Fake News and Advertising on Social Media: A Study of the Anti-Vaccination Movement

Lesley Chiou* and Catherine Tucker†

July 6, 2018

Abstract

Online sources sometimes publish information that is false or intentionally misleading. We study the role of social networks and advertising on social networks in the dissemination of false news stories about childhood vaccines. We document that anti-vaccine Facebook groups disseminate false stories beyond the groups as well as serving as an "echo" chamber. We also find that after Facebook's ban on advertising by fake new sites, the sharing of fake news articles on Facebook fell by 75% on Facebook compared to Twitter.

*Economics Department, Occidental College, CA

†MIT Sloan School of Management, MIT, Cambridge, MA.

‡We thank Yinbo Gao, Tamara Kawash, Gyan Prayaga, and Andrea Tuemmler for excellent research assistance.

1 Introduction

The Internet has significantly changed the type of news that consumers receive. In the past consumers relied on traditional media, such as radio and television, which involved relatively fewer and more established sources of news. Nowadays consumers are exposed to online sources of information, through, for example, social networking sites, which allow any individual to share content without "fact-checking or editorial judgment" (Allcott and Gentzkow, 2017a). Many worry that online sources may publish false information, but present it as facts or "real" news. We document how anti-vaccine groups on Facebook disseminate false information to users, and we also study whether Facebook's ban on the advertising of "fake" news prevents the spread of false or misleading news stories on childhood vaccines.

While nearly 60% of adults in the US have searched for health information online in the past year (Fox and Duggan, 2013), online information for consumer health is "often unreliable" and difficult for consumers to assess (Fu et al., 2016; Fahy et al., 2014). Studies demonstrate how consumers do not accurately determine the reliability of health content on the Internet (Allam et al., 2014; Knapp et al., 2011; Kutner et al., 2006). In particular, individuals do not take into account the credibility of the content when presented with online information that is critical of vaccination (Nan and Madden, 2012; Betsch et al., 2010, 2013; Allam et al., 2014).

We focus on childhood vaccines for several reasons. First significant confusion and misleading information on the Internet surrounds the adverse effects of vaccinations on children. For example online articles allege that the vaccine for measles, mumps, and rubella causes autism even though academic studies in the medical literature have since debunked these myths. Second although the Centers for Disease Control and Prevention (CDC) recommend that individuals receive their first vaccinations during childhood, parents report concerns

about safety as among primary reasons why their children are not vaccinated (Smith et al., 2016). Finally vaccines represent an important health concern for the general public because when children receive vaccinations, they also protect the community through herd immunity by preventing further spread of the disease to those individuals unable to be vaccinated.

We explore the role of Facebook groups in spreading false information. We collect data on the content and types of posts shared by Facebook groups that promote the discussion of anti-vaccine beliefs. We find that a handful of authors account for a disproportionately large number of posts and that the posts focus on promoting articles from fake news sites and other online social media. Our results suggest that anti-vaccine groups on Facebook serve as an alternative channel of information for users—both as an "echo" chamber (when users "like" anti-vaccine posts by other users) and as a means of disseminating false stories (when users share a post with others in their social network).

We then study the role of advertising in propagating fake news. In response to criticism over the potential influence of fake news on political outcomes, Facebook banned fake news ads from their advertising networks on November 14, 2016 (Dillet, 2016; Seetharaman, 2016; Wingfield et al., 2016). The intervention marks a major shift in policy from one of the largest social networking sites in the US and occurred when scrutiny heightened over the role that online misinformation may have played in the outcome of the 2016 US presidential election. Since the ban, Facebook does not display ads that link to websites with misleading or illegal content. Because this ban is unrelated to health news, this provides us with the opportunity to study how an exogenous shifter of fake news on health topics affects the sharing of this news.

Our paper is the first to our knowledge that empirically tests the role of advertising in the dissemination of fake news. Theoretically, the effect of advertising on the popularity of fake news on social media is unclear. On one hand, the popularity of fake news may occur in the absence of advertising as users share articles with others in their social network. On the

other hand, advertising may convince users to share an article that they would not otherwise.

To circumvent challenges in measuring the effects of advertising (Gordon et al., 2017; Lewis and Reiley, 2014; Lewis et al., 2011), we exploit a difference-in-differences framework. We study how Facebook's ban on the advertising of fake news affects shares of fake news on Facebook, and we use another prominent social media platform, Twitter, that did not experience any policy change during this period as a control group. We compare the number of shares on Facebook with Twitter for news stories about childhood vaccines before and after Facebook's advertising ban on fake news. Our results suggest that the advertising ban is particularly effective; the shares of fake news articles on Facebook drop by 75% compared to Twitter after the ban.

Our study directly relates to future policymaking. For instance, German regulators are considering regulation that requires Facebook to pay a fine of 500,000 euros for each fake news post that appears on their site (Olsen, 2016). Legislators intend to introduce a bill that will compel Facebook to compensate individuals who have been negatively affected by "fake or defamatory" stories. Recently the Parliament in Malaysia passed the world's first legislation that outlaws fake news: anyone who publishes or circulates misleading information faces up to six years in prison (Beech, 2018). Finally, France's Parliament is also debating a bill aimed at fake news; the bill would allow judges to block content deemed false (Nossiter, 2018).

This paper relates to several strands of literature. An established literature on bias in the media industry dates back to the growth of radio and television (Gentzkow and Shapiro, 2006; Mullainathan and Shleifer, 2005; Baron, 2004; Besley and Prat, 2004). In the few studies that address fake news, Allcott and Gentzkow (2017a) examine whether exposure to fake news influenced electoral outcomes in November 2016 election, and Vosoughi et al. (2018) and Friggeri et al. (2014) examine how false and true news and rumors propagate on social media. Finally, a well-developed literature on the effectiveness and regulation of

online advertising exists (Chiou and Tucker, 2016; Goldfarb and Tucker, 2011, 2015).

2 Fake News and Health Information on Social Media

2.1 Facebook and Twitter

Facebook and Twitter rank as the two largest social media platforms in the US. Users rely on both platforms to obtain news. Approximately two-thirds of US adults use Facebook, and half of Facebook users read news on its site (Pew, 2014). Twitter users account for 16% of US adults, and half of the users read news on its site.

The prevalence and prominence of social networking sites leads an "individual user with no track record or reputation can in some cases reach as many readers as Fox News, CNN, or *New York Times*" (Allcott and Gentzkow, 2017b). Critics allege that most traffic to fake news sites originates from Facebook and that Facebook referrals account for a larger fraction of referrals to fake news sites than real news sites (Wong, 2016; Shavit, 2016). Given that nearly 2 billion monthly users view the "Trending Topics" section of Facebook, Facebook faces increased scrutiny and criticism when fake news stories appear (Chaykowski, 2016).

Figure 1 illustrates a screenshot of an ad on Facebook for a fake news story about vaccines. The top of the ad contains the word "Sponsored" to indicate that this post is an advertisement. This ad links to a news article on `vaccinesrevealed.com` that alleges how vaccines are "neither safe, nor effective."

On November 15, 2016, in response to the concerns about the influence of fake news on the US presidential election, Facebook banned advertisers from running ads that link to fake news stories (Dillet, 2016; Seetharaman, 2016). The ban coincided with Facebook releasing its official policy on fake news sites as described in the Appendix A-1. Facebook bans ads that contain deceptive, false, or misleading content, including deceptive claims, offers, or methods, and it explicitly added fake news sites to the category of "misleading or false content," and .

Figure 1: Screenshot of Advertising of Fake News on Facebook

Note: Source is `gofundme.com`.

2.2 Misleading Health Information on the Internet

Vaccines protect the health of individuals as well as members of the community by preventing further spread of the disease. In fact, some vulnerable populations (people with allergies or weakened immune systems due to cancer, HIV/AIDS, and certain diseases) are unable to receive vaccinations and thus rely on community protection from the disease. The level of vaccination required to achieve this type of community protection or "herd" immunity ranges between 83 to 95 percent (of Health and Services, 2018).

Stories with false or misleading information surround health topics on the Internet. Public health officials voice concerns about the influence of fake news because one-third of US consumers use social media for health care information (Miller, 2017), and more than 40% of consumers say that "information found via social media affects the way they deal with their health."

In particular, false news stories surround the safety of vaccinations. Esposito et al. (2014) finds that the "dissemination of misinformation and anecdotal reports of alleged vaccine reactions by the media, the Internet and anti-vaccination groups leads parents to question the need for immunization." For instance, the vaccine for measles, mumps, and rubella is among the most "frequently omitted of the recommended vaccines, usually because of concerns about the vaccine safety." Fake news articles allege that the vaccine may cause autism even though the medical literature has since debunked such claims.

A growing body of evidence demonstrates that consumers struggle to evaluate the credibility and accuracy of online content. Experimental studies find that exposure to online information that is critical of vaccination leads to stronger anti-vaccine beliefs, since individuals do not take into account the credibility of the content (Nan and Madden, 2012; Betsch et al., 2010, 2013; Allam et al., 2014). Survey evidence also shows that only half of low-income parents of children with special healthcare needs felt "comfortable determining the quality of health websites" (Knapp et al., 2011). Since only 12% of US adults are proficient in health literacy with 36% at basic or below basic levels (Kutner et al., 2006), Fu et al. (2016) state that the influence of "low-quality antivaccine web pages that promote compelling but unsubstantiated messages."

Public health officials across the world express concerns about how fake news may influence parents' decision to vaccinate their children. The president of the Irish Medical Association states that the uptake rates for the HPV vaccine are declining to a "worrying extent" due to false stories about the risks from vaccinations, and he further expresses that

fake news jeopardizes the future health of young women (Power, 2017). Health Minister Beatrice Lorenzin of Italy indicates that the current measles epidemic in Italy and the corresponding declining vaccination rates presents an "emergency generated by fake news" (Press, 2017).

3 Social Sharing of Fake News within Facebook Groups

We first document how users disseminate false or misleading information in Facebook groups. Facebook groups serve as a place of communication for people to share common interests and to express their opinions. In Facebook groups, users may organize around a common cause, issue, or activity, and they may express objectives, discuss issues, post photos, and share related content. These groups are similar to discussion forums where groups of users may share photos, links, and updates (Singh, 2014).

We identify anti-vaccine Facebook groups by performing a keyword search on "anti vaccine" in Facebook and filtering the search results by Facebook groups. Using a script, we collect data from these groups for all posts between May and October 2017.[1]

For each post, we observe the author, message, date of the post, and the type of post. The types of posts include those that display an album, event, link, note, photo, status, or video. A note is a longer message that users access on a separate page; it may be edited and formatted as well as set to different privacies. A status is a short message that users post at the top of the page of the group. We also observe the cumulative number of likes, comments, and reactions for the post as of October 2017. On Facebook, users can respond to a post by "liking" it, commenting on it, or designating a reaction (e.g., "wow," "love," "haha", "sad," "angry") that captures their emotional response to the post.

Figure 2 shows a typical post from an anti-vaccine group.[2] Here an author posts a false

[1]We focus on Facebook groups with at least 100 users. Note that we only observe Public groups on Facebook, as Private groups do not have group identification numbers, which are required in our script to extract the data.

[2]We have redacted the author's name and image.

claim from a fake news site that a "significant link" exists between children who are vaccinated and mental disorders. At the bottom of the post, users may respond by liking the post, writing a comment, or sharing it with others. The fake news article focused on the study's finding of a small correlation found between receiving the MMR vaccine and a subsequent diagnosis of brain-related disorders; the article concludes that vaccines present a real health risk and should not be administered to children. This assertion is misleading because the study ultimately concludes that the modest magnitude of their findings is outweighed by the clear public health benefits of vaccines in preventing mortality and morbidity in childhood infectious diseases, and the study recommends that families maintain vaccination schedules according to CDC guidelines.

Our data suggest that members primarily use Facebook groups to share information. As Figure 3 indicates, a vast majority (over 40%) of posts consists of links. To a lesser extent, users share status, photos, and videos, and they very infrequently post events. This suggests that Facebook groups primarily disseminate information rather than organize events or activities around a common cause.

Table 1 lists the top 20 domains shared in the Facebook groups. Many links direct users back to social media such as Facebook (26%) and Youtube. Interestingly the top links also include a link to governmental fact-based organization (i.e., National Institutes of Health). The National institutes of Health supports the medical literature on vaccination, so users may be linking to fact-based data and then drawing misleading conclusions.

A small number of authors account for a disproportionately large number of posts. In fact, the top 10 authors post approximately one-third of all posts even though the top 10 authors account for less than 1% of users that post on the Facebook groups.

Most users respond to posts with an emotional reaction or sharing the post with others. Table 2 reports the average number of responses for a post on the Facebook groups. Users may respond to posts by either writing a comment, sharing it with others in their social

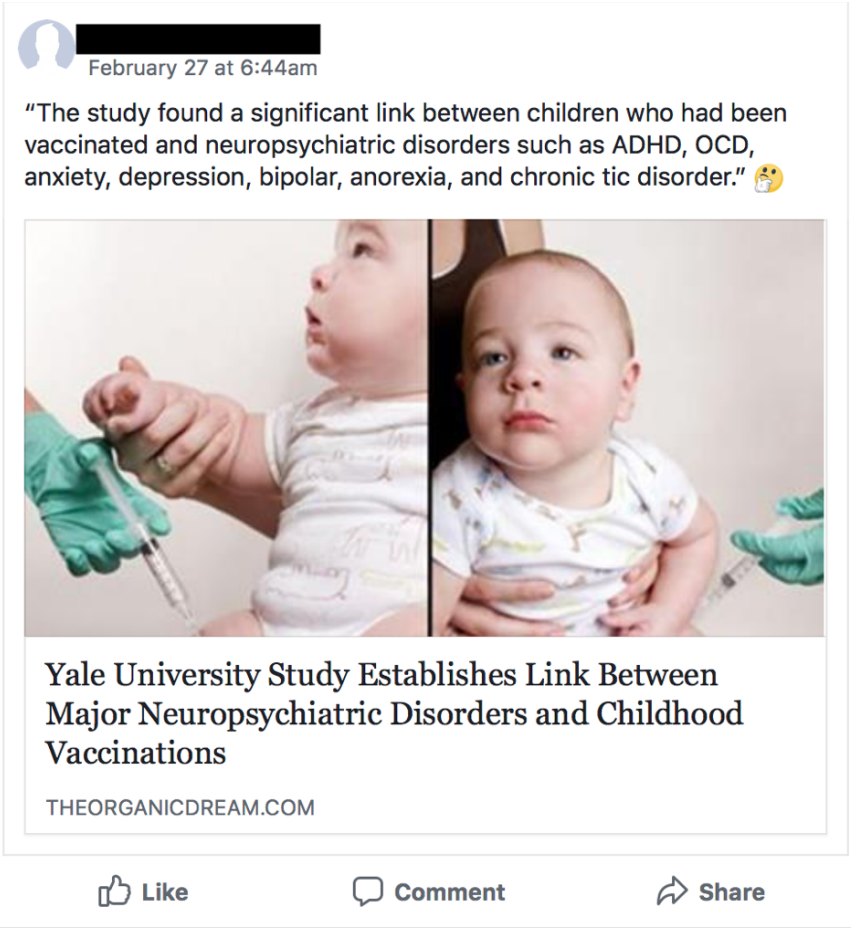

Figure 2: Screenshot of post on Anti-vaccine Facebook groups

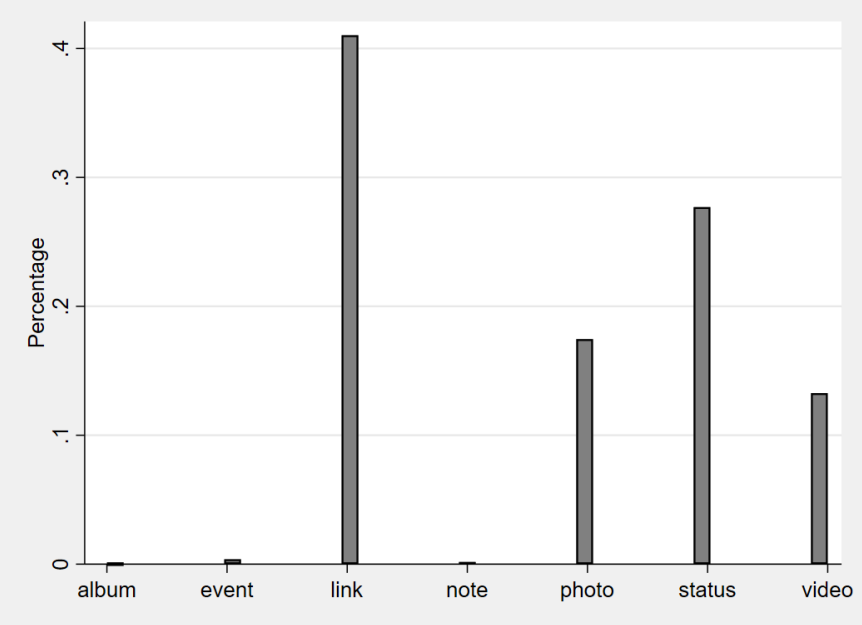

Figure 3: Types of posts on Anti-vaccine Facebook groups

Table 1: Top 20 domains shared in Facebook groups

	Pct
facebook.com	0.36
youtube.com	0.057
naturalnews.com	0.022
youtu.be	0.018
bolenreport.com	0.016
bit.ly	0.015
ncbi.nlm.nih.gov	0.012
vaccineimpact.com	0.012
vaxxter.com	0.011
greenmedinfo.com	0.0095
ourfamilymagazine.com	0.0095
thevaccinereaction.org	0.0091
yournewswire.com	0.0085
inmymindtoday.com	0.0081
go.thetruthaboutvaccines.com	0.0080
healthimpactnews.com	0.0075
m.youtube.com	0.0059
ow.ly	0.0048
vactruth.com	0.0048
t.co	0.0044

Table 2: Summary statistics for Facebook groups

	Mean	Std Dev	Min	Max
Number of comments	1.69	7.80	0	578
Number of shares	2.35	17.4	0	2101
Number of reactions	7.63	23.3	0	2075
Observations	24025			

Notes: Each observation represents a post on the top 20 anti-vaccine Facebook groups.

network, or indicating an emotional reaction (i.e.., likes, loves, wows, hahas, sads, angrys, special). A typical post receives 7 emotional reactions and 2 shares. To a lesser extent, users post comments; a post receives about 1.7 comments on average. Figure 4 reveals how the vast majority of reactions are "likes" that indicate users respond positively to posts on the Facebook groups.

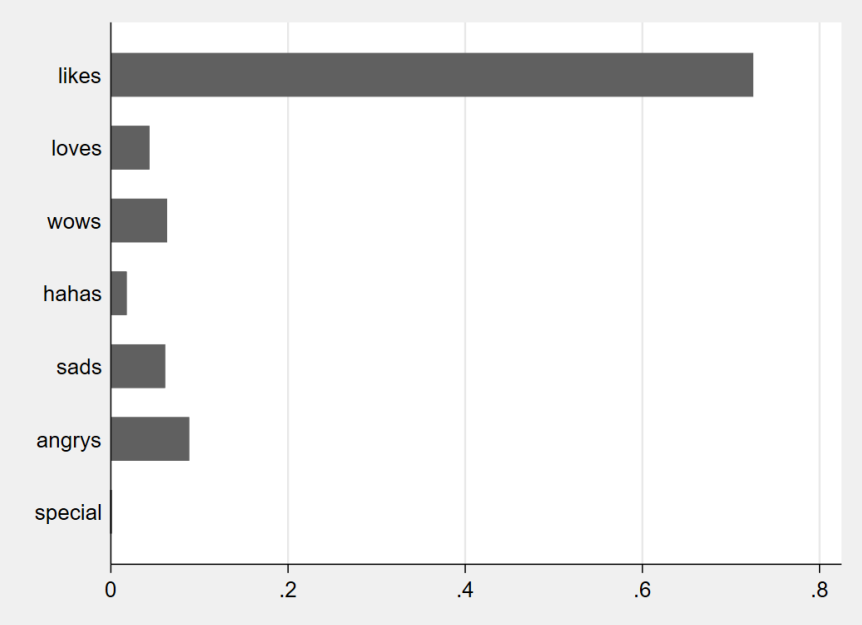

Figure 4: Types of reactions on Anti-vaccine Facebook groups

Our evidence points to Facebook groups as "echo chambers" whereby users with similar views reinforce those views and opinions (Sunstein, 2017). Users primarily respond to the content by "liking" it rather than writing a comment to invite further dialogue. On average, the number of reactions per post (7.6) exceed the number of comments (1.7).

We also find that the Facebook groups allows a relatively small number of users to promote and disseminate information beyond the members of the group. A small fraction of users post a disproportionately large amount of content on the Facebook groups, and the content focuses on promoting links to either fake news sites or other posts on social media such as Facebook and Youtube. Members then respond by sharing the original post from the Facebook groups to others in their social network, thereby extending the reach of the post to outside the Facebook groups. As a handful of authors promote fake news that ultimately reach members beyond the Facebook group, Facebook groups serve almost serve a similar purpose to 'advertising'. However, as we document below advertising itself of fake news provides a distinctive channel for the spread of misinformation.

4 Does Advertising Spread Fake News?

4.1 Data on the Popularity of Fake News on Facebook and Twitter

We investigate the effect of advertising of fake news articles on Facebook and Twitter for common childhood vaccines. First, we obtain a list of common vaccines used in the United States for children ages 0-18 years old from the website for the Centers for Disease Control and Prevention.[3] Table A-1 in the Appendix lists the keywords.

Then we collect data from Hoaxy, a search engine that tracks "online misinformation" (Shao et al., 2016). Indiana University Network Science Institute (IUNI) and the School of Informatics and Computing's Center for Complex Networks and Systems Research (CNetS) created the platform to track stories from "sources that often publish inaccurate, unverified, or satirical claims according to lists compiled and published by reputable news and fact-checking organizations" (Hoaxy, 2016).

We submit a query to Hoaxy for each vaccine keyword, and Hoaxy returns a list of the most relevant articles related to the keyword. We perform our queries on January 11-15, 2017. For each article, Hoaxy reports the title of the article, url, cumulative number of Facebook shares, and cumulative number of tweets on Twitter at the time of the query. We identify articles based upon their titles that relate to vaccines and remove those that do not.[4]

Our sample consists of a cross-section of articles. Since we collect data after Facebook's initial ban on advertising of fake news, our data contain articles published before and after the Facebook ban on November 14, 2016.[5] Table 3 reports the summary statistics. The variable *shares* measures the number of shares that an article receives. The indicator variable *PostBan* equals one if the article was published after Facebook's advertising ban, and the

[3] `https://www.cdc.gov/vaccines/vpd/vaccines-list.html`

[4] Some articles are political stories that we remove from our sample because they do not relate to vaccines, e.g., "Hillary Clinton Diagnosed with Pneumonia."

[5] We only examine articles with a positive number of shares on Facebook, since we study how Facebook's advertising ban affects shares on Facebook relative to Twitter.

Table 3: Summary statistics

	Mean	Std Dev	Min	Max
shares	2409.5	11342.1	1	168397
Postban	0.64	0.48	0	1
Facebook	0.50	0.50	0	1
days	391.3	39.0	341	460
Observations	354			

Notes: Each observation represents a website and keyword combination.

indicator variable *Facebook* equals one for shares on Facebook and equals zero for shares on Twitter. The variable *days* measures the number of days that the article has been published online.

In our sample, an article receives on average 2000 shares on either Facebook or Twitter. Nearly 20% of the articles in our sample were published after the advertising ban. We have a matched sample with exactly half of observations for shares on Facebook and the other half for shares on Twitter. The articles have been in publication for approximately seven months on average.

Table A-2 in the Appendix describes the demographics of users of Facebook and Twitter. Reassuringly the demographics are relatively similar with the exception that Facebook has a lower proportion of users that are male. Facebook and Twitter have a similar age profile of their users, and both platforms have a sizable fraction (35 to 45 percent) of users in the highest income bracket. Overall the table suggests that users of Twitter provide a plausible control for users of Facebook.

4.2 Estimating the Effect of Advertising on Sharing

As a preliminary analysis, we compare the number of shares of articles with fake news published before and after the advertising ban on Facebook compared to Twitter. Figure 5 depicts the the average logarithm of shares across keywords for Twitter and Facebook before and after the advertising ban. While shares for Twitter remain similar before and after the

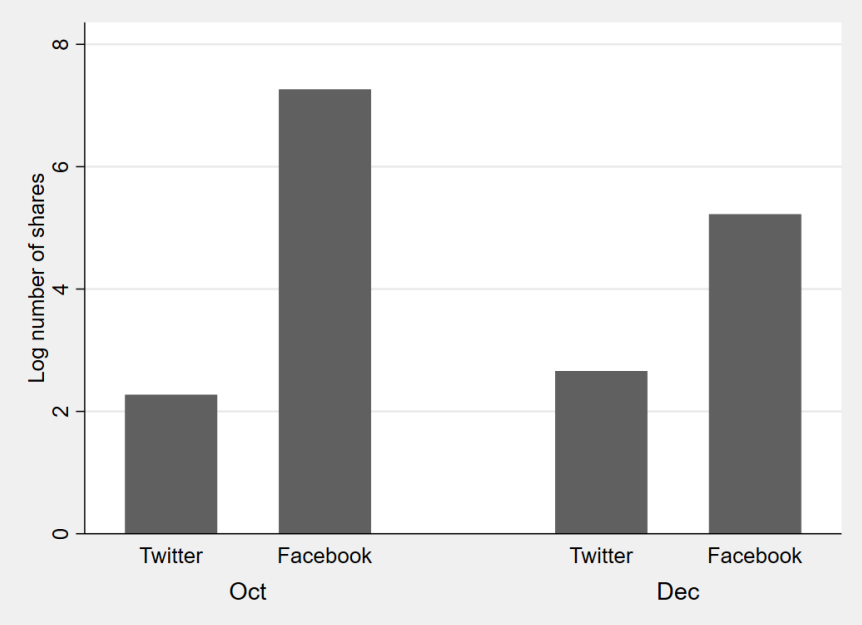

Figure 5: Logarithm of shares on Facebook drop relative to Twitter comparing the month before and after Facebook's ban on advertising by fake news sites

advertising ban on Facebook, shares for Facebook decline sharply after the ban.

For a formal regression analysis, we estimate the following equation. For every search term j on platform k, we regress the logarithm of the number of shares received by news article i:

$$\begin{aligned} log(shares_{ijk}) &= \beta_0 + \beta_1 Facebook_k \times PostBan_i + \beta_2 Facebook_k \\ &+ \beta_3 days_k + \gamma_i + \alpha_j + \epsilon_{ijkt} \end{aligned} \tag{1}$$

where *Facebook* is an indicator variable equal to 1 if the platform is Facebook and 0 if Twitter; *PostBan* is an indicator variable that equals one if the news article was published after Facebook's ban of advertising by fake news sites. The variable *days* is the number of days that the article has been in publication. The control γ is a fixed effect for the article's month of publication, and α is a fixed effect for keyword. We cluster our standard errors by

platform and keyword.

Our identification strategy compares shares on Facebook and Twitter for articles published before and after Facebook's advertising ban. We exploit the advertising ban as an exogenous shifter of advertising for fake news for vaccines because the ban was unrelated to health topics and enacted due to concerns about political misinformation. The number of shares on Facebook for articles published after the advertising ban on Facebook reflect popularity in the absence of advertising while the number of shares on Facebook for articles published prior to the advertising ban include the influence of advertising. We control for seasonal differences in popularity of articles by using shares on Twitter over this same time period as a control group. Because we observe shares in Facebook and Twitter for the same article and control for the number of days in publication, we can isolate the effect of advertising.

We interpret our estimated coefficients of the semi-log specification as the "ratio-of-ratios" (Mullahy, 1999). For instance, to determine the effect of the advertising ban on shares, we compute the corresponding ratio-of-ratios:

$$\frac{\left\{\frac{E[shares|Facebook=1,PostBan=1]}{E[shares|Facebook=1,PostBan=0]}\right\}}{\left\{\frac{E[shares|Facebook=0,PostBan=1]}{E[shares|Facebook=0,PostBan=0]}\right\}} = exp(\beta_1). \tag{2}$$

In Equation (2) above, the fraction in the numerator (proportionately) compares the expected number of shares on Facebook before and after the advertising ban. The fraction in the denominator compares the expected number of shares to the control Twitter before and after the advertising ban. The formula avoids the "retransformation bias" for estimating the number of shares from the semi-log regression and offers a natural interpretation for the estimated coefficients directly (Mullahy, 1999).

In other words, $exp(\beta_1)$ captures the extent to which shares on Facebook fall proportionately relative to shares on Twitter after the advertising ban. If the expression is less than

Table 4: Facebook shares drop relative to Twitter after Facebook's ban on advertising of fake news

	(1)	(2)	(3)
PostBan × Facebook	-1.415***	-1.415***	-0.694***
	(0.461)	(0.341)	(0.143)
PostBan × Facebook × Early Childhood			-0.974**
			(0.383)
Facebook × Early Childhood			1.439***
			(0.363)
Facebook	3.970***	3.970***	2.845***
	(0.336)	(0.298)	(0.257)
days	0.00102	-0.0156	-0.0156
	(0.0200)	(2.708)	(2.704)
Month Fixed Effects	Yes	Yes	Yes
Keyword Fixed Effects	Yes	Yes	Yes
Article Fixed Effects	No	Yes	Yes
Observations	354	354	354
R-Squared	0.430	0.804	0.809

Notes: Robust standard errors. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The dependent variable is the logarithm of the number of shares of a news article from a platform—either Facebook or Twitter. The regressions estimate click behavior before and after the advertising ban by Facebook.

one, then the "interaction" between the advertising ban and Facebook is negative; shares on Facebook decrease compared to Twitter after the advertising ban. If the expression is equal to one, then no interaction effect exists. If the expression is greater than one, then the interaction is positive; shares on Facebook increase compared to Twitter after the advertising ban.[6]

Table 4 reports the results of our regression. Column (1) examines the period before and after Facebook started banning ads on fake news on November 15, 2016. The coefficient on $Facebook \times PostBan$ indicates that after the advertising ban, shares on Facebook are 25% or $exp(-1.415) \approx 0.25$ of the level prior to the ban. In other words, shares on Facebook drop by 75%.

[6]This interpretation is equivalent to a traditional difference-in-differences setup where a positive coefficient on the interaction term ($exp(\beta_1) > 1$) implies a positive effect on the treatment group; a zero coefficient ($exp(\beta_1 = 1)$ implies no effect.

Overall, a direct ban on ads of fake news dramatically decreases the number of shares of fake news by 75%. We explore the magnitude in two ways. First, we perform a back-of-the-envelope calculation of the total decline in shares from the advertising ban. Since our sample includes articles that were shared a total of 1.6 million times on Facebook before any of the advertising bans, a decrease of 75% equates to a decline in total shares of 1.12 million for the fake news sites in our sample.

Second, we calculate a benchmark of how referrals from Facebook to fake news sites change after the advertising ban. We collect additional data from comScore that tracks the incoming traffic to each of the fake news sites in our sample from Hoaxy.[7] Table 5 compares the average percentage of incoming traffic to the fake news sites on Hoaxy that originate from Facebook and the top three search engines. We consider the top three search engines, since they represent platforms that consumers use for information and that also feature fake news sites and stories. Facebook accounts for a large fraction (13 percent) of incoming referrals while the other three search engines each account for significantly fewer (less than 10% of referrals). If incoming traffic declines by the same proportion as the number of shares, then Facebook's share of incoming traffic would drop from 13% to 3%. In other words, the advertising ban reduces incoming traffic from Facebook to that comparable of other search engines.

Our results suggest that advertising has a large influence on the spread of false news on social media. Approximately of 75% of the popularity of fake news may be attributed to advertising. The policy measure of banning advertising of fake news presents an effective way of mediating the popularity of false information online.

[7]We use comScore's Search Planner database, which tracks the online activity of a panel of more than 2 million users based in the US and aggregates their search patterns to the search-term level for resale to commercial clients. ComScore recruits its panel members through affiliate programs and partnering with third party application providers. ComScore emphasizes and discusses the representativeness of its sample to the general population in its Marketer User Guide. Several academic studies use the comScore data and note it as a "highly regarded proprietary [source] for information on the size and composition of media audiences" (Gentzkow and Shapiro, 2011; Montgomery et al., 2004; De Los Santos et al., 2012).

Table 5: Incoming traffic for fake news sites

	Pct
Facebook	13.3
Google	6.51
Bing	2.46
Yahoo	1.63

5 Robustness and Falsification Checks

5.1 Controlling for Differences between Articles

We perform several robustness and falsification checks in this section. First, we consider whether underlying differences in popularity between articles drive our results; did the decline in shares of fake news on Facebook occur because articles published before and after the ban differ in underlying popularity? As a robustness check, we include fixed effects for each article in Column (2) of Table 4. This identification strategy compares the number of shares on Facebook and Twitter for the exact same article, so we difference out any article-specific characteristics that are constant over time. The results are qualitatively similar; the point estimate for $PostBan \times Facebook$ in Column (2) is similar to the specification in Column (1) without article fixed effects. This provides some reassurance that unobserved differences in articles do not drive our main results.

5.2 Vaccines in Early vs. Later Childhood

We examine whether the effect of the ban was larger for vaccines administered earlier in childhood. We partition our list of vaccines into two groups: those that are typically administered to young children and those that are administered later in life.[8] The idea is that parental concern and fake news may focus more on vaccines administered to the most vulnerable—young children as opposed to later in life. Therefore we would expect the effect of the ban to be larger for vaccines administered earlier in childhood. In Column (3) of

[8]The vaccines administered later in childhood and adulthood are for flu, HPV, and shingles.

Table 6: No negative pre-trend for Facebook shares prior to its ban on advertising by fake news sites

	(1)	(2)
FakePost × Facebook	0.464	0.464
	(0.519)	(0.390)
Facebook	3.275***	3.275***
	(0.348)	(0.356)
days	0.0175	2.874
	(0.0257)	(2.613)
Month Fixed Effects	Yes	Yes
Keyword Fixed Effects	Yes	Yes
Observations	286	286
R-Squared	0.412	0.772

Notes: Robust standard errors. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The dependent variable is the logarithm of the number of shares of a news article from a platform—either Facebook or Twitter.

Table 4, the negative coefficient on the interaction of $PostBan \times Facebook \times Early$ supports our hypothesis that the ban on fake news has a larger effect on articles about vaccines administered earlier in childhood.

5.3 Checking for a Pre-Trend

We investigate whether the decline in Facebook shares relative to Twitter after the advertising ban may be due to a pre-existing negative trend in shares for Facebook relative to Twitter. In Table 6, we perform a falsification check by restricting our sample to articles published in the four months prior to the advertising ban. We create a indicator variable $FakePost$ that equals one for articles published in the second part of this time period. The coefficient on the interaction between the $FakePost$ and $Facebook$ is not statistically significant, indicating no evidence of a pre-existing negative trend.

5.4 Placebo Test with Medical Conditions

We conduct a placebo test for the effect of the advertising ban on medical conditions. We consider common medical conditions for children, such as ear infections, as a placebo group because these conditions may be less susceptible to false news as opposed to the more con-

Table 7: Summary statistics for placebo group of health conditions

	Mean	Std Dev	Min	Max
Number of comments	1.69	7.80	0	578
Number of shares	2.35	17.4	0	2101
Number of reactions	7.63	23.3	0	2075
Observations	24025			

Notes: Each observation represents a post on the top 20 anti-vaccine Facebook groups.

troversial health topic of vaccines. We expect that the advertising ban would have a smaller (if any) effect on the sharing of fake news for medical conditions than for vaccines.

We identify health conditions listed under the "Most Viewed Health Topics" on the Dr. Sears website (`http://www.askdrsears.com/`). Dr. Sears is a prominent pediatrician and his website "Ask Dr. Sears" serves as a online resource for parents about childrens' health. In the Appendix, Table A-3 lists the keywords for these medical conditions.

Then we perform keyword searches for these health conditions on Hoaxy and collect data for the top 20 articles of each keyword in January 2017. We run a regression similar to equation 1 using this dataset. Table 7 lists the summary statistics. Our results in table 8 suggest that the advertising ban did not have an effect on the sharing of fake news articles for medical conditions because the estimated coefficient is smaller in magnitude than the results for vaccine keywords and is also not statistically significant.

6 Conclusion

This paper examines how false information spreads on social networking sites, using Facebook as a case study. We first we examine the role of anti-vaccination groups on Facebook in spreading false information. The Facebook groups serve as an "echo chamber" when members "like" posts from other users and also serve as a channel of disseminating fake news when users share the posts with others in their social group. A small fraction of authors account for a large majority of posts, which reinforces the concern that social media allows an individual to reach a wide audience and share information without editorial or fact-checking input.

Table 8: No effect on placebo group of health conditions

	(1)	(2)
PostBan × Facebook	-0.592	-0.592
	(0.445)	(0.414)
Facebook	1.987***	1.987***
	(0.375)	(0.430)
days	0.00199	1.321
	(0.0204)	(0.826)
Month Fixed Effects	Yes	Yes
Keyword Fixed Effects	Yes	Yes
Article Fixed Effects	No	Yes
Observations	186	186
R-Squared	0.401	0.694

Notes: Robust standard errors. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$. The dependent variable is the logarithm of the number of shares of a news article from a platform—either Facebook or Twitter.

We also study the effectiveness of advertising in spreading fake news about vaccines on social media sites. We examine a policy experiment where Facebook banned ads containing links to fake news sites in response to criticism over the influence of fake news on the US presidential election. Our results indicate that this ban on advertising led to a dramatic decline of 75% in the number of shares on Facebook relative to Twitter, which had no change in its advertising policy during this time.

Our results suggest potential ways of curbing the influence of fake news on social networking sites. We explore how false information spreads within Facebook groups dedicated to promoting false information. In addition, we illustrate how advertising regulations may effectively curtail the popularity of articles with fake news.

Our study also suggests another potential mechanism for counteracting fake news: creating Facebook pages of real news and using this advertising to disseminate accurate information. Future work can focus on whether positive advertising can counteract the effects of negative advertising.

In the context of our study on childhood vaccines, fake news may potentially harm consumers, as public health officials fear that fake news' misleading claims about the safety

of vaccines lowers vaccination rates. The community as a whole depends upon vaccinations by individuals, which makes this is a pressing public concern.

Consumers find health information to be extremely difficult to evaluate and to determine validity, so consumers benefit when firms and policymakers take direct action by preventing the spread of fake news articles.

References

Allam, A., P. Schulz, and K. Nakamoto (2014). The impact of search engine selection and sorting criteria on vaccination beliefs and attitudes: Two experiments manipulating Google output. *Journal of Medical Internet Research 16*, e100.

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

Allcott, H. and M. Gentzkow (2017b). Social Media and Fake News in the 2016 Election. *Stanford University, working paper*.

Baron, D. (2004). Persistent Media Bias. *Journal of Public Economics 90*, 1–36.

Beech, H. (2018). As Malaysia Moves to Ban 'Fake News,' Worries About Who Decides the Truth. *New York Times*.

Besley, T. and A. Prat (2004). Handcuffs for the Grabbing Hand? Media Capture and Government Accountability. *London School of Economics*.

Betsch, C., F. Renkewitz, T. Betsch, and C. Ulshofer (2010). The Influence of Vaccine-critical Websites on Perceiving Vaccination Risks. *Journal of Health Psychology 26*, 446–455.

Betsch, C., F. Renkewitz, and N. Haase (2013). Effect of Narrative Reports about Vaccine Adverse Events and Bias-Awareness Disclaimers on Vaccine Decisions: A Simulation of an Online Patient Social Network. *Medical Decision Making 33*, 14–25.

Chaykowski, K. (2016). Facebook Says Fake News Sites Aren't Allowed on Its Advertising Network. *Forbes*.

Chiou, L. and C. Tucker (2016). How does pharmaceutical advertising affect consumer search? *working paper, MIT*.

De Los Santos, B., A. Hortacsu, and M. R. Wildenbeest (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review 102*(6), 2955–80.

Dillet, R. (2016). Google and Facebook Ban Fake News Sites from Their Advertising Networks. *TechCrunch*.

Esposito, S., N. Principi, and G. Cornaglia (2014). Barriers to the vaccination of children and adolescents and possible solutions. *Clinical Microbiology and Infection 20*, 25–31.

Fahy, E., R. Hardikar, A. Fox, and S. Mackay (2014). Quality of Patient Health Information on the Internet: Reviewing a complex and Evolving Landscape. *Australasian Medical Journal 7*, 24–28.

Fox, S. and M. Duggan (2013). Pew Internet and American Life Project.

Friggeri, A., L. A. Adamic, D. Eckles, and J. Cheng (2014). Rumor cascades. In *ICWSM*.

Fu, L., K. Zook, Z. Spoehr-Labutta, P. Hu, and J. Joseph (2016). Search Engine Ranking, Quality, and Content of Web Pages That Are Critical Versus Noncritical of Human Papillomarvirus Vaccine. *Journal of Adolescent Health 58*, 33–39.

Gentzkow, M. and J. Shapiro (2006). Media Bias and Reputation. *Journal of Political Economy 114*(2), 280–316.

Gentzkow, M. and J. M. Shapiro (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics 126*, 1799–1839.

Goldfarb, A. and C. Tucker (2011). Advertising Bans and the Substitutability of Online and Offline Advertising. *Journal of Marketing Research 48*(2), 207–227.

Goldfarb, A. and C. Tucker (2015). Standardization, standards and online advertising. *Management Science 61*, 2707–2719.

Gordon, B., F. Zettelmeyer, N. Bhargava, and D. Chapsky (2017). A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook. *working paper*.

Hoaxy (2016). Frequently Asked Questions About Hoaxy.

Knapp, C., V. Madden, and H. Wang (2011). Internet use and eHealth literacy of low-income parents whose children have special health care needs. *Journal of Medical Internet Research 13*, e75.

Kutner, M., E. Greenberg, Y. Jin, and C. Paulsen (2006). The health literacy of America's adults: Results from the 2003 National Assessment of Adult Literacy. *US Department of Education, National Center for Education Statistics 13*.

Lewis, R., J. Rao, and D. Reiley (2011). Here, There, and Everywhere: Correlated Online Behaviors Can Lead to Overstimates of the Effects of Advertising. *Proceedings of the 20th International Conference on the World Wide Web*, 157–166.

Lewis, R. and D. Reiley (2014). Online ads and offline sales: measuring the effect of retail advertising via a controlled experiment on Yahoo! *Quantitative Marketing and Economics 12*, 235–266.

Miller, M. (2017). 33 percent of us consumers use social media for health care info. *Search Engine Watch*.

Montgomery, A. L., S. Li, K. Srinivasan, and J. C. Liechty (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science 23*(4), 579–595.

Mullahy, J. (1999, November). Interaction effects and difference-in-difference estimation in loglinear models. NBER Technical Working Papers 0245, National Bureau of Economic Research, Inc.

Mullainathan, S. and A. Shleifer (2005). The Market for News. *American Economic Review 95*, 1031–1053.

Nan, S. and K. Madden (2012). HPV Vaccine Information in the Blogosphere: How Positive and Negative Blogs Influence Vaccine-Related Risk Perceptions, Attitudes, and Behavioral Intentions. *Health Communication 27*, 829–836.

Nossiter, A. (2018). Macron Pushes Bill Aimed at 'Fake News' as Critics Warn of Danger. *New York Times*.

of Health, U. D. and H. Services (2018). Vaccines Protect Your Community.

Olsen, P. (2016). Germany Wants Facebook to Pay For Fake News. *Forbes*.

Pew (2014). How Social Media is Reshaping News.

Power, J. (2017). HPV vaccine uptake rate falls 15 percent among young girls. *The Irish Times*.

Press, A. F. (2017). Italy makes childhood vaccinations compulsory. *CTV News*.

Seetharaman, D. (2016). Facebook Bans Fake News Sites from using its Advertising Network. *The Wall Street Journal*.

Shao, C., G. L. Ciampaglia, A. Flammini, and F. Menczer (2016). Hoaxy: A platform for tracking online misinformation. In *Proceedings of the 25th International Conference Companion on World Wide Web*, WWW '16 Companion, pp. 745–750.

Shavit, N. (2016). Data on Facebook's Fake News Problem. *Jumpshot*.

Singh, K. (2014). What's really the difference between facebook pages, groups and profiles? *Mashable*.

Smith, L., R. Amlot, J. Weinman, J. Yiend, and G. Rubin (2016). A systematic review of factors affecting vaccine uptake in young children. *Vaccine 35*, 6059–6069.

Sunstein, C. (2017). *Republic.com*. Princeton University Press.

Vosoughi, S., D. Roy, and S. Aral (2018). The spread of ture and false news online. *Science 359*, 1146–1151.

Wingfield, N., M. Isaac, and K. Benner (2016). Google and Facebook Take Aim at Fake News Sites. *New York Times*.

Wong, J. (2016). Almost All the Traffic to Fake News Sites is from Facebook, New Data Show. *The Medium*.

A Appendix

A-1 Facebook policy on fake news sites (excerpts from April 6, 2017)

Working to Stop Misinformation and False News

By Adam Mosseri, VP, News Feed

We know people want to see accurate information on Facebook and so do we.

False news is harmful to our community, it makes the world less informed, and it erodes trust. Its not a new phenomenon, and all of us tech companies, media companies, newsrooms, teachers have a responsibility to do our part in addressing it. At Facebook, were working to fight the spread of false news in three key areas:

- disrupting economic incentives because most false news is financially motivated;
- building new products to curb the spread of false news; and
- helping people make more informed decisions when they encounter false news.

Disrupting Economic Incentives

When it comes to fighting false news, one of the most effective approaches is removing the economic incentives for traffickers of misinformation. Weve found that a lot of fake news is financially motivated. These spammers make money by masquerading as legitimate news publishers and posting hoaxes that get people to visit their sites, which are often mostly ads.

Some of the steps were taking include:

- Better identifying false news through our community and third-party fact-checking organizations so that we can limit its spread, which, in turn, makes it uneconomical.
- Making it as difficult as possible for people posting false news to buy ads on our platform through strict enforcement of our policies.
- Applying machine learning to assist our response teams in detecting fraud and enforcing our policies against inauthentic spam accounts.
- Updating our detection of fake accounts on Facebook, which makes spamming at scale much harder.
- Update on May 10, 2017: Weve made updates so people see fewer posts and ads in News Feed that link to low-quality web page experiences.
- Update on August 9, 2017: Weve made updates to address cloaking so that what people see after clicking an ad or post matches their expectations.
- Update on August 28, 2017: Weve made an update in which repeat offenders that repeatedly share stories marked as false will no longer be allowed to advertise on Facebook.

Building New Products

Were building, testing and iterating on new products to identify and limit the spread of false news. We cannot become arbiters of truth ourselves its not feasible given our scale, and its not our role. Instead, were working on better ways to hear from our community and work with third parties to identify false news and prevent it from spreading on our platform.

Some of the work includes:

- Ranking Improvements: Were always looking to improve News Feed by listening to what the community tells us. Weve found opportunities like the fact that if reading an article makes people significantly less likely to share it, that may be a sign that a story has misled people in some way. Were continuing to test this signal and others in News Feed ranking in order to reduce the prevalence of false news content.

  Update on June 30, 2017: We made an update to help show more informative links in News Feed.

- Easier Reporting: Weve always relied on our community to determine what is valuable and what is not. Were testing ways to make it easier to report a false news story if you see one on Facebook, which you can do by clicking the upper right hand corner of a post. Stories that are flagged as false by our community then might show up lower in your feed.

- Working with Partners: We believe providing more context can help people decide for themselves what to trust and what to share. Weve started a program to work with independent third-party fact-checking organizations, who are signatories of Poynters Code of Principles. Well use the reports from our community, along with other signals, to send stories to these organizations. If the fact-checking organizations identify a story as false, it will get flagged as disputed and there will be a link to a corresponding article explaining why. Stories that have been disputed also appear lower in News Feed.

Helping People Make More Informed Decisions

Though were committed to doing everything we can to reduce the spread of false news to as close to zero as possible, we also need to make sure we take steps to address the problem when people do encounter hoaxes. To that end, were exploring ways to give people more context about stories so they can make more informed decisions about what to read, trust and share and ways to give people access to more perspectives about the topics that theyre reading.

Some of the work weve been focused on includes:

- Facebook Journalism Project: We are committed to collaborating with news organizations to develop products together, providing tools and services for journalists, and helping people get better information so they can make smart choices about what they read. We are convening key experts and organizations already doing important work in this area, such as the Walter Cronkite School of Journalism and Mass Communication at Arizona State University, and have been listening and learning to help decide what new research to conduct and projects to fund. Working with the News Literacy Project, we are producing a series of public service announcements (PSAs) to help inform people on Facebook about this important issue.

- News Integrity Initiative: Weve joined a group of over 25 funders and participants including tech industry leaders, academic institutions, non-profits and third party organizations to launch the News Integrity Initiative, a global consortium focused on helping people make informed judgments about the news they read and share online. Founding funders of this $14-million fund include Facebook, the Craig Newmark Philanthropic Fund, the Ford Foundation, the Democracy Fund, the John S. and James L. Knight Foundation, the Tow Foundation, AppNexus, Mozilla and Betaworks. The initiatives mission is to advance news literacy, to increase trust in journalism around the world and to better inform the public conversation. The initiative, which is administered by the CUNY Graduate School of Journalism, will fund applied research and projects, and convene meetings with industry experts.
- Update on August 3, 2017: We made an update where if an article has been reviewed by fact checkers, we may show the fact checking stories below the original post in Related Articles.

We need to work across industries to help solve this problem: technology companies, media companies, educational organizations and our own community can come together to help curb the spread of misinformation and false news. By focusing on the three key areas outlined above, we hope we will make progress toward limiting the spread of false news and toward building a more informed community on Facebook.

Table A-1: Keywords for vaccines

keyword
chicken pox
diptheria
dtp
flu
haemophilus influenzae
hepatitis a
hepatitis b
hib
hpv
human papillomavirus
measles
meningitis
meningococcal
mmr
mumps
pertussis
pneumococcal
pnuemonia
polio
rotavirus
rubella
shingles
smallpox
tetanus
varicella
whooping cough

Table A-2: Demographic description of users

Measure	Facebook	Twitter
Male	47.2	55.6
Age 18-24	12.1	14.9
Age 25-34	17.2	19.1
Age 35-44	16.4	15.6
Age 45-54	18	17.2
Age 55+	29.7	25.2
Income <25k	10.9	7.5
Income 25-60k	26.1	20.6
Income 60-100k	28.0	27.0
Income >100k	35	44.9

Source: comScore

Note: This table reports the fraction of users within each demographic category. Statistics are reported for users of Facebook and Twitter.

Table A-3: Keywords for placebo group of health conditions

Health Keywords
allergies
asthma
breastfeeding
coughs and colds
dehydration
diarrhea
ear infection
fever
head injuries
lice
rashes
ringworm
sore throat
vomiting

Notes: This table lists the health keywords for the "Most Viewed Health Topics" on the Dr. Sears website `http://www.askdrsears.com/`.