Science AAAS

Supplementary Materials for

Supersharers of fake news on Twitter

Sahar Baribi-Bartov, Briony Swire-Thompson, Nir Grinberg

Corresponding author: Nir Grinberg, nirgrn@bgu.ac.il

Science **384**, 979 (2024)
DOI: 10.1126/science.adl4435

The PDF file includes:

Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S15
References

Other Supplementary Material for this manuscript includes the following:

MDAR Reproducibility Checklist

Materials and Methods

Twitter Panel

The foundation of this work is a large-scale online panel of registered U.S. voters on Twitter. The panel was constructed by matching U.S. voter records to Twitter accounts using full names and location information. Records were matched if a person's full name matched in both datasets and they were the only person with that name at either the city or state-level geographic area in both datasets. This approach was first published by Grinberg et al. *(4)* and was later expanded and used by Shugars et al. and Green et al. *(33, 34)*. Importantly, a comparison of the panel to a gold-standard, random-sampling survey conducted by Pew Research Center showed that the panel is largely reflective of registered U.S. voters on Twitter with only small differences due to the reliance on unique names and possibly uneven response rates *(35)*. See SM.1 for additional details about panel construction and representativeness.

The sample used throughout this work consists of 664,391 panel members who were minimally active on Twitter around the 2020 presidential election from July to November 2020 (inclusive) by liking or tweeting at least once during this period. This large sample of panelists enabled us, for the first time, to identify a statistically meaningful sub-sample of 2,107 fake news supersharers, which is the focus of this work.

The panel also enables a unique linkage between sociodemographic information from administrative records about registered voters with their online activity on Twitter, including profile information and online social network. In particular, we include in our models variables that describe individuals' age, gender, race, party registration, residence in a swing state[1], and the number of Twitter followers. We also use data from the 2018 Census to augment this information with regional variables describing educational attainment, income, and urban/rural classification at the tract level (see SM.1 for more details).

Identifying Political News

We used a machine learning classifier to identify political news and validated its accuracy against human coders (as per 4, 36). We trained a logistic regression classifier using the text of tweets with URLs (including retweets, quotes, and replies), where we considered as positive examples tweets that contain whitelist terms and a random sample of tweets as negative examples. The whitelist terms contained general political keywords (e.g., election2020, mail-in), candidate names and handles (e.g., joebiden, trump), and various hashtags (e.g., votebidenharris2020). The classifier was then trained separately on each day of data to identify additional terms that are associated with politics. Finally, crowdworkers evaluated a sample of more than 2,000 tweets, stratified over time, to determine whether those tweets were about the U.S. presidential election, U.S. politics in general, non-U.S. politics, or other topics. Compared to the gold standard of human judgment, the classifier resulted in a precision of 88.8% and recall of 80.0% for political tweets. The classifier was able to retrieve 96.4% of election-related tweets. For more details about the classifier and its validation, see SM.2.

[1] Based on Wikipedia list of 2020 swing states available at https://en.wikipedia.org/wiki/Swing_state.

Identifying Fake News

We adopt a source-level definition of fake news *(4, 6)* defining it as news outlets with the trappings of legitimately produced news that lack the "editorial norms and processes for ensuring the accuracy and credibility of information" *(17)*. Operationally, we relied on the manual labeling of sources conducted by Grinberg et al. *(4)* and extended our list of sources by using NewsGuard's more comprehensive and up-to-date ratings. NewsGuard is an organization that employs trained journalists and experienced editors to review and rate news sources regularly. NewsGuard's list contained over 6,000 news sources in five different countries and their methodology is described in full[2].

NewsGuard ratings include nine criteria that describe a source's failure to adhere to basic journalistic practices of credibility and transparency such as "does not repeatedly publish false content", "gathers and presents information responsibly" or "regularly corrects or clarifies errors", and it allows the exclusion of satirical sites and other social media platforms[2]. We combined the nine adherence criteria into a single binary classification by training a logistic regression classifier with the labels from Grinberg et al. *(4)*. We tested the robustness of our results using multiple operationalizations of fake news (see SM.3 for details). The main results are reported using domains that violate the same criteria as the sources in Grinberg et al.'s list, excluding satire, and fail at least three of NewsGuard's journalistic criteria. Results were extremely similar when using NewsGuard's Red label for low-credibility sites.

Quantifying Individual Exposure

Precise information about individual-level exposure to content on social media is not publicly available on any API or platform. Previous research worked around this limitation by estimating individuals' potential exposure by using the content available to individuals from the accounts they follow *(2, 4)*. Similar to Grinberg et al. *(4)*, we based our estimates on the 10% random sample of Twitter content (Decahose) to evaluate the composition of content available to panelists from their Twitter friends. It is important to note that this approach does not provide information about content that panelists necessarily saw, but rather an estimate of the composition of available content from their peers. Notably, this approach does not take into account out-of-network content that Twitter's recommendation algorithm may introduce to a user's feed, nor does it consider the ordering of content. Nevertheless, peer content has been dominant on Twitter throughout the years. In 2023, Twitter revealed that about 50% of the content available to users in its algorithmic "For you" feed comes from peers[3]. The other main interface to content, now called the "Following" feed, has always been populated with peer content.

Moreover, examining peer content requires network information, which we were only able to retrieve through the Twitter API for 641,144 users (96.5% of panelists). To reliably estimate the fraction of fake news in political news available to panelists from their network, we limit the analysis to individuals with at least 100 political news from their network during the entire election period, although the results are robust to different thresholds (see SM.10). For each individual, we calculated the fraction of tweets linking to fake news sources out of all political

---

[2] https://www.newsguardtech.com/ratings/rating-process-criteria, also available in the Internet Archive at https://web.archive.org.

[3] https://blog.twitter.com/engineering/en_us/topics/open-source/2023/twitter-recommendation-algorithm, also available in the Internet Archive.

news available from their network as well as the fraction of fake news directly coming from supersharers.

Supplementary Text

S.1 Panel Construction and Representativeness

We rely on a panel of 1.5 million Twitter users and their activity on Twitter, dating back to 2017. The panel was created by linking U.S. voter registration records to Twitter accounts. The panel contains the profile information and online activity of Twitter accounts associated with people who live in the United States and are registered to vote. The voter file used was provided by TargetSmart, one of the leading companies providing up-to-date U.S. voter records. The voter records include various information including each individual's name, age, gender, race, and party registration. The first step was collecting data from 290 million profiles using Twitter's 10% Decahose sample *(4, 33, 35)*. These profiles represent a set of accounts that were active between January 2014 and March 2017. The matching was done by linking accounts by the first and last name and location (city and state or just state if the city was not present in the Twitter profile info) of the person, only if it was unique in both datasets. Additional details about the panel construction can be found in SM.1 and SM.8 of Grinberg et al. *(4)*, and in the Materials and Methods of Shugars et al. *(33)*. In the current work, we analyzed the activity of all panel members who were minimally active during the 2020 presidential elections as described in the Materials and Methods section.

Panel Representativeness: The panel covers all 50 states in the U.S. as well as the District of Columbia, and accounts for approximately 3% of all adult U.S. Twitter users *(33)*. Importantly, Hughes et al. *(35)* showed that the panel has little demographic bias compared to a standard probability sample of survey respondents who are registered to vote. By comparing the panel to a sample constructed by Pew Research Center, they showed that the panel used in this work is slightly over-representing white and female individuals, under-representing Hispanics and Asian users, and appropriately representing African-Americans. To explain the differences, Hughes et al. compared registered Twitter users to all registered voters, evaluating whether the samples were different to begin with or whether the differences were more likely caused by data processing procedures. They showed that the voter records had a higher proportion of registered voters that are white, female, and democratic, and a notably smaller proportion who are Asian American or who identify with other racial groups. The high proportion of white and female registered voters in the voter record can help explain the differences in the panel. Other differences such as the under-representation of Asian Americans could stem from demographic biases in the unique name and location restrictions of the matching process or uneven response rates in the survey data *(34)*. Notably, the panel only includes registered U.S. voters, which were estimated in 2012 to be about 78% of all eligible voters in the U.S. *(37)*. Prior work had also noted that while Twitter and Facebook are not representative of the general population, differences in political attention, values, or behavior mostly vanish when controlling for a few demographic variables *(38)*.

Supplementing income and educational attainment variables: In addition to age, race, gender, and party of registration, we wanted to understand how supersharers vary in terms of their income and education, which are not available in voter records. To supplement the individual-level information, we collected information about income and education attainment from the U.S. Census Bureau at the tract level (2018 Census data, 5-year aggregation). The census data for income and education attainment is categorical, and we applied two transformations to obtain a numeric value: (i) taking the most likely category, and (ii) taking the weighted average of categories. For income, we represented each income bracket (category) by the middle of the bin value, with the highest income bracket of $200,000 or more represented by $200,000. For education attainment, we mapped the categories into the number of years that typically takes to complete the level of education as specified in Table S1. The income and education variables we computed at the tract level were then linked back to individuals based on their Federal Information Processing Standard (FIPS) code, which uniquely identifies geographic areas in the U.S. If the tract-level data was unavailable, we backtracked to the county level information, then further to the state level if necessary.

Descriptive Statistics about Panelists: Table S2 below provides summary statistics about the panel as a whole and the various sub-sample groups used for comparison throughout this work.

Education							
Category	**Years**	**Category**	**Years**	**Category**	**Years**	**Category**	**Years**
No schooling	0	4th grade	4	10th grade	10	1+ year college	14
Nursery school	0	5th grade	5	11th grade	11	Associate's degree	14
Kindergarten	0	6th grade	6	12th grade	12	Bachelor's degree	15
1st grade	1	7th grade	7	High school diploma	12	Master's degree	17
2nd grade	2	8th grade	8	GED	12	Professional degree	15
3rd grade	3	9th grade	9	<1 year college	13	Doctorate degree	22

Table S1.
Mapping of educational attainment categories to education years.

	Panel	Panel subsample	SS-NF	Avg. fake news sharer	Supersharers
Sample size (N)	664,391	11,199	11,199	10,464	2,107
Average age	41.3	41.2	51.4	53.0	58.2
% Female	49.9%	49.9%	52.6%	49.1%	58.8%
% White	83.8%	83.8%	87.0%	90.2%	91.7%
% registered Republicans	14.0%	14.0%	10.11%	25.95%	38.03%
Avg. education years	13.7	13.7	13.9	13.7	13.6
Avg. income	85,171	85,475	89,364	84,953	85,906
% urban	90.7%	90.7%	94.3%	90.5%	89.6%
% in Swing States	33.2%	33.4%	31.4%	38.3%	39.1%
# of followers (median)	163	168	720	358	1148
# Daily tweets with URLs shared	0.94	0.83	7.41	5.70	18.49
# Daily political news shared	0.39	0.30	5.01	4.41	15.86
# Daily fake news articles	0.01	0.01	0.13	0.35	2.76

Table S2.

Demographics and additional statistics about the full panel and the subgroups used throughout the work: a random sub-sample of panelists (Panel subsample), supersharers, SS-NF (the heaviest sharers of non-fake political news), and the "average" fake news sharers.

S.2 Political Classifier

Using the same approach utilized in prior work *(4, 36)*, we trained a logistic regression classifier using the text of tweets containing URLs. We did not distinguish between tweet types, i.e. we included original tweets, retweets, quotes, and replies. The classifier was trained using tweets that contain whitelist terms as positive examples and a random sample of tweets as negative examples. The classifier was trained separately on each day of data to identify the additional terms that are associated with politics each day. Empirically, we found that L2 regularization with a lambda value of 0.01 resulted in the best performance, and therefore used this model as our classifier. The whitelist terms were composed of general political keywords, candidate names and handles, and hashtags related to the 2020 election. The full list of keywords is shown in Table S3.

We evaluated the classifier using crowdsourced labels of 2,065 tweets, which were stratified over the days of the study. All tweets were annotated by at least two workers on Amazon's Mechanical Turk and disagreements were resolved by additional labeling by the authors. Following the coding scheme of Grinberg et al. *(4)*, crowdworkers were asked to assign tweets into the category that best describes the tweet from the following options: (i) U.S. presidential election, (ii) U.S. politics, (iii) Non-U.S. politics, (iv) Other, and (v) I don't know. Based on the human labels, we found that the classifier was able to recall 96.4% of all U.S. presidential election tweets. When considering the broader category of political tweets, tweets labeled as pertaining to the U.S. presidential election or U.S. politics more generally, the classifier achieves a precision of 88.8% and a recall of 80%. These results exceed the performance reported in Grinberg et al. *(4)* for the same task and demonstrate that the classifier aligns well with the concept of U.S. politics and can retrieve most of the relevant content.

While the overall performance of the classifier was relatively high, it is still possible that the classifier could have been biased. The classifier could be identifying fake news content as political at a higher or lower rate, and it could be biased based on the political leaning of the content. To evaluate these potential biases, we examined the classifier labels relative to the crowdsourced ones on different subsets of the annotated tweets. In particular, we examined whether the classifier exhibited bias in detecting fake news, and whether there is bias in detecting political content from left- versus right-leaning sources.

Figure S1 shows that such a bias is unlikely to have occurred in identifying fake news. The figure presents the percentage of content from fake news sources in the annotated set for political content (left) and non-political content (right). We found no statistically significant difference in the set of identified fake news when using crowdworkers or classifier labels for both political and non-political content. Specifically, we found that 4.7% of the tweets labeled by crowdworkers as political contained links to fake news sources, compared to 5.3% in the tweets labeled by the classifier as political (P=0.36). In non-political tweets, the classifier and crowdworkers effectively found no fake news.

Next, we examined the classifier potential bias in identifying political content from left- versus right-leaning sources. To that end, we used the publicly available list of domain alignment scores[4] collated by Robertson et al. *(39)*. We considered sources with a negative alignment score of -0.1 or lower as left-leaning, and sources with a positive alignment score of 0.1 or higher as right-leaning.

---

[4] https://github.com/gitronald/domains/tree/master

Figure S2 shows the percentage of annotated tweets that were labeled as political (left) or non-political (right) in tweets linking to left-leaning sources (top) and right-leaning sources (bottom). We found no statistically significant differences in any of the four comparisons conducted (P=0.20 for political left; P=0.23 for political right; P=0.44 for non-political left; P=0.30 for non-political right). Relative to crowdworkers political labels, the classifier under-identified right-leaning content (31.4% vs. 33.6%) and over-identified left-leaning content (37.4% vs. 34.6%). In terms of non-political content, the classifier over-identified right-leaning content (27.5% vs. 26.7%) and under-identified left-leaning content (16.4% vs. 16.6%). Given that these differences are small (<3%) and not statistically significant, we conclude that it is highly unlikely that such a small difference is a dominant factor behind the large party registration differences we have identified between groups.

Finally, based on the observation that the fraction of content linking to fake news sources in political news increased after the election (see Figure 1A in the main body), we examined a random sample of nearly 300 post-election tweets that linked to fake news sources. We observed that most of the tweets supported various election fraud arguments.

General terms
election, presdebate, vpdebate, democratic, gop, dnc, rnc, politics, political, voter, senate, senator, 2020election, election2020, electionday, votebymail, votersuppression, ballot, mailin, mail-in, mail in, russiahoax, qanon, obamagate, mailinballots, nakedballots, presidential, vote-by-mail, votingsquad, votethemout, wewillvote, blackvotesmatter.
Elected officials
mike pence, michael pence, mikepence, michaelpence, pence, kamala harris, kamala, harris, spike cohen, angela walker, kamalaharris, senkamalaharris
Candidates
joe biden, joebiden, biden, votebiden, bluewave2020, votebidenharris2020, ridinwithbiden, nomalarkey, biden2020, bidenharris2020, bidenforpresident, bidenkamala2020, joebiden2020, bidenharris, votehimout, Dumptrump, nevertrump, bluewave, fucktrump, bidenwarroom, voteblue, demconvention, votebluetosaveamerica, wakeupamerica, trumpisanationaldisgrace, trumpvirus, trumpisalaughingstock, traitortrump, jo jorgensen, jojorgensen, joanne marie jorgensen, jorgensen2020, beboldvotegold, donald trump, don trump, realdonaldtrump, donaldtrump, donaldjtrump, donald j trump, trump, trumpwarroom, teamtrump, the donald, trump2020, maga, draintheswamp, keepamericagreat, neverbiden, trumppence2020, makeamericagreatagain, kag, presidenttrump, notmypresident, americafirst, redwave, votered, sleepy joe, sleepyjoe, hidenbiden, creepyjoebiden, bidenukrainescandal, rnc2020, kag2020, maga2020, trump2020landslide, tulsatrumprally, voteredtosaveamerica, trumpforpresident, backtheblue, howiehawkins, howie hawkins, howiehawkins2020, hawkins2020

Table S3.

The lists of whitelist terms used to identify political tweets with high probability.

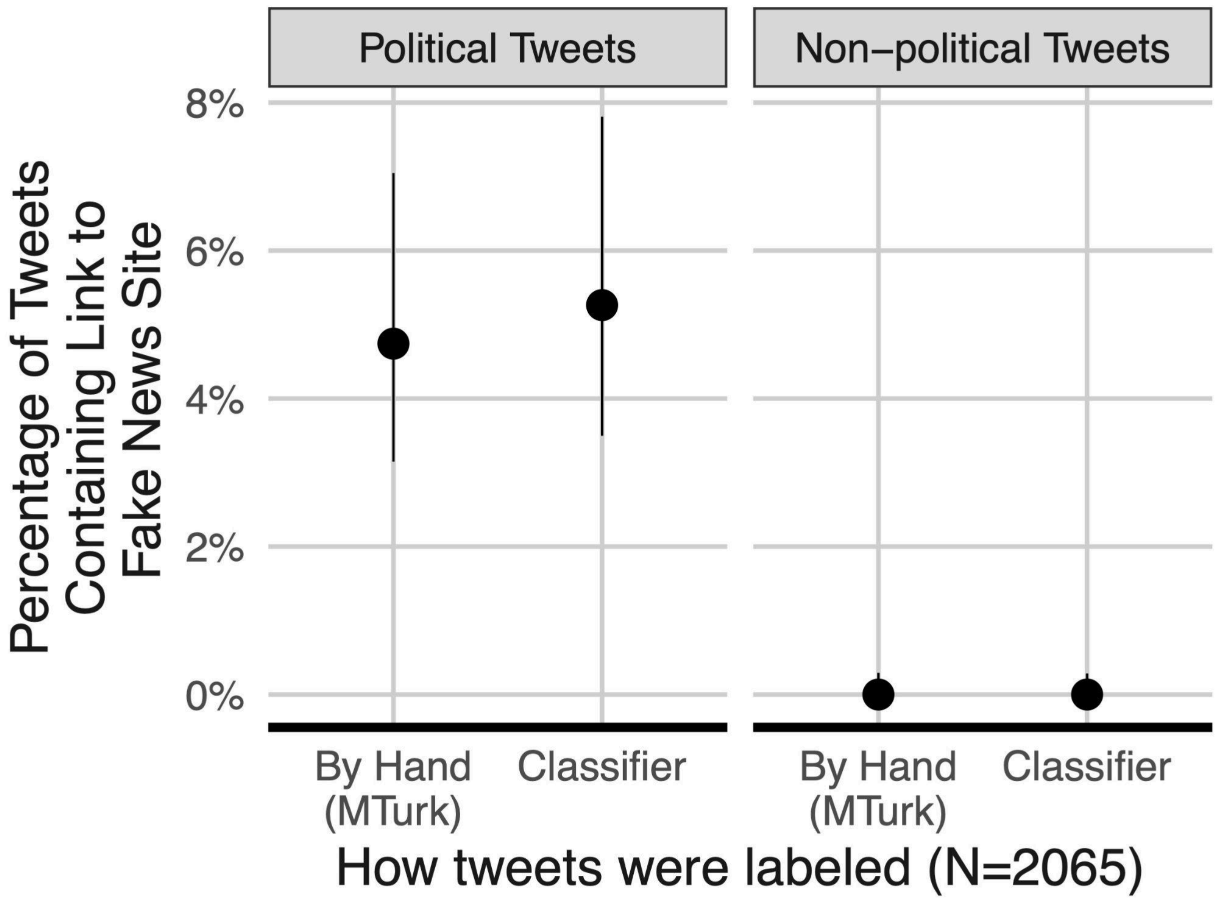

Figure S1.
The percentage of content from fake news sources (y-axis) in annotated tweets for political (left panel) and non-political (right panel) tweets.

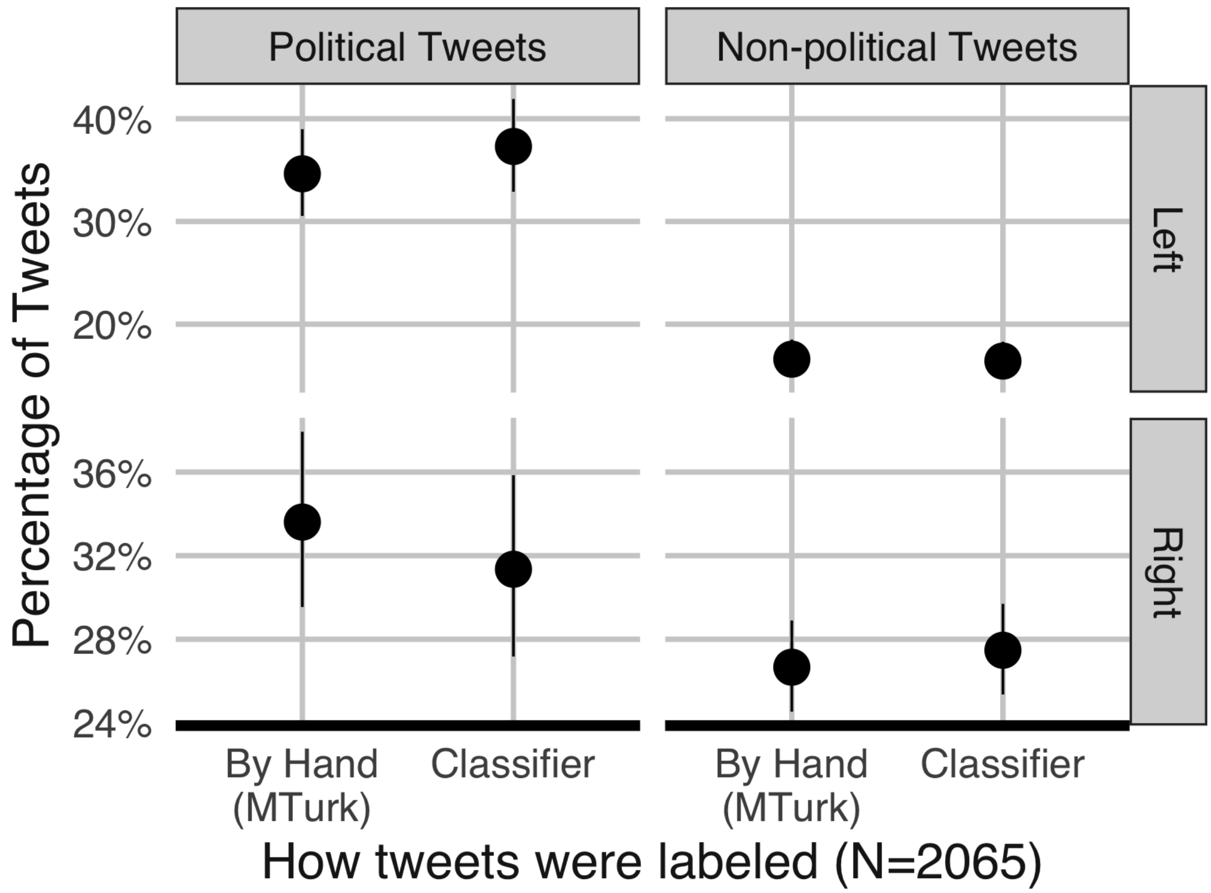

Figure S2.
The percentage of tweets from left-leaning (top) or right-leaning (bottom) sources, and separated to tweets labeled political (left) and non-political (right).

S.3 Identifying Fake News Sources and Robustness Checks

In this section, we describe how we identified fake news sources by expanding and updating the manual labeling of sources conducted by Grinberg et al. *(4)* using NewsGuard ratings, and evaluated the robustness of our findings to an alternative operationalization.

As described in the main body, we rely on a definition given by Lazer et al. *(17)* for fake news sources that was operationalized in prior work *(4)*. In order to extend this to include fake news sources that operated in the 2020 election, we train a model that learns the association between NewsGuard ratings to the labels from Grinberg et al. *(4)*. In particular, we trained a logistic regression classifier on the list of domains from Grinberg et al., where Black, Red or Orange labeled sources were considered as 'fake' and all other sources non-fake. As features, the model used NewsGuard's nine criteria for adherence to journalistic practices (e.g., author attribution, truthful headlines, offering of corrections).

This initial model successfully predicted the labels of held-out sources from the Grinberg et al. list with 96% accuracy [95% CI of (83%, 99%)]. Nevertheless, we wanted the final model to be particularly careful about inaccurately labeling domains as fake news (type I error). Therefore, we considered a site as fake if and only if it met two stringent criteria: (i) The model was "confident" in labeling the site as fake, i.e., assigning a predicted probability that is considerably higher than 0.5 – we used 0.67 as the threshold, and (ii) The site failed at least three out of the nine NewsGuard criteria, ensuring that no site is considered as fake solely based on a single criterion or two.

This model resulted in labels that were highly consistent with NewsGuard's credibility ratings and aligned more closely with prior academic work. Table S4 shows the number of fake and non-fake sources that match NewsGuard's low-credibility and high-credibility labels. The numbers along the diagonal show that our labels and the NewsGuard labels are identical in 97.7% of the cases (5384/5509). The vast majority of the remaining sites (96/125) are sites that NewsGuard considered as low-credibility, but our more stringent model considered as not meeting the bar for inclusion as fake sources. Manually examining these cases revealed that most of these sources failed multiple transparency criteria (e.g., not disclosing ownership or properly listing authors) and some journalistic standards (e.g., not issuing corrections), but still often produced factual headlines and content. Notable examples in this category include domains like washingtontimes.com and breitbart.com. Our model labeled an additional 29 news sites as fake sources that NewsGuard considered as not crossing the threshold of low credibility, all of which publish deceptive headlines, mostly without issuing corrections, or disclosing ownership. Notable sites include dailycaller.com, deadspin.com, and occupydemocrats.com, which prior work considered as sources of fake news *(4)*.

To assess the robustness of our results to alternative definitions of fake news sources, we repeated our main regression analysis where instead of the more stringent model described above, we used NewsGuard low-credibility labels. In other words, we repeated the analyses where fake sources were those with a NewsGuard score below 60. The results of these multi-level logistic regressions are presented in Table S5. Section S4 contains the full description of these individual-level regression models and the full regression results in Table S6. Note that the coefficients' sign, magnitude, and significance levels are largely the same, which demonstrates the robustness of our findings to this alternative operationalization of fake news sources. It should also be noted that NewsGuard's credibility ratings are highly consistent with five independent source-credibility assessments by experts, showing the reliability of the definition *(40)*. It is an open question how comprehensive the coverage is of source credibility raters, and whether it has any systematic bias against a particular political ideology. Future research should examine this along with the availability and production of misinformation by different ideologies.

Finally, to assess the robustness of our source-level definition at the level of a story, we manually labeled a random sample of 50 tweets with links to fake news sources, and 50 random tweets with links to non-fake sources. For each news story, two of the authors read the article and conducted a web search to find whether the article's main claim had been falsified by third-party fact-checkers (e.g., organizations like Politifact, the Associated Press, or FactCheck.org). We found that the main claim in 17 of the 50 tweets (34%) linking to fake news sources could be directly traced to a known falsehood that was published on a trusted fact-checker's website. We found that only one of the 50 tweets (2%) linking to non-fake news sources could be tied to a known falsehood. The fake news set had a significantly higher number of stories publishing false claims ($P<0.001$). These results provide further support for the definition of fake news at the level of a source, and highlight that the veracity of stories from fake news sources should often be called into question.

		NewsGuard label	
		High-credibility (score >= 60)	Low-credibility (score < 60)
Classifier label	Non-fake	3,116 (99.1%)	96 (4.1%)
	Fake	29 (0.9%)	2,268 (95.9%)

Table S4.

Confusion matrix comparing NewsGuard labels with the classification of sources into fake and non-fake based on NewsGuard's nine flags of adherence to journalistic practices. Percentages in each column show that the classifier label is consistent with NewsGuard's label in more than 95% of sources, in both the low- and high-credibility groups.

Relative to	SS-NF	Panel	Avg. fake news sharers
Constant	-2.77*** (0.43)	-6.35*** (0.48)	-5.02*** (0.43)
Age (30-44)	0.82*** (0.22)	1.96*** (0.22)	0.83*** (0.22)
Age (45-65)	1.69*** (0.21)	3.75*** (0.21)	1.46*** (0.21)
Age (Over 65)	2.01*** (0.22)	4.69*** (0.22)	1.76*** (0.21)
Female	0.16** (0.06)	0.25*** (0.06)	0.31*** (0.05)
Gender Unknown	0.49 (0.25)	0.02 (0.27)	0.29 (0.24)
White	0.44*** (0.11)	0.51*** (0.12)	0.17 (0.11)
Race N/A	0.34 (0.21)	0.42 (0.25)	0.43* (0.21)
Republican	2.81*** (0.10)	2.38*** (0.11)	1.67*** (0.10)
Independent	1.58*** (0.11)	1.25*** (0.13)	1.29*** (0.12)
Other party	1.66*** (0.22)	2.02*** (0.26)	1.28*** (0.22)
Party N/A	1.75*** (0.09)	1.49*** (0.11)	1.24*** (0.10)
Avg education	-0.26*** (0.03)	-0.31*** (0.03)	-0.11*** (0.03)
Standarized residual income	0.13*** (0.03)	0.03 (0.03)	0.10*** (0.03)
Location urban	-0.30*** (0.10)	0.02 (0.11)	-0.04 (0.09)
Location N/A	2.72 (2.45)	2.00 (1.44)	2.19 (1.36)
Swing state	0.17** (0.06)	0.27*** (0.07)	0.05 (0.05)
Log(1 + # followers)	0.42*** (0.04)	1.41*** (0.04)	0.70*** (0.03)
Observations	13,305	13,305	12,571
Log-likelihood	-4,532.71	-3,405.90	-4,950.04
Akaike IC	9,103.42	6,849.80	9,936.09
Note:	*p<0.05; **p<0.01; ***p<0.001		

Table S5.
Regression results when considering as fake news sources with NewsGuard score below 60. The three logistic regression models (in columns) distinguished supersharers from a reference group: the heaviest sharers of non-fake political news (SS-NF), a random subsample of panelists (Panel), and the "average" fake news sharers. See SM.4 for more details about the covariates used in these models. All findings in the main text are replicated.

S.4 Full Regression Models Specifications

The results identifying distinct sociodemographic characteristics of supersharers are based on three multilevel logistic regressions. In each of these regressions, the model distinguishes supersharers from a different reference group using the same set of sociodemographic covariates. The three reference groups are: (i) SS-NF, the most prolific sharers of non-fake political content, (ii) a random subsample of panelists of the same size as the SS-NF group, and (iii) avg. fake news sharers, the set of users who shared at least 3 political tweets linking to fake news sources, and are not supersharers. The selection of panelists to the SS-NF group uses the same definition utilized to identify supersharers (80% of shared volume) with the only difference being the kind of political news sources people link to (non-fake for the SS-NF group versus fake for the supersharers). It should be noted that the comparison groups are of different sizes, reflecting the fact that political news sharing of non-fake sources is much more common than fake news sharing. If groups were of the same size, we would be ignoring this imbalance and effectively comparing more prominent individuals in political news sharing (a large category) with individuals further down the tail of the fake news sharing distribution (a smaller category). The key benefit of our current operationalization is that it preserves individuals' relative position in the cumulative production of content. Nonetheless, differences between supersharers and the SS-NF group are robust to different specifications where individuals are matched based on their propensity to share political news, as can be seen in SM.12.

The covariates used in our models include individual-level variables describing panelists' age, gender, race/ethnicity, registration with a political party, and the number of followers on Twitter. We also included covariates that describe the geographic areas (based on FIPS codes) where panelists are registered. Specifically, we linked to panelists tract-level information about the average number of education years attained, average annual income (in USD), an indication of urban/rural classification, and an indication of residence in a Swing State. In the few cases where tract-level information was unavailable, we deferred to county-level information, then state-level information. See Section S1 for more details about the derivation of these variables. Since income and education are strongly correlated, we included education as a primary variable in the regression and income as a residual variable after accounting for the tract's level of education. For proper scaling of covariates, the residual income was standardized.

We considered additional model specifications. We examined models with interaction terms and where age was included as a continuous numeric variable instead of a categorical age bin. None of the interaction terms were found statistically significant ($P>0.05$), and therefore omitted from the models. Age as a continuous variable produced consistent results with the models specified below. We chose to include age as a categorical variable in our final models because it provides a more granular view of the differences in specific age groups.

The full regression results are presented in Table S6, where each column represents a different model distinguishing supersharers from a different reference population.

	Is supersharer?		
Relative to	SS-NF	Panel	Avg. fake news sharers
Constant	-2.29*** (0.42)	-6.35*** (0.47)	-4.97*** (0.42)
Age (30-44)	0.91*** (0.22)	2.04*** (0.22)	0.92*** (0.22)
Age (45-65)	1.79*** (0.21)	3.75*** (0.21)	1.50*** (0.21)
Age (Over 65)	2.16*** (0.21)	4.84*** (0.22)	1.86*** (0.21)
Female	0.15** (0.05)	0.23*** (0.06)	0.3*** (0.05)
Gender Unknown	0.30 (0.26)	-0.20 (0.27)	0.16 (0.24)
White	0.47*** (0.11)	0.62*** (0.12)	0.18 (0.11)
Race N/A	0.38 (0.20)	0.54 (0.24)	0.47* (0.20)
Republican	2.72*** (0.09)	2.14*** (0.10)	1.47*** (0.09)
Independent	1.39*** (0.11)	1.07*** (0.12)	1.06*** (0.11)
Other party	1.55*** (0.21)	1.73*** (0.26)	1.09*** (0.21)
Party N/A	1.63*** (0.09)	1.29*** (0.10)	1.06*** (0.09)
Avg education	-0.27*** (0.03)	-0.32*** (0.03)	-0.11*** (0.03)
Standarized residual income	0.16*** (0.03)	0.07 (0.03)	0.13*** (0.03)
Location urban	-0.34*** (0.09)	0.10 (0.11)	-0.08 (0.09)
Location N/A	2.60 (2.48)	2.12 (1.47)	2.08 (1.34)
Swing state	0.10 (0.05)	0.22*** (0.07)	-0.02 (0.05)
Log(1 + # followers)	0.31*** (0.03)	1.45*** (0.04)	0.74*** (0.03)
Observations	13,306	13,306	12,571
Log-likelihood	-4,788.29	-3,513.38	-5,123.30
Akaike IC	9,612.58	7,062.76	10,282.60
Note:	*p<0.05; **p<0.01; ***p<0.001		

Table S6.
Logistic regression models comparing supersharers to the heaviest sharers of non-fake political news (SS-NF), a random subsample of panelists (Panel), and the "average" fake news sharers.

S.5 Robustness Checks for Supersharer Definition

While our operationalization of supersharers follows a standard Pareto-principle by focusing on the individuals that account for 80% of the political news linking to fake news sources, it is important to consider the sensitivity of results to other operationalizations. Specifically, we assessed the robustness of our results in two different ways.

First, we varied the threshold from 75% to 95% in increments of one percent and fitted the main logistic regression model comparing supersharers to a random sample of panelists. We chose this range to balance statistical power on one end and avoid mixing with the average fake news sharers group on the other end. For brevity, Table S7 shows a subset of the thresholds tested, where changes in significance levels were observed. Across the entire 75-95% range, the regression coefficients remained with the same significance levels, sign, and magnitude (approximately). The only exception was the gender coefficient that was no longer significant at 89% and even changed sign towards the 95% threshold, when including more sharers of fake news from the tail of the distribution. This highlights that the over-representation of females in the supersharers group is a unique characteristic of the very top of the fake news sharing distribution that weakens as we include more "average" fake news sharers in this group.

Second, we examined whether there is a threshold when supersharers are no longer distinguishable from "average" fake news sharers. To that end, we divided the avg. fake news sharers group into deciles by the volume of fake news they are sharing. We fitted the same logistic regression model specified in Section S4 to distinguish supersharers from each decile of the avg. fake news sharers separately. Table S8 reports these regression results for the first five deciles of the avg. fake news sharers group, moving from the most prolific users (10th decile) after the supersharers to the less prolific users (6th decile). The 1st to 5th declines are omitted because they were largely the same as the 6th decile.

As one can see in Table S8, supersharers are statistically different from the avg. fake news sharers across the different deciles and along all variables reported in the main text (age, gender, party registration, and education). Supersharers have a significantly higher proportion of older adults, females, registered Republicans, and individuals from tracts with lower education attainment than most deciles of the average fake news sharers. Table S8 further substantiates that race does not significantly distinguish supersharers from any of the deciles of avg. fake news sharers. Supersharers have a significantly higher proportion of individuals of White ethnicity than both SS-NF and the Panel (as shown in Table S6), but this characteristic is not unique to the supersharers as it is common among sharers of fake news more generally.

In summary, the two sensitivity analyses reported in this section show the robustness of our results across a wide range of thresholds for considering individuals as supersharers. The over-representation of females is a unique property of the supersharers group that is particularly strong at the top of the fake news sharing distribution. White ethnicity does not significantly distinguish supersharers from the avg. fake news sharers, although it did significantly distinguish supersharers from both the panel and SS-NF groups.

	Supersharers relative to Panel				
TH	0.75	0.85	0.88	0.89	0.95
Constant	-6.81***(0.55)	-5.04*** (0.4)	-4.40*** (0.36)	-4.03*** (0.35)	-2.59*** (0.30)
Age (30-44)	2.16***(0.29)	1.93*** (0.16)	1.84*** (0.14)	1.81*** (0.13)	1.41*** (0.07)
Age (45-65)	4.09***(0.28)	3.50*** (0.16)	3.34*** (0.13)	3.28*** (0.13)	2.72*** (0.07)
Age (Over 65)	5.15***(0.29)	4.51*** (0.17)	4.34*** (0.14)	4.23*** (0.14)	3.64*** (0.09)
Female	0.28***(0.07)	0.19*** (0.05)	0.10* (0.05)	0.10 (0.05)	-0.12** (0.04)
Gender Unknow	0.23 (0.29)	-0.08 (0.24)	-0.23 (0.22)	-0.25 (0.22)	-0.23 (0.15)
White	0.63***(0.14)	0.49*** (0.10)	0.50*** (0.09)	0.49*** (0.09)	0.51*** (0.06)
Race N/A	0.28 (0.29)	0.24 (0.21)	0.17 (0.19)	0.18 (0.19)	0.13 (0.06)
Republican	2.22***(0.12)	2.04*** (0.09)	1.90*** (0.08)	1.90*** (0.08)	1.47*** (0.06)
Independent	1.13***(0.14)	0.87*** (0.10)	0.75*** (0.09)	0.77*** (0.09)	0.41*** (0.07)
Other party	1.91***(0.28)	1.36*** (0.23)	1.11*** (0.21)	1.12*** (0.20)	0.96*** (0.15)
Party N/A	1.32***(0.11)	1.18*** (0.08)	1.05*** (0.07)	1.04*** (0.07)	0.73*** (0.05)
Avg education	-0.34***(0.04)	-0.33*** (0.03)	-0.31*** (0.02)	-0.32*** (0.02)	-0.27*** (0.02)
Standarized residual income	0.03 (0.04)	0.05 (0.03)	0.05 (0.03)	0.05* (0.02)	0.01 (0.02)
Location urban	0.01 (0.12)	0.04 (0.10)	0.02 (0.09)	0.02 (0.09)	-0.00 (0.06)
Location N/A	2.47 (1.51)	1.57 (1.45)	1.22 (1.43)	1.11 (1.43)	1.48. (1.17)
Swing state	0.31***(0.07)	0.24*** (0.06)	0.23*** (0.05)	0.24*** (0.05)	0.22*** (0.04)
Log(1 + # followers)	1.54***(0.05)	1.37*** (0.05)	1.28*** (0.03)	1.26*** (0.03)	1.06*** (0.03)
Observations	12,775	14,130	14,896	15,227.00	19,239.00
Log-likelihood	-2,803.65	-4,553.62	-5,432.78	-5,794.72	-9,491.94
Akaike IC	5,643.31	9,143.23	10,901.55	11,625.42	19,019.89
Note:	*p<0.05; **p<0.01; ***p<0.001				

Table S7.

Logistic regression models comparing supersharers to a random sample of panelists, where each model represents a different volume threshold for the definition of supersharers. Select thresholds from the range of 75-95% are shown. See SM.4 for more details about the covariates used in these models.

Supersharers relative to avg. fake news sharers deciles					
Decile	10	9	8	7	6
Constant	-1.48* (0.65)	-1.69** (0.63)	-1.93** (0.65)	-2.85***(0.65)	-1.99** (0.64)
Age (30-44)	0.32 (0.31)	0.39 (0.29)	0.62* (0.29)	0.83** (0.27)	1.15***(0.27)
Age (45-65)	0.73* (0.29)	0.81** (0.29)	0.94***(0.29)	1.25***(0.27)	1.68***(0.26)
Age (Over 65)	0.84** (0.30)	0.99***(0.29)	1.14***(0.29)	1.51***(0.28)	2.11***(0.27)
Female	0.02 (0.08)	0.29***(0.08)	0.18* (0.08)	0.27** (0.08)	0.31***(0.08)
Gender Unknow	0.39 (0.42)	1.13* (0.52)	-0.46 (0.33)	0.21 (0.41)	0.09 (0.38)
White	0.23 (0.15)	0.08 (0.16)	0.18 (0.16)	0.02 (0.16)	-0.02 (0.17)
Race N/A	0.75* (0.35)	0.54 (0.34)	0.63. (0.35)	0.37 (0.35)	0.05 (0.34)
Republican	0.69***(0.13)	0.97***(0.13)	1.23***(0.13)	1.47***(0.13)	1.53***(0.13)
Independent	0.77***(0.17)	0.68***(0.16)	1.21***(0.17)	1.16***(0.16)	1.02***(0.16)
Other party	1.19** (0.39)	0.80* (0.34)	0.59. (0.31)	0.91** (0.31)	1.81***(0.41)
Party N/A	0.55***(0.13)	0.68***(0.13)	0.92***(0.12)	1.11***(0.12)	1.03***(0.13)
Avg education	-0.02 (0.04)	-0.05 (0.04)	-0.09* (0.04)	-0.06 (0.04)	-0.14** (0.04)
Standarized residual income	0.06 (0.04)	0.05 (0.04)	0.12** (0.04)	0.16***(0.04)	0.13** (0.04)
Location urban	-0.26. (0.14)	0.04 (0.15)	0.19 (0.13)	-0.13 (0.14)	-0.09 (0.15)
Location N/A	10.96 (196.96)	11.01 (196.97)	0.57 (1.56)	11.06 (196.97)	10.83 (196.97)
Swing state	-0.03 (0.08)	0.00 (0.08)	-0.11 (0.08)	0.05 (0.09)	-0.09 (0.08)
Log(1 + # followers)	0.42***(0.05)	0.44***(0.05)	0.57***(0.05)	0.71***(0.05)	0.66***(0.05)
Observations	3,154	3,154	3,154	3154	3154
Log-likelihood	-1,929.27	-1,903.61	-1,864.76	-1799.92	-1772.03
Akaike IC	3,894.54	3,843.22	3,765.52	3635.85	3580.06
Note:	*p<0.05; **p<0.01; ***p<0.001				

Table S8.
Regression coefficients comparing supersharers with each decile of the avg. fake news sharers group, moving from the most prolific (10th decile) to the less prolific users. See SM.4 for more details about the covariates used in these models.

S.6 Evaluating Use of Automation

In this section, we evaluate the use of automation by our supersharers. We describe how we determined that bot detection tools are not well-calibrated for our sample of supersharers, which motivated our use of more direct measures for quantifying the use of automation. Finally, we show that our main findings hold even if a small fraction of fully-automated supersharer accounts are still present in the sample.

To evaluate how effective bot detection tools were at detecting bots in our sample, we used version 4 of Botometer *(20)*, a state-of-the-art bot detection tool. We took a random sample of 100 accounts from each comparison group with a high Botometer score of 0.7 or higher, and blindly coded whether the tweets posted by the account seemed automated as indicated by having a large amount of content posted in a short period of time, no genuine user content, an extremely narrow focus (e.g., advertising a commercial entity), or completely incoherent posts. If bot scores were well-calibrated, a score threshold of 0.7 should have resulted in a sample with at least 70% of accounts being automated. In practice, we found that only 7.1% (2.0%, 12.2%) of supersharer accounts seemed potentially automated. The SS-NF group had a similar estimate of 4.0% (1.0%, 8.1%), which was not statistically different from the supersharers groups (P=0.35). The panel subsample did have a significantly lower percentage of bots 1.2% (0%, 3.5%; P<0.001), but still far from the 70% projected by Botometer. Several reasons could justify the mismatching scores, from the unique construction of the panel using official voter records, to the selection of supersharers based on their activity. Determining the reason for this discrepancy is an interesting avenue for future research. For our purposes, a viable upper bound for the percentage of supersharers accounts that are fully automated is 7.1% (2.0, 12.2).

Temporal Analysis and Automation Apps: To investigate whether supersharers' activity was automated, we conducted a temporal analysis of posting times and examined the apps used to post to Twitter. In particular, we analyzed the time of the day when users post, the time interval between posts, the length of posting sessions, and the average number of posting sessions each day. We considered a session as a sequence of posts that appeared within 45 minutes of each other. The results of this analysis are shown in Figure S3 for the panel random subsample (grey), the group of supersharers (purple), and the SS-NF group (yellow). Panel A shows the percentage of posts at each hour of the day in local time. Panels B-D, respectively show histograms time intervals between consecutive posts, sessions' length, and the average number of sessions each day.

If supersharers' activity was largely automated, one would expect to observe a deviation in their temporal patterns from other groups *(21)*. However, Figure S3 shows that there is neither an abnormal clustering of posting times nor a substantial difference from the distribution of other groups. Supersharers do seem slightly more active on the platform as indicated by shorter intervals between posts (Panel B), longer posting sessions (Panel C), and more sessions each day (Panel D). Yet, none of these differences seem sufficiently large to support the conclusion that their activity comes from a fundamentally different generating process with extensive use of automation tools.

Moreover, investigating the apps used to post on Twitter as indicated in tweets' metadata showed that supersharers use automation apps *less* than other groups. We manually examined the apps used by panelists and evaluated whether it has a feature of scheduling or automatically tweeting on behalf of the user. For example, apps that we labeled as enabling auto-tweeting include TweetDeck, TwinyBots, and MeetEdgar. As shown in Table S10, we found that a similar percentage of SS-NF and panel users used automation apps (8.9% for SS-NF and 6.8% for the panel; P=0.08), supersharers use automation apps significantly less than the SS-NF, but not compared to the panel (3.6%; P<0.05).

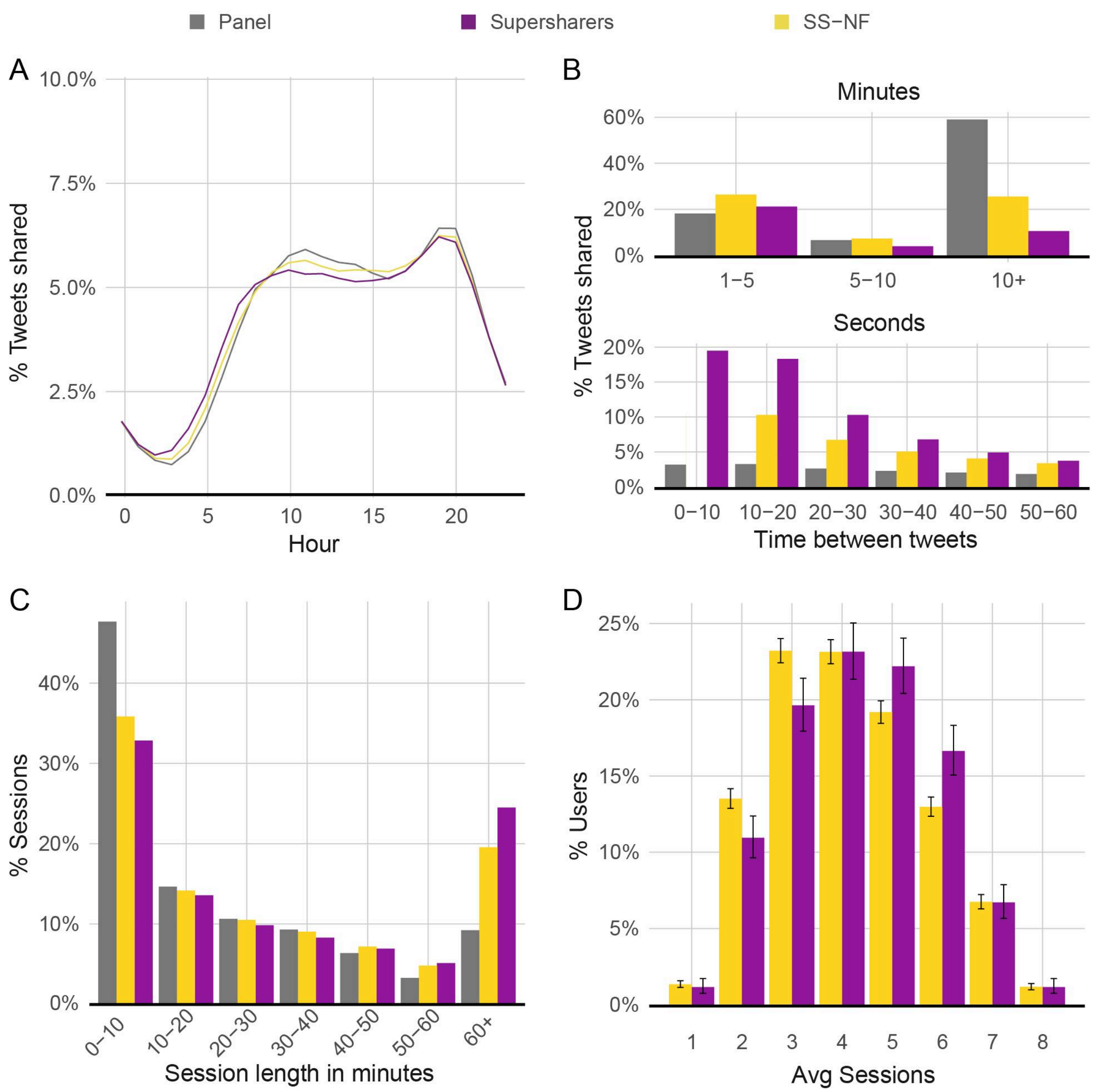

Figure S3.
(A) Percent of tweets shared by the hour of the day; (B) Percent of tweets shared within minutes (top) or seconds (bottom) of a previous post; (C) Percent sessions by length in minutes; and (D) Percent users by the amount of posting sessions per day. Error bars were omitted when estimates were smaller than 0.1%. Panel D omits the random subsample of panelists, where nearly 80% of users have no more than one posting session a day, which obscured the more interesting comparison between the supersharers and the SS-NF groups.

Bootstrapping Regression Estimates: We evaluated the robustness of our main findings by bootstrapping the regression coefficients with 10,000 replicas sampled with replacement. By bootstrapping with replacement, we vary the percentage of potentially-automated accounts in each replica, which helps assess the sensitivity of results to varying levels of potentially-automated accounts. We use the same model specification used in the main text and described in SM.4. Table S9 reports the bootstrapped model estimates for the model comparing supersharers to the SS-NF group along with the estimates from the main model (without bootstrapping, as reported in Table S6). As one can see from the table, the estimates are nearly identical with no difference larger than 0.01.

Additional Differences in Posting: The absence of substantial differences in the use of automation led us to investigate additional differences in the way supersharers tweet. We examined the percentage of different types of tweets, i.e., original tweets, retweets, quotes, and replies, posted by individuals in each of the groups. As shown in Table S10, we found that supersharers retweet at a significantly higher rate than all other groups ($P<0.001$). Retweets were 74.7% (73.7%, 75.6%) of the tweets posted by the average supersharer. For comparison, 32.7% (32.0%, 33.3%) of tweets by the average panelist are retweets, this percentage is 59.9% (59.4%, 60.4%) for the SS-NF group. Supersharers also had the lowest percentage of original tweets of all groups with an average of 6.26% (5.66%, 6.86%) of original tweets [compared to 10.34% (10.04%, 10.65%) for the SS-NF group and 26.70% (26.06%, 27.34%) for the panel]. Compared to other types of tweets, retweets are a relatively effortless way of generating a large quantity of content rapidly, which could explain the slightly shorter time intervals observed for supersharers between consecutive tweets.

Is supersharer?		
Relative to	Bootstraped bots estimates	SS-NF
Constant	-2.67*** (0.42)	-2.68*** (0.42)
Age (30-44)	0.91*** (0.23)	0.91*** (0.22)
Age (45-65)	1.73*** (0.22)	1.74*** (0.21)
Age (Over 65)	2.11*** (0.22)	2.11*** (0.21)
Female	0.16** (0.05)	0.16** (0.05)
Gender Unknow	0.34 (0.25)	0.34 (0.26)
White	0.44*** (0.11)	0.43*** (0.11)
Race N/A	0.37 (0.20)	0.37 (0.20)
Republican	2.64*** (0.09)	2.64*** (0.09)
Independent	1.36*** (0.11)	1.36*** (0.11)
Other party	1.48*** (0.22)	1.48*** (0.21)
Party N/A	1.58*** (0.09)	1.58*** (0.09)
Avg education	-0.25*** (0.03)	-0.25*** (0.03)
Standarized residual income	0.15*** (0.03)	0.15*** (0.03)
Location urban	-0.34*** (0.10)	-0.34*** (0.09)
Swing state	0.10 (0.06)	0.10 (0.06)
Log(1 + # followers)	0.45*** (0.04)	0.45*** (0.04)
Observations	13,306	13,306
Note:	*p<0.05; **p<0.01; ***p<0.001	

Table S9.
Bootstrapped estimates from 10,000 samples with replacement of the logistic regression model comparing supersharers to the heaviest sharers of non-fake political news (SS-NF). The column on the left shows the bootstrapped estimates and std. error, while the column on the right shows the estimates of the main model (from Table S6) for ease of comparison. See SM.4 for more details about the covariates used in these models.

	Panel (N=11,199)	SS-NF (N=11,199)	Supersharers (N=2,107)
% Bots (upper bound)	1.2 (0, 3.5)	4.0* (1.0, 8.1)	7.1* (2.0, 12.2)
% Users using automation app	4.11 (4.11, 4.12)	8.74* (8.73, 8.74)	3.51 (3.50, 3.52)
Average % of retweets	32.7 (32.0, 33.3)	59.9* (59.4, 60.4)	74.7*, † (73.7, 75.6)

Table S10.
Statistics describing account automation and retweeting rate (rows) across three user groups (columns). Estimates along with 95% CIs are presented.
* significantly higher than the panel subsample (P<0.001).
† significantly higher than SS-NF (P<0.001).

S.7 Suspended Accounts

To evaluate whether Twitter took moderation actions against the accounts in our sample, we collected data from the Twitter Compliance API in December 2021 about suspended accounts in our sample. We found that the panel subsample had the lowest percentage of suspended accounts (0.1%). The SS-NF group had 3.6% of accounts suspended, and the avg. fake news sharers group had 7.1% of suspended accounts. Supersharers had the highest percentage of suspended accounts at 23%. However, further investigation revealed that 85% of the suspensions of supersharers happened in January, 2021. These findings suggest that Twitter did take considerable measures to moderate and limit the activity of supersharers, but only acted months after the election. At the time of our collection, nearly a year after the January 6th events, more than half (62.6%) of supersharer accounts were still operational and actively posting content on Twitter.

S.8 Concentration of Fake News

In this section, we compare the concentration of political fake news among panelists to other topics identified using a topic model, with the explicit goal of understanding how common it is for certain topics on Twitter to exhibit the same level of concentration observed for fake news. To do so, we used a state-of-the-art language model and clustering technique to extract prominent topics in the content shared by panelists and calculated the Gini coefficient as a measure of concentration across panelists. In particular, we used the language model of Conversational BERT[5] (*41*), a BERT-based model that was further pre-trained on social media content, to extract vector embeddings for panel tweets. We did this for a 10% random sample of panel tweets (about 12 million tweets) due to the time-consuming and expensive nature of the inference process using this large language model. Then, we used BERTopic (*42*) to identify topics, which uses UMAP dimensionality reduction *(43)* and K-Means clustering. We repeated the clustering process across 10 different folds of the data and experimented with different numbers of topics to ensure the robustness of our findings. We qualitatively examined the clusters for semantic coherence and removed clusters that were too sparse (shared by fewer than 100 panelists). Finally, we computed the Gini coefficient for each topic, where a zero coefficient represents perfect equality, i.e., all panelists sharing exactly the same number of tweets in this topic, and as the coefficient approaches one the disparities between panelists increase, i.e., a small percent of panelists are responsible for the vast majority of shared content in this topic.

Figure S4 shows the distribution of concentration levels (Gini coefficients) across topics as well as the concentration of content from fake news sources (vertical, red, dashed line). The different curves represent the distribution of Gini coefficients for topics extracted using different models (with 1000 topics in black and 300 topics in red). For example, the red curve shows that the most common concentration level in the K=300 model is about 0.25 and that very few topics exhibit a concentration level greater than 0.5. The figure also shows in green the distribution of topics appearing in roughly the same number of tweets (±10%) as the fake news tweets. Similarly, it shows in blue the distribution of topics with the same number of users (±10%) sharing the topic as the number of people sharing fake news.

The figure highlights a few findings. Across all topics and models, fake news sharing is more concentrated than 99% of topics. Fake news concentration in the population is higher than topics of similar volume and topics shared by a similar number of users (±10%). Overall, these

[5] Available at https://huggingface.co/DeepPavlov/bert-base-cased-conversational

findings show that the level of concentration of fake news is not typical on Twitter, and it cannot be simply explained by the "size" of the topic.

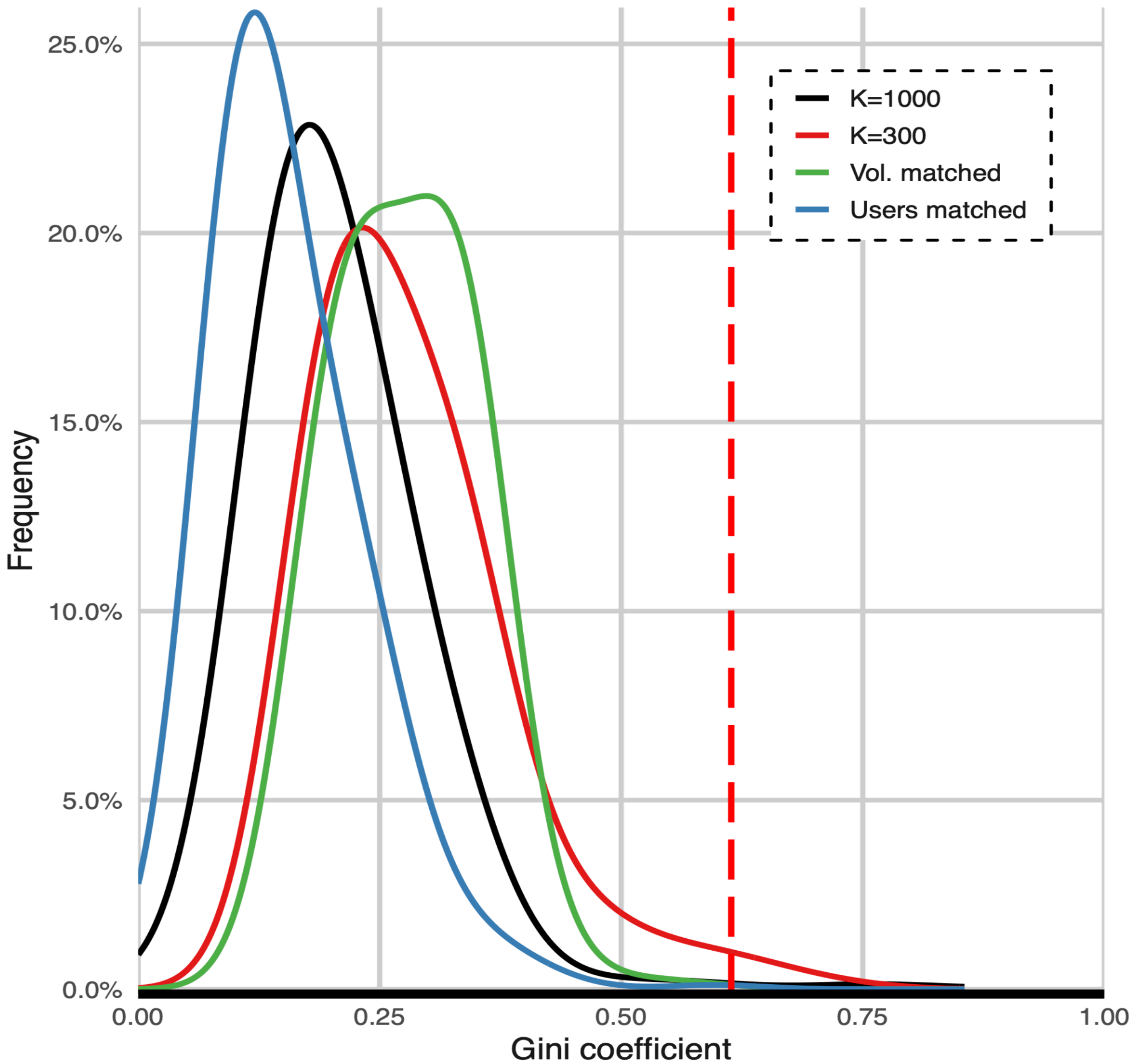

Figure S4.
Concentration across panel members of content from fake news sources (vertical, red, dashed line) compared to the concentration of other topics. Topics were extracted by clustering Conversational BERT embeddings for a 10% random sample of tweets shared by panelists, and measuring the Gini coefficient for each topic. Black and red curves show the distribution of Gini coefficients for models with 1000 and 300 topics, respectively. Blue and green curves show the distribution of Gini coefficients for the subset of topics with similar (±10%) number of tweets or number of users sharing, respectively, as observed for political news linking to fake news sources.

S.9 Platform Interventions

In this section, we evaluate the cost and benefit of possible platform interventions. In particular, we evaluate the impacts of banning/suspending users and enforcement of different limits on retweet rates.

The main paper reported that a suspension of supersharers would result in a reduction of 24.4% of the fake news available to their average follower. However, that requires identifying supersharers ahead of time. To get a more accurate estimate of a policy social media platforms could implement in practice, we used the first month of the study (August) to identify supersharers and the following months (September-November) to evaluate the impact of user suspension. Following the same definition used throughout the research, August supersharers were identified as the top users who shared 80% of the political news linking to fake news sources by the entire panel. We have identified 1,639 users as supersharers using this approach, most of them (1,421) were in the final set of 2,107 supersharers used throughout the work. If Twitter were to suspend those 1,639 users in August, in the subsequent election months of Sep-Nov it would have removed 66.82% of the fake news shared by registered U.S. voters on the platform. Of course, the same policy could be applied in subsequent months to identify more supersharers, which would increase these estimates. However, suspensions do not guarantee that a user would not find other ways to circumvent the suspension, e.g., by posting from other accounts.

The second intervention policy we consider involves limiting retweets to a daily limit. For our purposes, we do not distinguish between political and non-political tweets or tweets with and without links. To decide on meaningful thresholds, we examined the distribution of the maximal number of retweets per day shared by individuals in each of the groups. We found the threshold of 50 retweets per day to separate well panelists from the supersharer and SS-NF groups, and 100 retweets per day to separate well supersharers from the two other groups. We also observed that hardly any user in the panel or SS-NF groups retweeted more than 200 times a day. For each of the three daily retweet limits (50, 100, and 200), we calculated how many tweets would exceed the limit, on how many days, for how many users.

Table S11 provides statistics about the impact of limiting retweets at three different levels. It shows the number of affected users, the number of days users would have been impacted, and the percentages of fake news and non-fake news tweets that would have been impacted. For example, limiting users to 50 retweets per day would affect 87.7% of supersharers, 47.7% of people in the SS-NF groups, and less than 1% of panelists. It would have impacted 42.4% of the fake news shared by supersharers, but also 45.6% of the non-fake content they share. Supersharers would have faced this limit on 36.1 days on average, while members of the SS-NF group would have encountered it only on 9.3 days, on average.

A general tradeoff emerges from Table S11 between the specificity of the intervention and a reduction in fake news. The 50 retweets limit would impact fake news the most, but would also impact many users in the SS-NF group. In contrast, the 200 retweets limit would impact a lower fraction of fake news (12.5%) but could target supersharers more specifically. Across the different thresholds, it is notable that a considerable amount of non-fake content would also be affected by those interventions. Of course, other interventions could be considered, e.g., ones that are limited to politics or the resharing of links, which may reduce the impact on non-fake content. A more gentle variant of a hard retweet limit can alert the user to the fact they had shared many tweets over a short period and invite them to post more at a later point in time. It is possible that such a slowdown can have positive impacts both on the individual involved and their network. However, interventions can also have negative impacts on users, potentially creating a feeling of censorship or unfair treatment.

	Retweet limits per day								
	50 retweets			100 retweets			200 retweets		
	SS	SS-NF	P	SS	SS-NF	P	SS	SS-NF	P
% of affected users	87.7	47.7	0.9	65.0	20.9	0.2	34.7	7.7	0.01
# of days facing limit (out of 122 days in sample) per user	36.1	9.3	0.06	19.4	3.5	0.008	8.35	0.03	10^{-4}
% fake news tweets removed	42.4	10.4	1.1	26.6	4.1	0.0	12.5	1.1	0.0
% non-fake news tweets removed	45.6	18.0	0.8	30.0	9.1	0.1	14.7	3.3	0.0

Table S11.
The impact of imposing retweet limits on Twitter per day at three different levels (in columns). Rows contain statistics on the number of affected users, number of days being impacted, and percentages of tweets with or without links to fake news sources being removed. Each cell contains statistics for the supersharers (SS), SS-NF, and panel subsample (P) groups.

S.10 Robustness Checks for Exposure Thresholds

In this section, we report additional findings using different thresholds for estimating the proportions of available content to panelists, and different thresholds for identifying the heaviest consumers of fake news in the panel.

A minimal threshold for the amount of political news coming from one's network is needed to get a reliable estimate of the proportion of fake news that people get from their network. The main text reports our estimates using a threshold of 100 political news for statistical purposes. Here, we report findings from using a lower threshold of 50 political news.

Using the threshold of 50 political news, we found that none of the estimates was meaningfully affected. The average panelist received 1.60% (compared to 1.66%) of political news linking to fake news sources from their network and the average follower of a supersharer had 4.12% (compared to 4.11% before). Supersharers provided 24.4% of the fake news to their followers (compared to 24.6% using the higher threshold). Finally, supersharers were the only provider of fake news to 11.3% of their followers, exactly the same percentage when using the 100 threshold.

Different thresholds can be drawn for considering the heaviest consumers of fake news in the panel. Table S12 reports statistics for three different levels: the top 20 percent of users (80th percentile), the top 10 percent (90th percentile), and the top five percent (95th percentile). For each subgroup, we report the minimal level of fake news availability that corresponds to the threshold, the number of users in the set, and the percentage of users in the set following a supersharer. For example, users in the 90th percentile of fake news availability had at least one in every 20 political news coming from their network linking to fake news sources. In other words, users in the top decile had at least 5% of political news linking to fake news from their network. There are 51.2 thousand individuals included in this top decile and 22.1 of them follow a supersharer. Across the different percentiles, we observe that roughly 20% of users follow a supersharer.

Percentile of fake news from peers	**80**	**90**	**95**
Minimal individual % of fake news from network	1.7	5.0	9.1
# Heavy consumers (thousands)	103.7	51.2	21.4
% of Heavy consumers following a supersharer	19.4	22.1	16.4

Table S12.
The heaviest consumers of fake news at three different levels (columns). For each percentile of the heaviest consumer, the table includes the corresponding (minimal) proportion of fake news individuals get from the network, the size of the group, and the percentage of users following a supersharer.

S.11 Network Influence and Engagement

To further understand the influence of users in their local network, we computed Pei et al.'s measure of network influence *(18)*. In their work, they showed that in the absence of full network information and across different networks, a local proxy can be an effective measure of a user's network influence. That local proxy was calculated as the sum of the nearest neighbors' degrees. In a directed graph like Twitter, this would be the sum of follower counts for followers of a focal user. Intuitively, this is similar to second-degree reach, except that some nodes can be counted multiple times. We calculated this measure of network influence for all panel members for whom we have complete network information.

Using this measure of network influence, we found that supersharers had a significantly higher network influence than the SS-NF group ($P<0.001$), and that both groups had a significantly higher influence than the panel ($P<0.001$). Table S13 reports the median value for individuals in each of the comparison groups. It shows that the median supersharer had 380,306 followers of followers, which is 29% higher than the median in the SS-NF group and 42 times larger than the network influence of the median user in the panel. To put these numbers on a relative scale, we calculated the percentile in the full panel for each of these medians, after random down-sampling of the larger groups to account for group size. We found that the median supersharers belong in the 86th percentile, the median SS-NF user in the 84th percentile, and the median of the panel subsample at the 51st percentile.

Next, we examined people's engagement with content from supersharers and the other reference groups. We considered engagement as any reply, quote, or retweet to a user in one of the comparison groups. We did not include in this analysis users who did not post political news or had no followers in the panel. The number of users who passed these criteria in each group is specified in Table S13. For each of these users, we computed two measures: (i) a binary indication of whether they had any engagement with their content, and (ii) the fraction of panelists who engaged with their content relative to their number of followers in the panel. The two separate measures help distinguish between users that had any engagement in our sample and the number of panelists engaging with the content per individual.

Using the above measures of engagement, we found that the supersharers group had a significantly higher percentage of people with engagement to their content relative to the panel ($P<0.001$), but no significant difference from the SS-NF group. Table S13 shows these differences explicitly: 8.6% of users in the panel subsample had any engagement with their content in our sample compared with about 50% in the SS-NF and supersharer groups. Among users with an engagement, supersharers had significantly more panelists engaging with their content compared to the panel subsample and SS-NF groups ($P<0.001$). Table S13 shows the average fraction (percent) of panelists engaging with a user's content relative to the number of followers in the panel. For supersharers, that percentage stood at 24.1%, which is higher by 8.9% than the average in the SS-NF group, and higher by 14% than the average in the panel subsample. Overall, these results show that more supersharers had an engagement with their content, and that more panelists engaged with each supersharer.

	Panel (N=11,199)	**SS-NF** (N=11,199)	**Supersharers** (N=2,107)
Median Network Influence (# of followers of followers)	9,034	295,159*	380,306*, †
# of users who posted political news and have panel followers	2,905	10,354	1,752
% of above users with engagement	8.6	49.8*	50.7*
% of panelists engaging relative to followers (with 95% CIs)	10.1 (7.7, 12.5)	15.2* (13.8, 16.6)	24.1*, † (21.1, 27.0)

Table S13.
Network influence and engagement for the panel, SS-NF group, and supersharers.
* significantly higher than the panel subsample (P<0.001).
† significantly higher than SS-NF (P<0.001).

S.12 Matched Samples

In this section, we use matched samples to examine two alternative explanations for the demographic findings reported in the main text.

First, we examined whether party identification explains the demographic findings. If this were the case, one would expect that none of the significant differences reported in the main text would hold for individuals of the same party. To test this alternative explanation, we conducted additional regression analysis that compared supersharers to the SS-NF group separately for registered Democrats and for registered Republicans. The specifications for the logistic regression model in both cases were identical to the ones described in SM.4, except for the covariate describing the registration party, which was no longer needed.

Table S14 reports the coefficients for the models distinguishing supersharers from the SS-NF groups, fitted separately for registered Democrats and for registered Republicans. For registered Republicans, the results show that all of the demographic differences reported in the main papers remain unchanged with the only exception being that the small effect found for residual income was no longer statistically significant in this subset of users. Among registered Democrats, age, and education differences remained the same, but gender and race were no longer significant. The fact that race differences are attributable to party identification was already noted in the main text and affirmed in SM.5 through the comparison with the Republican-leaning average fake news sharers group. Therefore, we conclude that the key demographic differences reported in the main text hold separately for registered Republicans, and to a lesser extent for registered Democrats with a notable difference in women over-representation.

The second alternative explanation we considered was that the demographic differences stem from differences in propensities to share political news. To test this, we implemented a propensity score model using the MatchIt package *(44)* to compare a matched sample of supersharers with the SS-NF group. In particular, the propensity model used a multilevel linear regression model with all the covariates of the full model (described in SM.4) and a dependent variable which is the number of political news shared (in logarithmic scale). We used a propensity score to match individuals in the two groups based on their propensity score distance. After matching, the comparison groups were indistinguishable in their level of political sharing, effectively controlling for the alternative explanation that the tendency to share political news leads users to share fake news.

The results of the matched analysis are presented in Table S15. The results show that the key demographic differences reported in the main paper (age, gender, registration party, education, residual income) remained the same when matching individuals based on their propensity to share political news. The only covariate that did change is race, which was no longer significant. Further examination revealed that race differences vanished due to the over-representation of registered Democrats in the matched sample and the under-representation of registered Republicans. Hence, race differences are, again, found to be attributable to party identification, a fact that was already noted in the main text and SM.5. Overall, this analysis affirms that the demographic differences reported in the main text hold even when controlling for individuals' propensity to share political news.

	Is supersharer?	
Relative to	SS-NF Democrats	SS-NF Republicans
Constant	-0.75 (1.07)	-1.36 (0.8)
Age (30-44)	0.37 (0.63)	0.54 (0.36)
Age (45-65)	1.35* (0.59)	1.33*** (0.34)
Age (Over 65)	2.18*** (0.60)	1.66*** (0.35)
Female	-0.07 (0.15)	0.21* (0.10)
Gender Unknown	-14.11 (449.38)	1.01* (0.47)
White	-0.06 (0.24)	0.16 (0.20)
Race N/A	0.38 (0.20)	0.54 (0.24)
Avg education	-0.33*** (0.06)	-0.11* (0.05)
Standarized residual income	0.14 (0.08)	0.05 (0.05)
Location urban	-0.22 (0.33)	-0.24 (0.20)
Location N/A	-12.91 (3956.18)	13.02 (324.74)
Swing state	0.06 (0.19)	0.01 (0.10)
Log(1 + # followers)	0.37*** (0.10)	0.41*** (0.06)
Observations	4,755	1,936
Log-likelihood	-752.90	-1,243.25
Akaike IC	1,533.80	2,514.51
Note:	*p<0.05; **p<0.01; ***p<0.001	

Table S14.
Logistic regression models comparing supersharers to the SS-NF group, separately for registered Democrats and for registered Republicans.

	Is supersharer?
Relative to	SS-NF matched
Constant	-2.61*** (0.41)
Age (30-44)	0.43* (0.21)
Age (45-65)	1.46*** (0.19)
Age (Over 65)	2.08*** (0.20)
Female	0.26*** (0.06)
Gender Unknown	0.45 (0.25)
White	0.17 (0.10)
Race N/A	-0.01 (0.21)
Republican	2.39*** (0.09)
Independent	1.18*** (0.10)
Other party	1.37*** (0.21)
Party N/A	1.38*** (0.08)
Avg education	-0.28*** (0.03)
Standarized residual income	0.12*** (0.03)
Location urban	-0.40*** (0.09)
Location N/A	1.91 (3.42)
Swing state	0.10 (0.06)
Log(1 + # followers)	0.73*** (0.04)
Observations	13,306
Log-likelihood	-4,308.23
Akaike IC	8,652.47
Note:	*p<0.05; **p<0.01; ***p<0.001

Table S15.
Logistic regression model comparing supersharers to the SS-NF group in a matched sample based on individuals' propensity to share political news.

S.13 Geographic Distribution

In order to examine geographic differences, we compared supersharers and the average fake news sharers group with the full panel. Figure S5 shows the difference between the percent of supersharers and the panel in each state (left), and the difference between the avg. fake news sharers and the panel (right). Florida, Arizona, and Texas have significantly higher percentages of supersharers than their share in the panel sample (over-represented by a factor of 1.3-2.1; P < 0.05). Average sharers of fake news are also over-represented in Florida and Arizona but not in Texas, which puts Texas as the only state having a higher concentration of supersharers relative to more moderate fake news sharers.

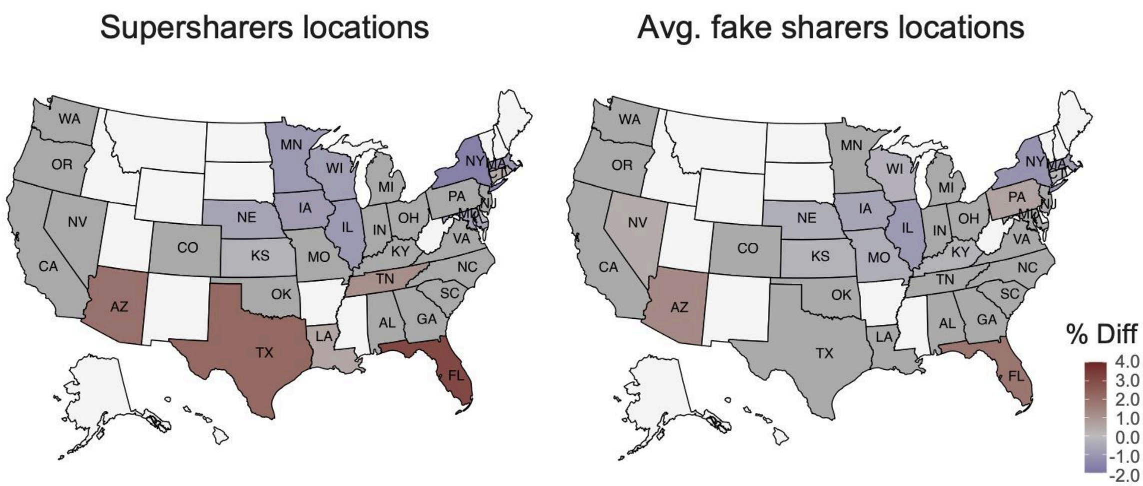

Figure S5. Concentration of supersharers (left) and avg. fake sharers (right) in U.S. states, compared to the panel. Color strength represents the magnitude of the difference with states shaded in red (purple) to represent over-representation (under-representation) compared to the panel. White represents states where our sample is statistically underpowered, and gray designates states where there is no statistically significant difference from the panel.

References and Notes

1. C. A. Bail, B. Guay, E. Maloney, A. Combs, D. S. Hillygus, F. Merhout, D. Freelon, A. Volfovsky, Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 243–250 (2020). doi:10.1073/pnas.1906420116 Medline

2. G. Eady, T. Paskhalis, J. Zilinsky, R. Bonneau, J. Nagler, J. A. Tucker, Exposure to the Russian Internet Research Agency foreign influence campaign on Twitter in the 2016 US election and its relationship to attitudes and voting behavior. *Nat. Commun.* **14**, 62 (2023). doi:10.1038/s41467-022-35576-9 Medline

3. C. Shao, G. L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, F. Menczer, The spread of low-credibility content by social bots. *Nat. Commun.* **9**, 4787 (2018). doi:10.1038/s41467-018-06930-7 Medline

4. 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). doi:10.1126/science.aau2706 Medline

5. A. Guess, B. Nyhan, J. Reifler, "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 U.S. Presidential Campaign" (European Research Council, Working Paper, 2018); https://about.fb.com/wp-content/uploads/2018/01/fake-news-2016.pdf.

6. A. Guess, J. Nagler, J. Tucker, Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Sci. Adv.* **5**, eaau4586 (2019). doi:10.1126/sciadv.aau4586 Medline

7. J. Habermas, *The Theory of Communicative Action, Volume 2: Lifeworld and System: A Critique of Functionalist Reason* (Beacon Press, 1984).

8. M. E. Roberts, *Censored: Distraction and Diversion Inside China's Great Firewall* (Princeton Univ. Press, 2018).

9. C. Lampe, N. Ellison, C. Steinfield, in *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work* (Association for Computing Machinery, 2006), pp. 167–170.

10. L. K. Fazio, D. G. Rand, G. Pennycook, Repetition increases perceived truth equally for plausible and implausible statements. *Psychon. Bull. Rev.* **26**, 1705–1710 (2019). doi:10.3758/s13423-019-01651-4 Medline

11. R. Dahlke, J. Hancock, "The effect of online misinformation exposure on false election beliefs," OSF Preprints (2022); https://doi.org/10.31219/osf.io/325tn.

12. D. A. Effron, M. Raj, Misinformation and Morality: Encountering Fake-News Headlines Makes Them Seem Less Unethical to Publish and Share. *Psychol. Sci.* **31**, 75–87 (2020). doi:10.1177/0956797619887896 Medline

13. A. Marwick, R. Lewis, "Media Manipulation and Disinformation Online" (Data & Society Research Institute, 2017); https://datasociety.net/library/media-manipulation-and-disinfo-online/.

14. A. Badawy, A. Addawood, K. Lerman, E. Ferrara, Characterizing the 2016 Russian IRA influence campaign. *Soc. Netw. Anal. Min.* **9**, 31 (2019). doi:10.1007/s13278-019-0578-6

15. G. King, J. Pan, M. E. Roberts, How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument. *Am. Polit. Sci. Rev.* **111**, 484–501 (2017). doi:10.1017/S0003055417000144

16. F. B. Keller, D. Schoch, S. Stier, J. Yang, Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign. *Polit. Commun.* **37**, 256–280 (2020). doi:10.1080/10584609.2019.1661888

17. D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, J. L. Zittrain, The science of fake news. *Science* **359**, 1094–1096 (2018). doi:10.1126/science.aao2998 Medline

18. S. Pei, L. Muchnik, J. S. Andrade Jr., Z. Zheng, H. A. Makse, Searching for superspreaders of information in real-world social media. *Sci. Rep.* **4**, 5547 (2014). doi:10.1038/srep05547 Medline

19. S. González-Bailón, D. Lazer, P. Barberá, M. Zhang, H. Allcott, T. Brown, A. Crespo-Tenorio, D. Freelon, M. Gentzkow, A. M. Guess, S. Iyengar, Y. M. Kim, N. Malhotra, D. Moehler, B. Nyhan, J. Pan, C. V. Rivera, J. Settle, E. Thorson, R. Tromble, A. Wilkins, M. Wojcieszak, C. K. de Jonge, A. Franco, W. Mason, N. J. Stroud, J. A. Tucker, Asymmetric ideological segregation in exposure to political news on Facebook. *Science* **381**, 392–398 (2023). doi:10.1126/science.ade7138 Medline

20. K.-C. Yang, O. Varol, P.-M. Hui, F. Menczer, in *Proceedings of the AAAI Conference on Artificial Intelligence* (AAAI Press, 2020), pp. 1096–1103.

21. C. M. Zhang, V. Paxson, in *International Conference on Passive and Active Network Measurement*, N. Spring, G. F. Riley, Eds. (Springer, 2011), pp. 102–111.

22. T. N. Ridout, E. F. Fowler, M. M. Franz, Spending Fast and Furious: Political Advertising in 2020. *Forum* **18**, 465–492 (2021). doi:10.1515/for-2020-2109.

23. A. Goerres, Why are Older People More Likely to Vote? The Impact of Ageing on Electoral Turnout in Europe. *Br. J. Polit. Int. Relat.* **9**, 90–121 (2007). doi:10.1111/j.1467-856x.2006.00243.x

24. R. Igielnik, "Men and women in the U.S. continue to differ in voter turnout rate, party identification" (Pew Research Center, 2020); https://web.archive.org/web/20240408224740/https://www.pewresearch.org/short-reads/2020/08/18/men-and-women-in-the-u-s-continue-to-differ-in-voter-turnout-rate-party-identification/.

25. M. Deckman, in *Tea Party Women: Mama Grizzlies, Grassroots Leaders, and the Changing Face of the American Right* (New York Univ. Press, 2016), pp. 30–97.

26. B. G. S. Casara, C. Suitner, J. Jetten, The impact of economic inequality on conspiracy beliefs. *J. Exp. Soc. Psychol.* **98**, 104245 (2022). doi:10.1016/j.jesp.2021.104245

27. M. A. Craig, J. A. Richeson, On the precipice of a "majority-minority" America: Perceived status threat from the racial demographic shift affects White Americans' political ideology. *Psychol. Sci.* **25**, 1189–1197 (2014). doi:10.1177/0956797614527113 Medline

28. M. Dobbs, J. DeGutis, J. Morales, K. Joseph, B. Swire-Thompson, Democrats are better than Republicans at discerning true and false news but do not have better metacognitive awareness. *Community Psychol.* **1**, 46 (2023). doi:10.1038/s44271-023-00040-x

29. R. K. Garrett, R. M. Bond, Conservatives' susceptibility to political misperceptions. *Sci. Adv.* **7**, eabf1234 (2021). doi:10.1126/sciadv.abf1234 Medline

30. S. Y. Syn, S. Oh, Why do social network site users share information on Facebook and Twitter? *J. Inf. Sci.* **41**, 553–569 (2015). doi:10.1177/0165551515585717

31. L. Cava, L. M. Aiello, A. Tagarelli, Drivers of social influence in the Twitter migration to Mastodon. *Sci. Rep.* **13**, 21626 (2023). doi:10.1038/s41598-023-48200-7 Medline

32. S. Baribi-Bartov, B. Swire-Thompson, N. Grinberg, Supersharers of Fake News on Twitter, dataset, Dryad (2024); https://doi.org/10.5061/dryad.44j0zpcmq.

33. S. Shugars, A. Gitomer, S. McCabe, R. J. Gallagher, K. Joseph, N. Grinberg, L. Doroshenko, B. Foucault Welles, D. Lazer, Pandemics, Protests, and Publics. *J. Quant. Descr. Digit. Media* **1**, 1–68 (2021). doi:10.51685/jqd.2021.002

34. J. Green, W. Hobbs, S. McCabe, D. Lazer, Online engagement with 2020 election misinformation and turnout in the 2021 Georgia runoff election. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2115900119 (2022). doi:10.1073/pnas.2115900119 Medline

35. A. G. Hughes, S. D. McCabe, W. R. Hobbs, E. Remy, S. Shah, D. M. J. Lazer, Using Administrative Records and Survey Data to Construct Samples of Tweeters and Tweets. *Public Opin. Q.* **85**, 323–346 (2021). doi:10.1093/poq/nfab020

36. E. Bakshy, S. Messing, L. A. Adamic, Political science. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015). doi:10.1126/science.aaa1160 Medline

37. S. Jackman, B. Spahn, Politically Invisible in America. *PS Polit. Sci. Polit.* **54**, 623–629 (2021). doi:10.1017/S1049096521000639

38. J. Mellon, C. Prosser, Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Res. Politics* **4**, 2053168017720008 (2017). doi:10.1177/2053168017720008

39. R. E. Robertson, S. Jiang, K. Joseph, L. Friedland, D. Lazer, C. Wilson, in *Proceedings of the ACM on Human-Computer Interaction*, K. Karahalios, A. Monroy-Hernández, A. Lampinen, G. Fitzpatrick, Eds. (Association for Computing Machinery, no. 148, 2018).

40. H. Lin, J. Lasser, S. Lewandowsky, R. Cole, A. Gully, D. G. Rand, G. Pennycook, High level of correspondence across different news domain quality rating sets. *PNAS Nexus* **2**, pgad286 (2023). doi:10.1093/pnasnexus/pgad286

41. M. Burtsev, A. Seliverstov, R. Airapetyan, M. Arkhipov, D. Baymurzina, N. Bushkov, O. Gureenkova, T. Khakhulin, Y. Kuratov, D. Kuznetsov, A. Litinsky, V. Logacheva, A. Lymar, V. Malykh, M. Petrov, V. Polulyakh, L. Pugachev, A. Sorokin, M. Vikhreva, M. Zaynutdinov, in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics-System Demonstrations* (Association for Computational Linguistics, 2018), pp. 122–127.

42. M. Grootendorst, BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv:2203.05794 [cs.CL] (2022).

43. L. McInnes, J. Healy, N. Saul, L. Großberger, UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **3**, 861 (2018). doi:10.21105/joss.00861

44. D. E. Ho, K. Imai, G. King, E. A. Stuart, MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *J. Stat. Softw.* **42**, 8 (2011). doi:10.18637/jss.v042.i08