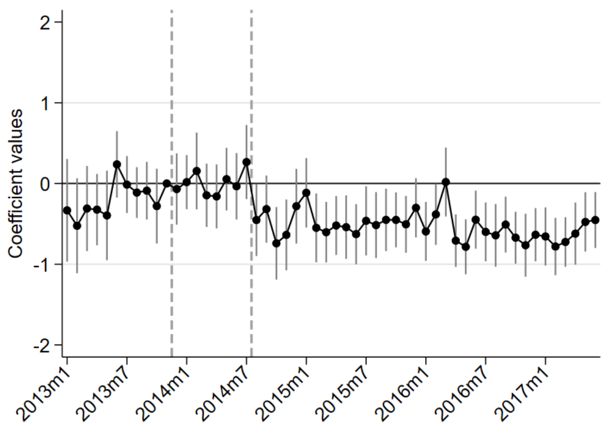
A: Log total number of Facebook posts / print articles



B: Log number of political posts and political articles

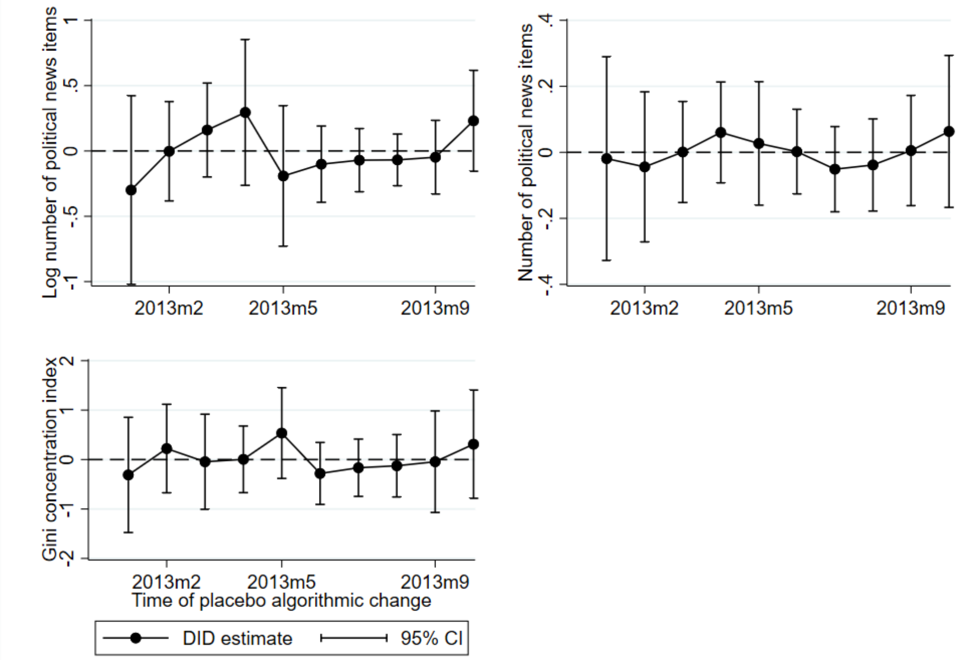


C: Gini concentration on Facebook and print



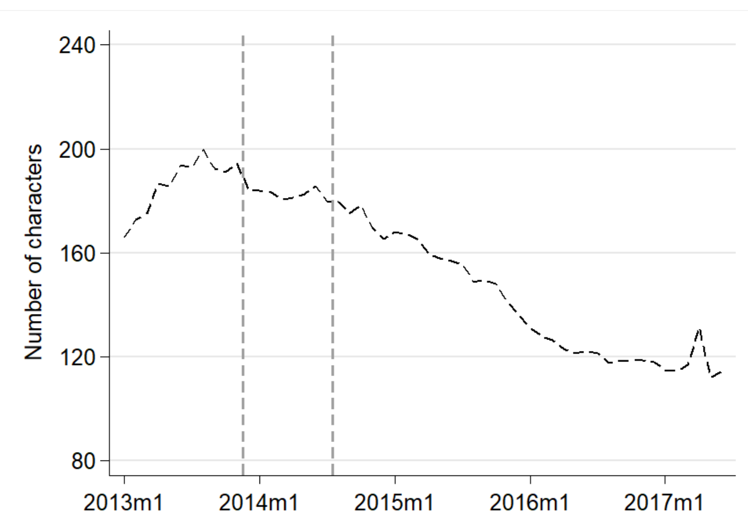
Notes: The graphs show the coefficients from OLS regressions of the variables stated above the graphs on interactions of the Facebook dummy with time dummies, conditional on outlet-channel fixed effects; see Equation (6) for details. The reference month is November 2013. The grey dashed lines mark the December 2013 and August 2014 updates of Facebook's news feed algorithm. The grey solid spikes denote the 90% confidence interval, based on standard errors clustered at the outlet-channel-level.

Figure A9: The effects of placebo algorithm updates in the pre-treatment period



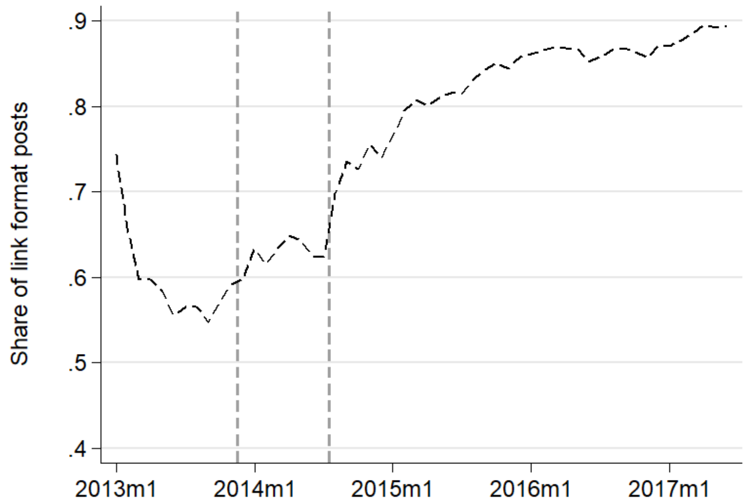
Notes: The graphs show coefficients of placebo algorithm updates from regressions focusing on the period before the first update (January 2013 – November 2013) and using specifications similar to Columns 3 (top left) and 4 (top right) of Table 2 and Column 1 of Table 3 (bottom). The horizontal axis denotes the time of the placebo algorithm update.

Figure A10: Average length of Facebook posts



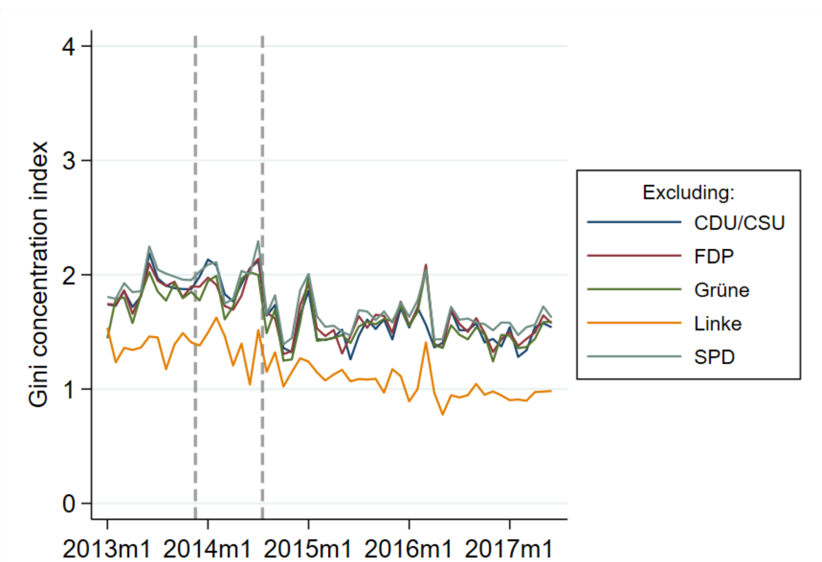
Notes: The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure A11: Use of link format over time



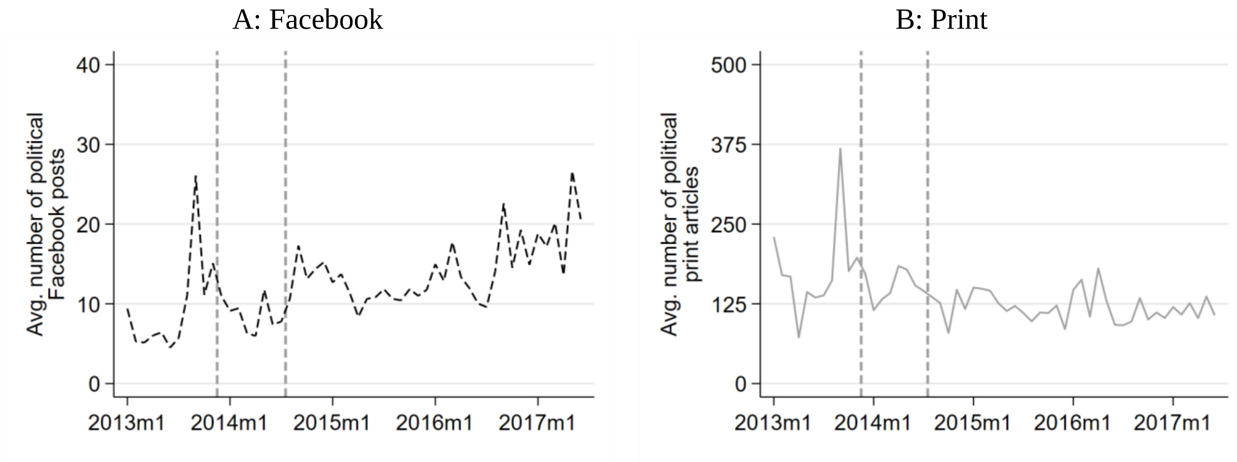
Notes: The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure A12: Concentration of political Facebook posts over time, excluding each party at a time



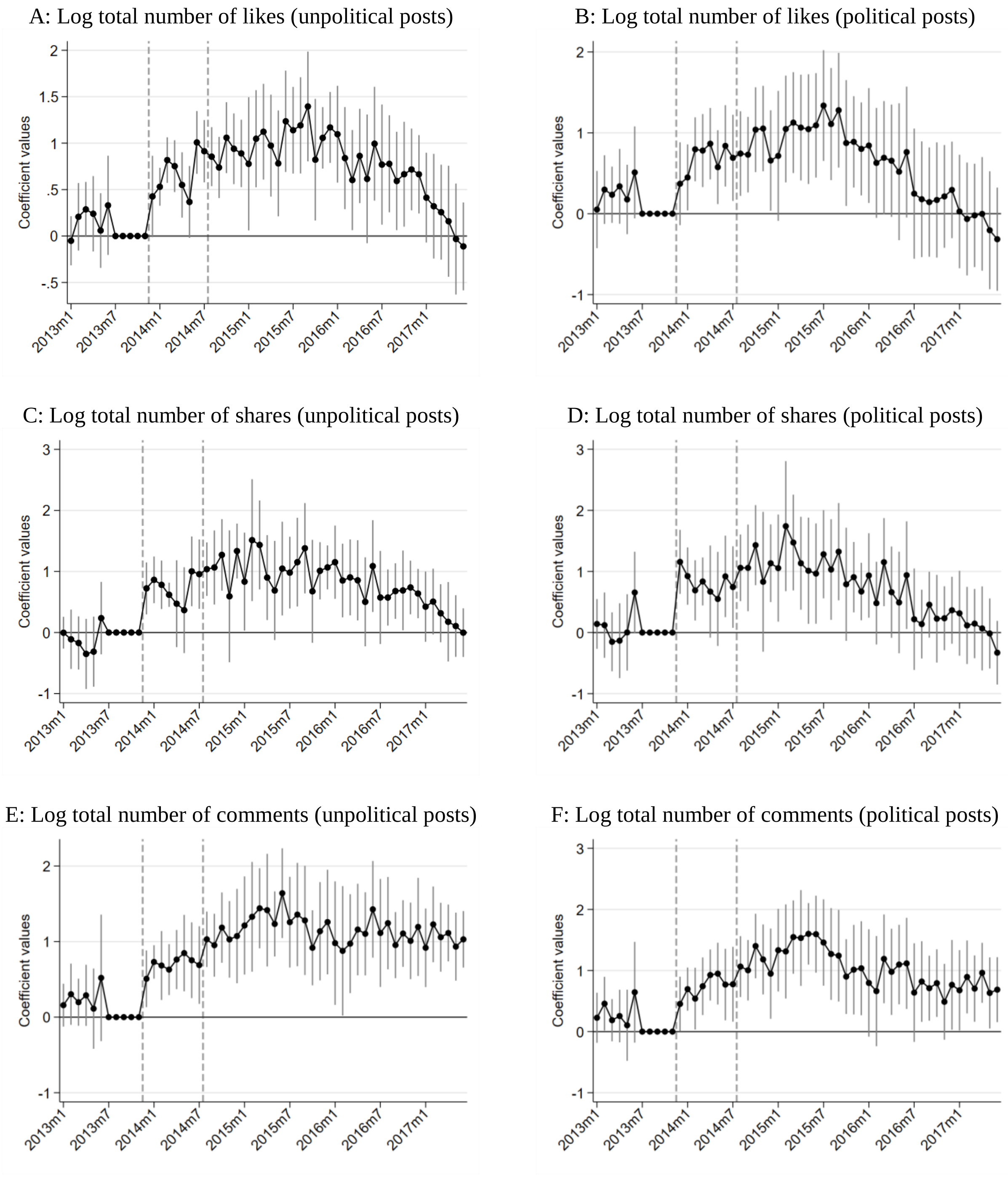
Notes: Lower concentration scores reflect more balanced coverage, whereas higher scores indicate more one-sided reporting. The grey dashed lines mark the December 2013 and August 2014 changes in Facebook's news feed algorithm.

Figure A13: Absolute occurrence of political news items (based on party mentions)



Notes: Political news items are Facebook posts or print articles that include mentions of the relevant political parties (i.e., CDU, CSU, FDP, Grüne, Linke, and SPD). The spike in September 2013 coincides with the elections to German parliament that year. The black solid line shows the linear fit and the shaded area denotes corresponding 95% confidence interval.

Figure A14: Tests for parallel pre-trends (engagement metrics)



Notes: The figure compares the monthly user engagement for the Facebook Pages of 37 newspapers relative to the Pages of 5 political parties (CDU, FDP, Grüne, Linke, and SPD), using the data described in Section 4.2.3 and Table 7. The coefficients are based on OLS regressions of the variables stated above the graphs on interactions of the Newspaper dummy with time dummies, conditional on Page fixed effects and the log total number of posts. The reference period (Jul 2013 – Nov 2013) stretches from the beginning of the campaign phase of the 2013 national elections to the last pre-treatment month. The grey dashed lines mark the December 2013 and August 2014 updates of Facebook's news feed algorithm. The grey solid spikes denote the 90% confidence interval, based on standard errors clustered by Page.