Algorithmic amplification of politics on Twitter

Ferenc Huszár[a,b,c,1,2] , Sofia Ira Ktena[a,1,3], Conor O'Brien[a,1] , Luca Belli[a,2] , Andrew Schlaikjer[a] , and Moritz Hardt[d]

[a]Machine Learning Ethics, Transparency, and Accountability Team, Twitter, San Francisco, CA 94103; [b]Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, United Kingdom; [c]Gatsby Computational Neuroscience Unit, University College London, London, W1T 4JG, United Kingdom; and [d]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720

Content on Twitter's home timeline is selected and ordered by personalization algorithms. By consistently ranking certain content higher, these algorithms may amplify some messages while reducing the visibility of others. There's been intense public and scholarly debate about the possibility that some political groups benefit more from algorithmic amplification than others. We provide quantitative evidence from a long-running, massive-scale randomized experiment on the Twitter platform that committed a randomized control group including nearly 2 million daily active accounts to a reverse-chronological content feed free of algorithmic personalization. We present two sets of findings. First, we studied tweets by elected legislators from major political parties in seven countries. Our results reveal a remarkably consistent trend: In six out of seven countries studied, the mainstream political right enjoys higher algorithmic amplification than the mainstream political left. Consistent with this overall trend, our second set of findings studying the US media landscape revealed that algorithmic amplification favors right-leaning news sources. We further looked at whether algorithms amplify far-left and far-right political groups more than moderate ones; contrary to prevailing public belief, we did not find evidence to support this hypothesis. We hope our findings will contribute to an evidence-based debate on the role personalization algorithms play in shaping political content consumption.

social media | algorithmic personalization | media amplification | political bias

Political content is a major part of the public conversation on Twitter. Politicians, political organizations, and news outlets engage large audiences on Twitter. At the same time, Twitter employs algorithms that learn from data to sort content on the platform. This interplay of algorithmic content curation and political discourse has been the subject of intense scholarly debate and public scrutiny (1–15). When first established as a service, Twitter used to present individuals with content from accounts they followed, arranged in a reverse chronological feed. In 2016 Twitter introduced machine learning algorithms to render tweets on this feed called Home timeline based on a personalized relevance model (16). Individuals would now see older tweets deemed relevant to them, as well as some tweets from accounts they did not directly follow.

Personalized ranking prioritizes some tweets over others on the basis of content features, social connectivity, and user activity. There is evidence that different political groups use Twitter differently to achieve political goals (17–20). What has remained a matter of debate, however, is whether or not any ranking advantage falls along established political contours, such as the left or right (2, 7), the center or the extremes (1, 3), specific parties (2, 7), or news sources of a certain political inclination (21). In this work, we provide systematic quantitative insights into this question based on a massive-scale randomized experiment on the Twitter platform.

Experimental Setup

Below, we outline this experimental setup and its inherent limitations. We then introduce a measure of algorithmic amplification in order to quantify the degree to which different political groups benefit from algorithmic personalization.

When Twitter introduced machine learning to personalize the Home timeline in 2016, it excluded a randomly chosen control group of 1% of all global Twitter users from the new personalized Home timeline. Individuals in this control group have never experienced personalized ranked timelines. Instead, their Home timeline continues to display tweets and retweets from accounts they follow in reverse chronological order. The treatment group corresponds to a sample of 4% of all other accounts who experience the personalized Home timeline. However, even individuals in the treatment group do have the option to opt-out of personalization (*SI Appendix*, section 1.A).

The experimental setup has some inherent limitations. A first limitation stems from interaction effects between individuals in the analysis (22). In social networks, the control group can never be isolated from indirect effects of personalization, as individuals in the control group encounter content shared by users in the treatment group. Therefore, although a randomized controlled experiment, our experiment does not satisfy the well-known Stable Unit Treatment Value Assumption from causal inference (23). As a consequence, it cannot provide unbiased estimates of causal quantities of interest, such as the average treatment

Significance

The role of social media in political discourse has been the topic of intense scholarly and public debate. Politicians and commentators from all sides allege that Twitter's algorithms amplify their opponents' voices, or silence theirs. Policy makers and researchers have thus called for increased transparency on how algorithms influence exposure to political content on the platform. Based on a massive-scale experiment involving millions of Twitter users, a fine-grained analysis of political parties in seven countries, and 6.2 million news articles shared in the United States, this study carries out the most comprehensive audit of an algorithmic recommender system and its effects on political content. Results unveil that the political right enjoys higher amplification compared to the political left.

effect. In this study, we chose to not employ intricate causal inference machinery that is often used to approximate causal quantities (24), as this would not guarantee unbiased estimates in the complex setting of Twitter's home timeline algorithm. Building an elaborate causal diagram of this complex system is well beyond the scope of our observational study. Instead, we present findings based on simple comparison of measurements between the treatment and control groups. Intuitively, we expect peer effects to decrease observable differences between the control and treatment groups; thus, our reported statistics likely underestimate the true causal effects of personalization.

A second limitation pertains to the fact that differences between treatment and control groups were previously used by Twitter to improve the personalized ranking experience. The treatment, that is, the ranking experience, has therefore not remained the same over time. Moreover, the changes to the treatment depend on the experiment itself.

Measuring Amplification

We define the reach of a set T of tweets in a set U of Twitter users as the total number of users from U who encountered a tweet from the set T.* Think of T, for example, as tweets from a group of politicians in Germany, and think of the audience U as all German Twitter users in the control group. We always consider reach within a specific time window, for example, a day.

We define the amplification ratio of set T of tweets in an audience U as the ratio of the reach of T in U intersected with the treatment group and the reach of T in U intersected with the control group. We normalize the ratio in such a way that amplification ratio 0% corresponds to equal proportional reach in treatment and control. In other words, a random user from U in the treatment group is just as likely to see a tweet in T as is a random user from U in the control group. An amplification ratio of 50% means that the treatment group is 50% more likely to encounter one of the tweets. Large amplification ratios indicate that the ranking model assigns higher relevance scores to the set of tweets, which therefore appear more often than they would in a reverse chronological ordering.

We often study the amplification ratio in cases where T is a set corresponding to tweets from a single Twitter account (individual amplification). When considering how groups of accounts are amplified, we have the choice between reporting distribution of amplification ratios of the individual accounts in the group or considering a single aggregate amplification ratio (group amplification), where T contains all tweets authored by any member of the group. We generally report both statistics. More detail on how we calculate amplification and a discussion of the difference between individual and group amplification is found in *SI Appendix*, section 1.D.

Results

We divide our findings into two parts. First, we study tweets by elected politicians from major political parties in seven countries which were highly represented on the platform. In the second analysis, which is specific to the United States, we study whether algorithmic amplification of content from major media outlets is associated with political leaning.

We first report how personalization algorithms amplify content from elected officials from various political parties and parliamentary groups. We identified Twitter account details and party affiliation for currently serving legislators in seven countries from public data (25–28) (*SI Appendix*, section 1.B). The countries in our analysis were chosen on the basis of data availability: These

countries have a large enough active Twitter user base for our analysis, and it was possible to obtain details of legislators from high-quality public sources. In cases where a legislator has multiple accounts—for example, an official and a personal account—we included all of them in the analysis. In total, we identified 3,634 accounts belonging to legislators across the seven countries (the combined size of legislatures is 3,724 representatives). We then selected original tweets authored by the legislators, including any replies and quote tweets (where they retweet a tweet while also adding original commentary). We excluded retweets without comment, as attribution is ambiguous when multiple legislators retweet the same content. When calculating amplification relating to legislators, we considered their reach only within their respective country.

To compare the amplification of political groups, we can either calculate the amplification of all tweets from the group (group amplification; Fig. 1 *A* and *B*) or calculate amplification of each individual in the group separately (individual amplification; Fig. 1*C*). The latter yields a distribution of individual amplification values for each group, thus revealing individual differences of amplifying effects within a group.

Fig. 1*A* compares the group amplification of major political parties in the countries we studied. Values over 0% indicate that all parties enjoy an amplification effect by algorithmic personalization, in some cases exceeding 200%, indicating that the party's tweets are exposed to an audience 3 times the size of the audience they reach on chronological timelines. To test the hypothesis that left-wing or right-wing politicians are amplified differently, we identified the largest mainstream left or center-left and mainstream right or center-right party in each legislature, and present pairwise comparisons between these in Fig. 1*B*. With the exception of Germany, we find a statistically significant difference favoring the political right wing. This effect is strongest in Canada (Liberals 43% vs. Conservatives 167%) and the United Kingdom (Labor 112% vs. Conservatives 176%). In both countries, the prime ministers and members of the government are also members of the Parliament and are thus included in our analysis. We, therefore, recomputed the amplification statistics after excluding top government officials. Our findings, shown in *SI Appendix*, Fig. S2, remained qualitatively similar.

When studying amplification at the level of individual politicians (Fig. 1*C*), we find that amplification varies substantially within each political party: While tweets from some individual politicians are amplified up to 400%, for others, amplification is below 0%, meaning they reach fewer users on ranked timelines than they do on chronological ones. We repeated the comparison between major left-wing and right-wing parties, comparing the distribution of individual amplification values between parties. When studied at the individual level, a permutation test detected no statistically significant association between an individual's party affiliation and their amplification.

We see that comparing political parties on the basis of aggregate amplification of the entire party (Fig. 1 *A* and *B*) or on the basis of individual amplification of their members (Fig. 1*C*) leads to seemingly different conclusions: While individual amplification is not associated with party membership, the aggregate group amplification may be different for each party. These findings are not contradictory, considering that different politicians may reach overlapping audiences. Even if the amplification of individual politicians is uncorrelated with their political affiliation, when we consider increases to their combined reach, group-level correlations might emerge. For a more detailed discussion, please refer to *SI Appendix*, section 1.E.3.

Our fine-grained data also allow us to evaluate whether recommender systems amplify extreme ideologies, far-left or far-right politicians, over more-moderate ones (3). We found that, in countries where far-left or far-right parties have substantial representation among elected officials (e.g., VOX in Spain,

---

*A tweet is counted as "encountered" by user A when 50% of the UI element containing the tweet is continuously visible on the user's device for 500 ms. See *SI Appendix, section 1* for details.
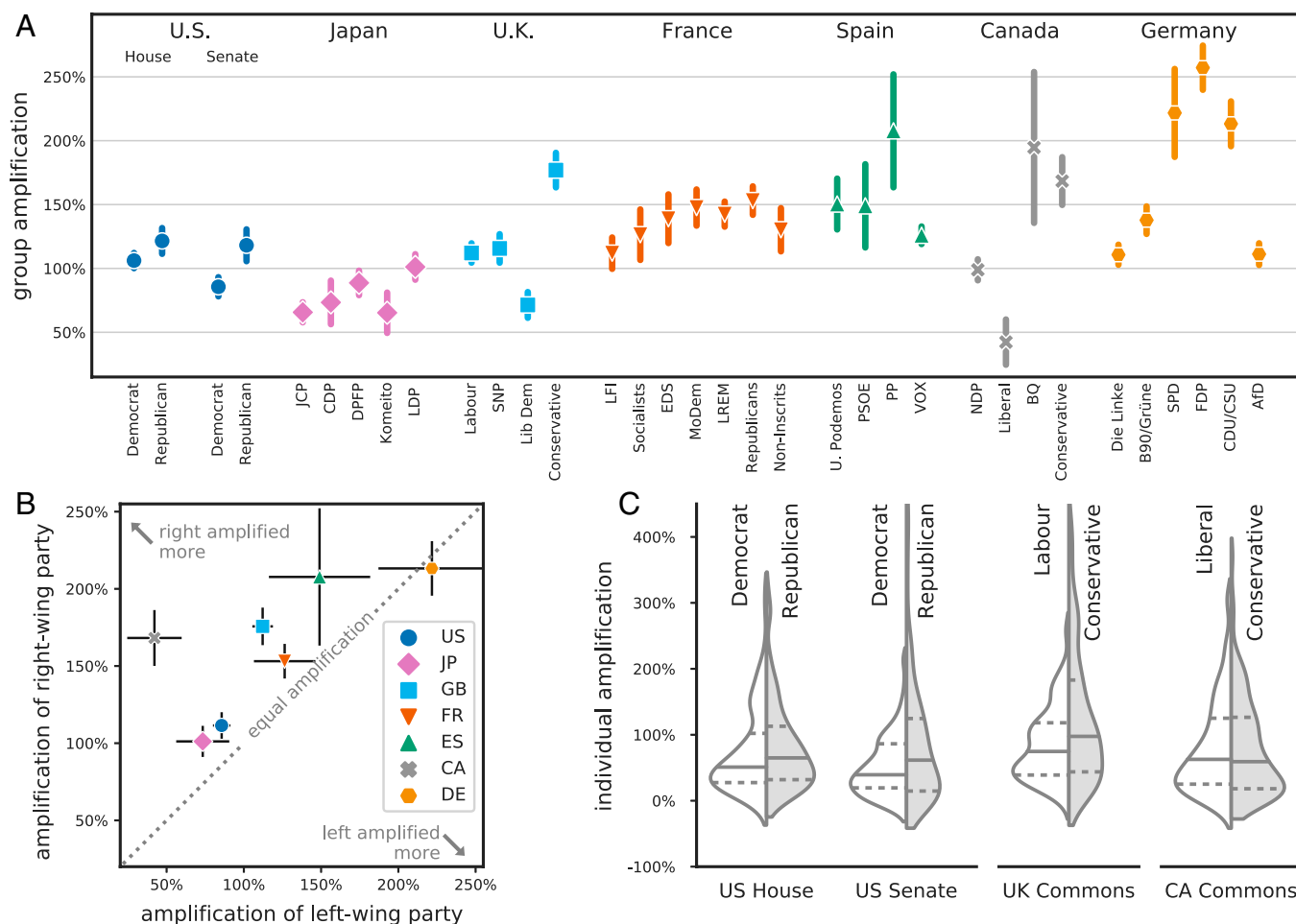
Fig. 1. Amplification of tweets from major political groups and politicians in seven countries with an active Twitter user base. (*A*) Group amplification of each political party or group. Within each country, parties are ordered from left to right according to their ideological position based on the 2019 Chapel Hill Expert Survey (29). A value of 0% indicates that tweets by the group reach the same number of users on ranked timelines as they do on chronological timelines. A value of 100% means double the reach. Error bars show SE estimated from bootstrap. Bootstrap resampling was performed over daily intervals as well as membership of each political group. (*B*) Pairwise comparison between the largest mainstream left- and right-wing parties in each country: Democrats vs. Republicans in the United States, Constitutional Democratic Party of Japan (CDP) vs. Liberal Democratic Party (LDP) in Japan, Labor vs. Conservatives in the United Kingdom, Socialists vs. Republicans in France, Spanish Socialist Worker's Party (PSOE) vs. People's Party (Partido Popular) in Spain, Liberals vs. Conservatives in Canada, and Social Democratic Party (SPD) vs. alliance of Christian Democratic Union and Christian Social Union (CDU/CSU) in Germany. In six out of seven countries, these comparisons yield a statistically significant difference, with right being amplified more, after adjusting for multiple comparisons. In Germany, the difference is not statistically significant. (*C*) Amplification of tweets by individual left- and right-wing politicians in the United States, United Kingdom, and Canada. Violin plots illustrate the distribution of amplification values within each party, solid lines show the median, dashed lines show 15th and 75th percentiles. There is substantial variation of individual amplification within political parties. However, there is no statistically significant dependence between an individual's amplification and their party affiliation, in any of the four comparisons. We used abbreviations LFI for La France Insoumise, EDS for Écologie Democratie Solidarité, PP for Partido Popular, and BQ for Bloc Québeqois.

Die Linke [The Left] and AfD [Alternative for Germany] in Germany, and La France Insoumise and Reasemblement national [National Rally] in France), the amplification of these parties is generally lower than that of moderate/centrist parties in the same country (*SI Appendix*, Fig. S1). Finally, we considered whether personalization consistently amplifies messages from the governing coalition or the opposition, and found no consistent pattern across countries. For example, in the United Kingdom, amplification favors the governing Conservatives, while, in Canada, the opposition Conservative Party of Canada is more highly amplified.

Tweets from legislators cover just a small portion of political content on the platform. To better understand the effects of personalization on political discourse, we extend our analysis to a broader domain of news content (30, 31). Specifically, we extend our analysis to media outlets with a significant audience in the United States (32). While the political affiliation of a legislator is publicly verifiable, there is no single agreed-upon classification of the political orientation of media outlets.

To reduce subjectivity in our classification of political content, we leverage two independently curated media bias–rating datasets from AllSides (33) and Ad Fontes Media (34), and present results for both.[†] Both datasets assign labels to media sources based on their perceived position on the US media bias landscape. The labels describe the overall media bias of a news source on a five-point scale ranging from partisan Left through Center/Neutral to partisan Right. We then identified tweets containing links to articles from these news sources shared by anyone between 1 April 2020 and 15 August 2020. We excluded tweets pointing to nonpolitical content such as recipes or sports. Wherever possible, we separated editorial content from

---

[†] Ad Fontes Media Bias Chart 5.0.

Huszár et al.
Algorithmic amplification of politics on Twitter

PNAS | 3 of 6
https://doi.org/10.1073/pnas.2025334119

general news coverage, as, in some cases, these had different bias ratings (*SI Appendix*, section 1.C). The resulting dataset contains AllSides annotations for 100,575,284 unique tweets pointing to 6,258,032 articles and Ad Fontes annotations for 88,818,544 unique tweets pointing to 5,100,381 articles.

We then grouped tweets by media bias annotation of their source and calculated the aggregate amplification of each bias category (Fig. 2). When using AllSides bias ratings (Fig. 2*A*), two general trends emerge: The personalization algorithms amplify sources that are more partisan compared to ones rated as Center. Secondly, the partisan Right is amplified marginally more compared to the partisan Left. The results based on Ad Fontes bias ratings (Fig. 2*B*) differ in some key ways. Most notable is the relatively low, 10.5%, amplification of the partisan Left compared to other categories. Among the remaining categories, the differences are not substantial, although the Neutral category is amplified significantly less than other categories.

Leave-one-out analysis of each media bias category (described in detail in *SI Appendix*, section 1.E.4) allows us to identify the most significant outliers in each category, also shown in Fig. 2. This analysis identified BuzzFeed News, LA Times, and Breitbart (based on both AllSides and Ad Fontes ratings) as negative outliers in their respective categories, meaning the amplification of their content was less than the aggregate amplification of the bias category they belong to. Meanwhile, Fox News and New York Post were identified as positive outliers. These outliers also illustrate that, just as we saw in the case of legislators, there is significant variation among news outlets in each bias category.

The fact that our findings differ depending on the media bias dataset used underlines the critical reliance of this type of analysis on political labels. We do not endorse either All-Sides or Ad Fontes as objectively better ratings, and leave it to the reader to interpret the findings according to their own assessment. To aid this interpretation, we looked at how AllSides and Ad Fontes ratings differ, where both ratings are available. We found that, while the two rating schemes largely agree on

rating the political right, they differ most in their assessment of publications on the political left, with a tendency for Ad Fontes to rate publications as being more neutral compared to their corresponding AllSides rating. Details are shown in *SI Appendix*, Figs. S3 and S4 and Table S1.

Discussion

We presented a comprehensive audit of algorithmic amplification of political content by the recommender system in Twitter's home timeline. Across the seven countries we studied, we found that mainstream right-wing parties benefit at least as much, and often substantially more, from algorithmic personalization than their left-wing counterparts. In agreement with this, we found that content from US media outlets with a strong right-leaning bias are amplified marginally more than content from left-leaning sources. However, when making comparisons based on the amplification of individual politician's accounts, rather than parties in aggregate, we found no association between amplification and party membership.

Our analysis of far-left and far-right parties in various countries does not support the hypothesis that algorithmic personalization amplifies extreme ideologies more than mainstream political voices. However, some findings point at the possibility that strong partisan bias in news reporting is associated with higher amplification. We note that strong partisan bias here means a consistent tendency to report news in a way favoring one party or another, and does not imply the promotion of extreme political ideology.

Recent arguments that different political parties pursue different strategies on Twitter (14, 15) may provide an explanation as to why these disparities exist. However, understanding the precise causal mechanism that drives amplification invites further study that we hope our work initiates.

Although it is the largest systematic study contrasting ranked timelines with chronological ones on Twitter, our work fits into a broader context of research on the effects of content personalization on political content (2, 3, 9, 21) and polarization (35–38). There are several avenues for future work. Apart from the
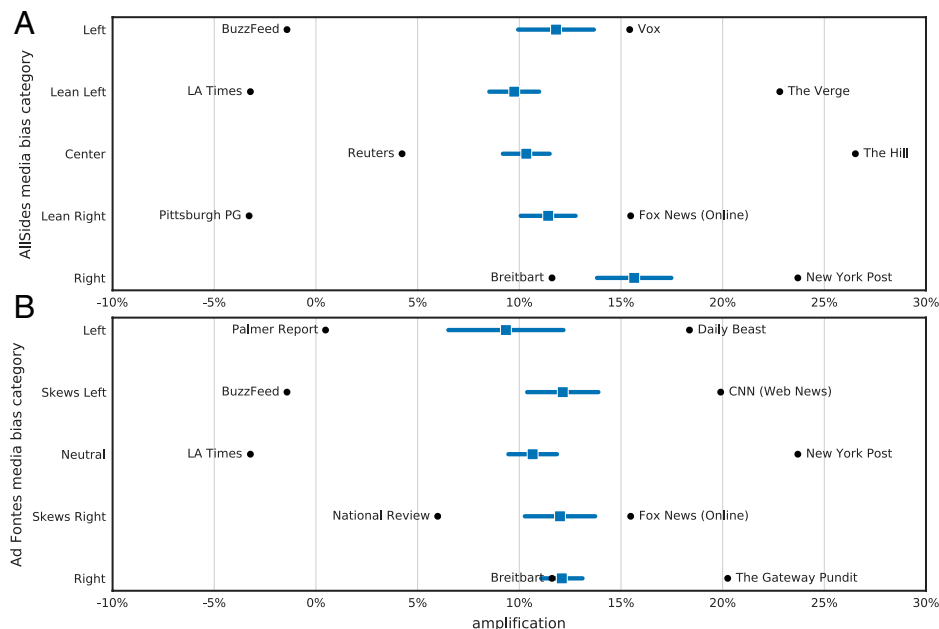

Fig. 2. Amplification of news articles by Twitter's personalization algorithms broken down by AllSides (*A*) and Ad Fontes (*B*) media bias ratings of their source. Blue squares denote the mean estimate of group amplification for each group of content, and error bars show the SD of the bootstrap estimate. Individual black circles show the amplification for the most significant positive and negative outliers within each group. For example, content from AllSides "Left" media bias category is amplified 12% by algorithms. The most significant negative outlier in this group is BuzzFeed, with an amplification of –2% compared to the chronological baseline. By contrast, Vox is amplified 16%. Negative and positive outliers are selected by a leave-one-out procedure detailed in *SI Appendix*, section 1.E.4.

Home timeline, Twitter users are exposed to several other forms of algorithmic content curation on the platform that merit study through similar experiments. Political amplification is only one concern with online recommendations. A similar methodology may provide insights into domains such as misinformation (39, 40), manipulation (41, 42), hate speech, and abusive content.

Materials and Methods

The Timelines Quality Holdback Experiment. Twitter has maintained the randomized experiment described in *Experimental Setup* since June 2016. Accounts were randomly assigned to treatment or control either at the experiment's onset or at the time the account was created. As of 5 June 2020, the experiment included 58 million unique Twitter user IDs (58,087,969, 5% of all accounts globally), of which 20% (11,617,373) are assigned to control, and 80% (46,470,596) are assigned to the treatment group. About 12% of studied accounts ($\sim$ 7 million) logged in within a single day of the study, and about 20% ($\sim$ 12 million) logged in within a single week. More information about the tweet selection, presentation, and ranking in either group, as well as the services and machine learning models influencing the content that users are exposed to through their Home Timeline, is provided in *SI Appendix*, section 1.A.

Ethical and Data Protection Reviews. The control group assessed was not created for the purpose of research but rather for the business purpose of improving the algorithm and providing a baseline to which it could be compared to monitor the ongoing performance of the algorithm. As such, this work was reviewed by Twitter's legal and privacy teams as part of its ordinary business operations (and not an IRB). As part of this review, a data protection impact assessment was conducted, and it was determined that additional notice and consent mechanisms were not required.

Obtaining Legislators' Twitter Details. We identified countries to include in our analysis based on the following criteria: 1) availability of data on politicians' Twitter accounts and 2) sufficient Twitter user base in the country. Screening for these criteria, we identified the following list of countries: United States, Japan, United Kingdom, France, Spain, Canada, Germany, and Turkey. Turkey was then excluded, due to limited availability of legislators' accounts for the current, 27th term (only about 18% of current legislators had a valid Twitter account). To identify members of the current legislative term in each country, we relied on Wikidata, public Twitter lists, and official government websites. While, in most countries, we were able to identify Twitter details of over 70% of all representatives following automated methods, our goal was to ensure that potentially missing accounts would not result in poor representation of certain minority groups in our dataset. We, therefore, focused manual annotation efforts on ensuring that accounts of legislators who belong to certain underrepresented groups are included in our dataset. In most countries, we were able to retrieve gender labels from Wikidata to aid with this process.

To test various hypotheses about the types of political parties algorithms might amplify more, we make some direct comparisons between parties in each country. We rely on the 2019 Chapel Hill Expert Survey (29) and Wikidata annotations to determine the ideological position of each party. More information on the data collection process from the aforementioned resources and groupings of parties is provided in *SI Appendix*, section 1.B.

Media Bias Ratings. We obtained media bias ratings for news sources from AllSides (33) and Ad Fontes Media (34). While the former includes news sources with a global audience, it focuses primarily on the US media landscape, and the media bias ratings relate to how the media bias of these sources is perceived in the United States. We excluded sites like Yahoo News and Google News, as well as podcasts, content studios, and activist groups whose original content was not clearly identifiable or attributable. To qualify for our analysis, the content from the publication source had to be clearly identifiable on the basis of URLs shared by users on the platform. AllSides provides categorical labels of media bias, while Ad Fontes provides numerical media bias ratings that are discretized into different categories based on the media bias chart.[‡]

To identify URLs that link to articles from each publication, we created regular expressions, which were matched against the text of the URL. We then identified tweets with content from these publications by screening public tweets created between 1 March and 30 June 2020, and matching any URLs these tweets contained against the regular expressions we curated. The resulting dataset contained AllSides annotations for 100,575,284 unique tweets pointing to 6,258,032 different articles and Ad Fontes annotations for 88,818,544 unique tweets pointing to 5,100,381 different articles. More information on the media bias ratings and the regular expressions used can be found in *SI Appendix*, section 1.C.

Measuring Amplification. Our measures of amplification are based on counting events called "linger impression," that is, events registered every time at least 50% of the area of a tweet is visible for at least 500 ms. Linger impressions are the best proxy available to us to tell whether a user has been exposed to the content of a tweet.

Let T denote a set of tweets. Let $U_{control}$ and $U_{treatment}$ denote the control and treatment groups of users, respectively, in the experiment. Note that, in our experiment, $|U_{treatment}| = 4|U_{control}|$. Let $U_{t,d}$ denote the set of users who registered a linger impression with tweet t on day d. For a set of tweets T, we further define $U_{T,d} = \bigcup_{t \in T} U_{t,d}$, the set of users who encountered at least one tweet from T on day d. We define the amplification of the set of tweets T on day d as

$$a_d(T) = \left(\frac{|U_{T,d} \cap U_{treatment}| + 1}{4|U_{T,d} \cap U_{control}| + 1} - 1 \right) \cdot 100\%.$$

Often, we consider amplification within a specific country. In this case, we calculate the above ratio based on impression events from an IP address that we identified to be within country c.

When we talk about the amplification of a group G of individuals, such as members of a political party, we mean the amplification of the set of all tweets created by this group T_G. The amplification for a group of users G is therefore $a(G) = a(T_G)$. Analogously, when referring to the amplification of an individual user i, we calculate this based on the set of tweets, T_i, or, equivalently, the group amplification for the singular set containing only i, that is, $a(i) = a(\{i\})$. A more detailed explanation of the group and individual amplification, as well as their differences, is presented in *SI Appendix*, sections 1.D and 1.E.

Data Availability. Aggregated study data are available upon request from the corresponding authors following the protocol outlined in *SI Appendix, section 3*.

[‡] https://www.adfontesmedia.com/interactive-media-bias-chart/.

1. S. Hong, S. H. Kim, Political polarization on twitter: Implications for the use of social media in digital governments. *Gov. Inf. Q.* **33**, 777–782 (2016).
2. The Economist, Twitter's algorithm does not seem to silence conservatives. *Economist*, 1 August 2020. https://www.economist.com/graphic-detail/2020/08/01/twitters-algorithm-des-not-seem-to-silence-conservatives. Accessed 1 September 2020.
3. Z. Tufekci, Youtube, the great radicalizer. *NY Times*, 10 March 2018. https://www.nytimes.com/2018/03/10/opinion/sunday/youtube-politics-radical.html. Accessed 1 September 2020.
4. M. D. Conover et al., Political polarization on twitter. *ICWSM* **133**, 89–96 (2011).
5. C. A. Bail et al., Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9216–9221 (2018).
6. I. Himelboim, S. McCreery, M. Smith, Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on twitter. *J. Computer Mediated Commun.* **18**, 154–174 (2013).
7. P. J. Hasson, *The Manipulators: Facebook, Google, Twitter, and Big Tech's War on Conservatives* (Regnery, 2020).
8. P. Barberá, J. T. Jost, J. Nagler, J. A. Tucker, R. Bonneau, Tweeting from left to right: Is online political communication more than an echo chamber? *Psychol. Sci.* **26**, 1531–1542 (2015).
9. D. O'Callaghan, D. Greene, M. Conway, J. Carthy, P. Cunningham, Down the (white) rabbit hole: The extreme right and online recommender systems. *Soc. Sci. Comput. Rev.* **33**, 459–478 (2015).
10. T. Gillespie, The politics of 'platforms.' *New Media Soc.* **12**, 347–364 (2010).
11. E. Bozdag, J. van den Hoven, Breaking the filter bubble: Democracy and design. *Ethics Inf. Technol.* **17**, 249–265 (2015).
12. M. H. Ribeiro, R. Ottoni, R. West, V. A. Almeida, W. Meira Jr., "Auditing radicalization pathways on YouTube" in *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery, 2020), pp. 131–141.
13. C. Shirky, The political power of social media: Technology, the public sphere, and political change. *Foreign Aff.* **90**, 28–41 (2011).
14. J. H. Parmelee, S. L. Bichard, *Politics and the Twitter Revolution: How Tweets Influence the Relationship between Political Leaders and the Public* (Lexington, 2011).
15. D. Freelon, A. Marwick, D. Kreiss, False equivalencies: Online activism from left to right. *Science* **369**, 1197–1201 (2020).
16. @mjahr, "Never miss important tweets from people you follow." *Twitter*, 10 February 2016. https://blog.twitter.com/en_us/a/2016/never-miss-important-tweets-from-people-you-follow. Accessed 1 August 2020.

Huszár et al.
Algorithmic amplification of politics on Twitter

PNAS | 5 of 6
https://doi.org/10.1073/pnas.2025334119

POLITICAL SCIENCES

COMPUTER SCIENCES

17. A. Badawy, E. Ferrara, K. Lerman, "Analyzing the digital traces of political manipulation: The 2016 Russian interference twitter campaign" in *The 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, U. Brandes, C. Reddy, A. Tagarelli, Eds. (Institute of Electrical and Electronics Engineers, 2018), pp. 258–265.

18. M. D. Conover, B. Gonçalves, A. Flammini, F. Menczer, Partisan asymmetries in online political activity. *EPJ Data Sci.* **1**, 6 (2012).

19. M. Vergeer, L. Hermans, S. Sams, Online social networks and micro-blogging in political campaigning: The exploration of a new campaign tool and a new campaign style. *Party Polit.* **19**, 477–501 (2013).

20. T. Graham, M. Broersma, K. Hazelhoff, G. Van'T Haar, Between broadcasting political messages and interacting with voters: The use of Twitter during the 2010 UK general election campaign. *Inf. Commun. Soc.* **16**, 692–716 (2013).

21. E. Bakshy, S. Messing, L. A. Adamic, Political science. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).

22. P. M. Aronow et al., Estimating average causal effects under general interference, with application to a social network experiment. *Ann. Appl. Stat.* **11**, 1912–1947 (2017).

23. D. R. Cox, *Planning of Experiments* (Wiley, 1958).

24. D. Eckles, B. Karrer, J. Ugander, Design and analysis of experiments in networks: Reducing bias from interference. *J. Causal Inference* **5**, 20150021 (2016).

25. D. Vrandečić, M. Krötzsch, Wikidata: A free collaborative knowledgebase. *Commun. ACM* **57**, 78–85 (2014).

26. UK Parliament, *MPs and lords.* https://members.parliament.uk. Accessed June 2020.

27. Ballotpedia, List of current members of the U.S. Congress. https://ballotpedia.org/List_of_current_members_of_the_U.S._Congress. Accessed June 2020.

28. House of Commons Canada, Current members of Parliament. https://www.ourcommons.ca/members/en/search. Accessed June 2020.

29. R Bakker, et al., 2019 Chapel Hill Expert Survey, version 2019.1. https://www.chesdata.eu/s/CHES2019V3.csv. Accessed 15 May 2021.

30. S. Messing, S. J. Westwood, Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online. *Communic. Res.* **41**, 1042–1063 (2014).

31. S. Flaxman, S. Goel, J. M. Rao, Filter bubbles, echo chambers, and online news consumption. *Public Opin. Q.* **80**, 298–320 (2016).

32. American Press Institute, The personal news cycle: How Americans choose to get their news. https://www.americanpressinstitute.org/publications/reports/survey-research/personal-news-cycle/. Accessed October 2020.

33. AllSides, How AllSides rates media bias. https://www.allsides.com/media-bias/media-bias-rating-methods. Accessed August 2020.

34. V. Otoro, How Ad Fontes rates the news: A methodology overview. https://adfontesmedia.com/how-ad-fontes-ranks-news-sources/. Accessed August 2020.

35. N. Perra, L. E. C. Rocha, Modelling opinion dynamics in the age of algorithmic personalisation. *Sci. Rep.* **9**, 7261 (2019).

36. E. Pariser, *The Filter Bubble: What the Internet Is Hiding From You* (Penguin UK, 2011).

37. M. Del Vicario et al., Echo chambers: Emotional contagion and group polarization on Facebook. *Sci. Rep.* **6**, 37825 (2016).

38. L. Boxell, M. Gentzkow, J. M. Shapiro, Greater Internet use is not associated with faster growth in political polarization among US demographic groups. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 10612–10617 (2017).

39. M. Del Vicario et al., The spreading of misinformation online. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 554–559 (2016).

40. N. Grinberg, K. Joseph, L. Friedland, B. Swire-Thompson, D. Lazer, Fake news on Twitter during the 2016 U.S. presidential election. *Science* **363**, 374–378 (2019).

41. S. C. Woolley, Automating power: Social bot interference in global politics. *First Monday* **21**, 4 (2016).

42. S. Aral, D. Eckles, Protecting elections from social media manipulation. *Science* **365**, 858–861 (2019).