Algorithms and Online News

Lesley Chiou[*] and Catherine Tucker[†,‡]

January 16, 2025

Abstract

Algorithms prioritize online content for users. We examine how changes in Facebook's algorithm for its News Feed shifts referrals to news sites. Our results indicate that after an algorithm change which demoted engagement bait, recently established news sites experienced a decline in referrals from Facebook relative to Twitter. We find some evidence that more Internet-savvy websites were able to recover more swiftly from the negative effects of the algorithm change. We also find that the number of user logoffs from Facebook declined relative to Twitter. However, we do not find evidence that another algorithm change that promoted content from friends affected referrals to news sites. Our findings have significant implications for recent regulations, such as Australia's News Media Bargaining Code, which aim to promote transparency and competition in response to algorithmic changes. These regulations suggest that differences in organizations' capacities to adapt to shifts in algorithms will affect market outcomes.

[*] Economics Department, Occidental College, CA

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

[‡] We thank Katelyn Fink, Jessica Lee, and Hannah Trautwein for excellent research assistance.

1 Introduction

Algorithms play an important role in the online consumption of information because they prioritize content delivered to users. In particular, social media sites such as Facebook use algorithms to determine which sources to present to users and in which prominent positions. Algorithms are key arbiters of content in the same way that editors played an important role for traditional news media. We examine how changes in Facebook's algorithm shift referrals to news sites and how news sites respond.

The news media industry provides an excellent context for this study. First, online news plays an important role within the news media industry. Almost 40% of consumers in the U.S. obtain news through social media, websites, and/or apps.[1] Second, algorithms potentially exert a strong in uence on behavior of news organizations. Publishers face strong incentives to respond to shifts in rankings because referrals generate tra c to their site and advertising revenues.

We study changes in Facebook's algorithm for its News Feed at the end of 2017 and beginning of 2018. Facebook's News Feed presents content that \matters the most" to users, each time they visit Facebook. It contains a \personalized, ever-changing collection of photos, videos, links, and updates from friends, family, businesses, and news sources."[2] In December 2017, Facebook demoted engagement bait on its News Feed in an e ort to reduce spammy content.[3] On January 31, 2018, Facebook prioritized content from friends and family in its News Feed to produce \more meaningful interactions between people."[4] Since space on News Feed is limited, content from news sites was in e ect demoted.

We exploit a di erence-in-di erences strategy to examine the consequences of an algo-

[1] Pew Research Center, \The Modern News Consumer," July 7, 2016. https://www.j8(o3ww.396/.tps:9i2e.tps:9i2e1i6.)0

rithm change on the news media industry. Our approach estimates how Facebook's algorithm changes a ected referrals of news sites from Facebook compared to referrals of news sites from Twitter, which experienced no algorithm change during this period.

We nd that after the algorithm demoted engagement bait, newer news sites experienced a decline in referrals from Facebook relative to Twitter. The e ect was short-lived for news sites that were more Internet savvy as measured by higher Search Engine Optimization (SEO) rankings. However, we do not nd evidence that a second algorithm change that promoted content from friends led to a decline in referrals to news sites by Facebook relative to Twitter. The nding is surprising given how much attention the second algorithm change attracted in the press; news sites expressed concerns that referrals to their sites would drop (Chaykowski, 2016; Mullin, 2018).

We also nd evidence that demoting engagement bait seemingly improved the quality of the user's experience on Facebook because logo s from Facebook declined relative to Twitter. We do not nd evidence that the algorithm change that promoted content from friends was correlated with any change in logo s to Facebook relative to Twitter.

Policymakers focus on algorithms because of concern over how algorithm changes might negatively a ect rms. For instance, in Australia, policymakers plan to include an amendment to the News Media Bargaining Code that would require platforms to notify news media businesses of algorithm changes 28 days in advance (Barbaschow, 2021). The noti cation would apply to \algorithm changes that are likely to materially a ect referral tra c to news, ..., and any substantial changes to the display and presentation of news and advertising directly associated with news." More broadly, our results are related to policy concerns about the long-run health of the news media industry. By shifting referrals, algorithms can potentially in uence the long-term e ect on the quantity and types of news consumed as well as incentives to produce news.

There has been limited empirical work that studies how algorithms a ect the news media industry. An exception is Calzada et al. (2024), which focuses on the short- and long-term e ects of changes to a search engine Google's algorithm on the news media visits and market structure. By contrast, our paper focuses on a social media site Facebook and which news media sites were able to respond to Facebook's algorithm changes and how quickly the sites reacted. Our study also relates more generally to prior work on how online platforms can in uence consumption of news media. In particular, these studies focus on how news aggregators can a ect consumption of online news (Athey et al., 2017; Chiou and Tucker, 2017; George and Hogendorn, 2020; Calzada and Gil, 2020). More generally, our paper also relates to prior work on how online rankings by search engines can a ect consumer choices; these studies primarily focus over consumers' purchase decisions and search engine revenues (De los Santos and Koulayev, 2017; Ghose et al., 2014; Ursu, 2018).

2 The Internet and the News Media Industry

2.1 Online news publishers and Facebook

Publishers increasingly rely on online platforms such as Facebook and Twitter for referrals to their sites. Whenever consumers navigate to a publisher's landing page, the publisher accrues revenues because consumers are exposed to online advertisements on the page.

Facebook's News Feed presents a list of stories in the middle of its home page that is constantly updated. The News Feed includes \stat updates, photos, videos, links, likes from people, Pages and groups that [users] follow on Facebook." The posts presented in the News Feed are \in uenced by [a user's] connections and activity on Facebook."[5] Thus the algorithm performs a key editorial role by curating the content and its position rank in the News Feed.

[5] https://www.facebook.com/help/1155510281178725

2.2 Facebook Algorithm Changes

At the end of 2017 and beginning of 2018, Facebook implemented two major algorithm changes to its News Feed. The rst algorithm change on December 18, 2017 combatted engagement bait on Facebook.[6] The idea was to remove \spammy posts on Facebook that goad [people] into interacting with likes, shares, comments, and other actions." Facebook refers to such tactics as \engagement bait," which \seeks to take advantage of our News Feed algorithm by boosting engagerment in order to get greater reach." Some examples of engagement bait include vote baiting (where a post asks you to vote), react baiting (where a post asks you to like it), and share baiting (where a post asks you to share it).

The second algorithm change on January 29, 2018 promoted \meaningful" posts from friends and family in News Feed.[7] Facebook further explained that because \space in News Feed is limited, showing more posts from friends and family...means we'll show less public content, including videos and other posts from publishers or businesses."[8] This policy change could signi cantly a ect news sites, as promoting content from friends and family resulted in the demotion of content from news publishers.

3 How does an algorithm change shift referrals to new sites?

3.1 Data on Referrals from Facebook and Twitter to News Sites

Our primary dataset derives from comScore. ComScore monitors the online behavior of a demic research and is recognized 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).

We identify all sites listed under the category of News Media in comScore. We focus on lower-level domains (.com, .net, etc.) because we are interested in tra c to speci c websites. ComScore also provides information aggregated to sites owned by the same entity.[9]

For each site, we query comScore for the monthly referrals from Facebook and Twitter during the months before and after Facebook's algorithm change from August 2017 to March 2018.[10] We observe the number of entries from Facebook or Twitter for each news sites in a given month. We create a balanced panel over our time period, and our nal sample includes sites with positive referrals in at least half of the time period.

We also restrict our sample to sites with referrals from both Facebook and Twitter because these groups represent our treatment and control for the natural experiments. Our nal sample contains a total of 136 sites. Table A-1 in the Appendix lists the top 40 sites with the highest average daily referrals. As expected the top sites include common and well-known news brands.

For each site, we compute the daily number of entries by dividing the monthly number of entries by the number of days in a month because months vary by the number of days. We also collect data on the year the news source was founded because this provides a measure of the age of the news source. We collect data from Woorank on the Search Engine Optimization (SEO) ranking for each site. The SEO ranking captures the ability of the website to optimize their website to receive tra c from a search engine's results page. We view this as a measure of Internet savviness of the site. A higher ranking indicates more Internet savviness.

[9] We remove weather sites because our focus is on not on websites that report statistics. We also include \USA Today" channel because no information is available on its lower-level domain usatoday.com.

[10] We end our sample at March 2018 because of Facebook's changes to its News Feed in April: https://about.fb.com/news/2018/04/news-feed-fyi-more-context/ .

Figure 1: Daily entries from Facebook and Twitter to news sites

errors by news site because of correlations in news consumption over time for the same site.

We interpret the coefficients β_1 and β_2 as the effect of the corresponding algorithm change on referrals by comparing referrals to the news site from Facebook and Twitter before and after Facebook's algorithm change. The algorithm change exogenously shifts the prominence of news sites because the change was motivated by demoting engagement bait and promoting content from friends and family in Facebook's News Feed. We control for seasonal differences in popularity of news sites by using referrals from Twitter as a control group.

Our estimated coefficients from the semi-log specification represent a "ratio-of-ratios" (Mullahy, 1999). For instance, to determine the effect of the algorithm change that demoted engagement bait on referrals, we compute the corresponding ratio-of-ratios:

$$\frac{\left\{\frac{E[\text{referrals} \mid Facebook=1; Post=1]}{E[\text{referrals} \mid Facebook=1; Post=0]}\right\}}{\left\{\frac{E[\text{referrals} \mid Facebook=0; Post=1]}{E[\text{referrals} \mid Facebook=0; Post=0]}\right\}} = \exp(\beta_1). \tag{2}$$

In Equation (2), the numerator compares the expected number of referrals from Facebook before and after the algorithm change, while the denominator does the same for Twitter, acting as a control. This comparison avoids "retransformation bias" from the semi-log regression and provides a straightforward interpretation of the estimated coefficients (Mullahy, 1999).

The value $\exp(\beta_1)$ measures the proportional change in Facebook referrals relative to Twitter after Facebook's algorithm change demoted engagement bait. A value below one indicates Facebook referrals dropped compared to Twitter, while a value of one shows no change. A value above one suggests Facebook referrals increased relative to Twitter. This mirrors a traditional difference-in-differences approach (Chiou and Tucker, 2013).[11]

Table 2 reports the results of the regression. Column (1) estimates the baseline regression

[11] A positive coefficient on the interaction term ($\exp(\beta_1) > 1$) implies a positive effect on the treatment group while a negative coefficient indicates a negative effect. A zero coefficient ($\exp(\beta_1$ ae48(acebookid)1stimuseline1

in Equation (1). The estimated coefficient on $PostFriends \times Facebook$ is small in magnitude and not statistically significant. The estimated coefficient on $PostBait$ is negative and statistically significant at 10% significance level. According to the results, referrals for news sites declined by 14% after the algorithm change that demoted engagement bait.[12]

The results suggest that the algorithm change which promoted posts from friends did not exert a measurable effect on referrals to new sites from Facebook. The results appear somewhat surprising given the amount of attention in the press and news industry over how the promotion of posts from friends would have dire consequences on news sites (Chaykowski, 2016; Mullin, 2018). In particular, concern existed that news sites would be shuttered after receiving a vast fall in referrals. Most publishers "expressed some concerns about unexpected and unexplained changes to ... Facebook search algorithms, most notably... Facebook News Feed," and they cite the Facebook algorithm change that promoted posts from friends and family as such an example (Competition and Authority, 201). Publishers argued that "a reduction in website traffic resulting from an algorithm change has a direct financial consequence for their business." Our estimates do not indicate that news sites experienced significant losses in the wake of the algorithm change that promoted posts from friends.

One possible explanation is that users substituted towards other means of accessing news on Facebook. Another explanation is that news articles are primarily shared organically through posts from friends, which are therefore unaffected by the first algorithm change.

By contrast, the algorithm change that demoted clickbait, which did not receive particular attention in the press, did seem to exert a negative effect on news sites. In Column (1), our results indicate that after the demotion of clickbait, referrals to news sites from Facebook dropped by 24% relative to Twitter.[13]

Overall, our results suggest that the algorithm change related to clickbait was more

---

[12] The estimated coefficient for β_1 is -0.149. Therefore $\exp(-0.149)$ equals 0.86. Referrals are 86% of their previous levels and thus decline by 14%.

[13] The calculation using the ratio-of-ratios is $\exp(-0.149) = 0.86$, and $1 - 0.86 = 0.24$.

Table 3: Falsification check: No evidence of a pre-trend

	(1)	(2)
FakePost × Facebook	-0.217	-2.109
	(0.142)	(4.178)
FakePost × Facebook × Year founded		0.000972
		(0.00218)
Website Fixed Effects	Yes	Yes
Observations	544	544
R-Squared	0.710	0.725

Notes: Robust standard errors. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The dependent variable is the logarithm of the number of daily entries in thousands plus one to a news site from a platform—either Facebook or

4.1 Are well-established news sites less a ected by the algorithm changes?

The prior section establishes that when Facebook’s algorithm demoted engagement bait, only the highest quartile of youngest news sources in the sample. In sum, our results are robust to including measures of age either linearly or non-linearly through quartiles.

4.2 How Do News Sites Respond to Algorithm Changes?

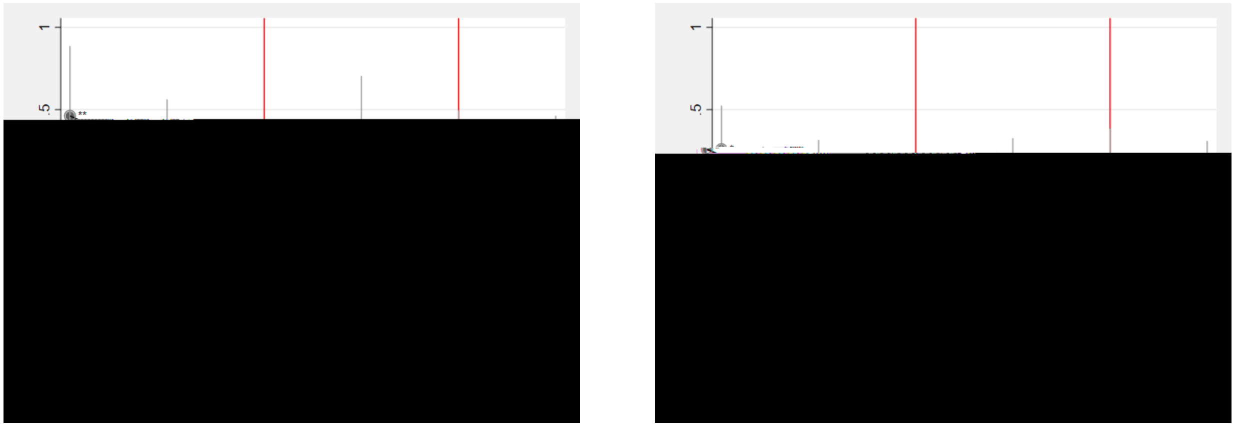
Given the decline in visits after the algorithm change that demoted clickbait, a natural question is whether news sites were able to respond to the decline in referrals and adapt to the algorithm change. To explore this, we examine how the algorithm change affected news sites by their SEO ranking in the months after the algorithm change. As described in Section 3.1, the SEO ranking measures the ability of a website to optimize its website to receive traffic from a search engine's results page. We view this as a measure of the Internet savviness of the site. A higher ranking indicates more Internet savviness. We would expect news sites that are more sophisticated in their web techniques to be able to respond and to recover from any negative effects of an algorithm change.

We run a regression similar to Equation (1) and include full interactions of monthly indicator variables instead of indicator variables for $PostBait$ and $PostFriends$. Note that the monthly indicator variables are interpreted relative to the month of December because this month is the omitted condition. Thus, the coefficients on the interactions of the monthly indicator variables captures the month-by-month response to an algorithm change. We run two separate regressions for news sites with SEO rankings above and below the median SEO ranking.

Figure 2 graphs the estimated monthly coefficients and confidence intervals for the interactions of monthly indicator variables and the indicator variable for Facebook. More specifically, Figure 2(a) graphs the estimated monthly coefficients for news sites with high SEO rankings while Figure 2(b) graphs the estimated monthly coefficients for news sites with low SEO rankings.

According to Figure 2(a), some evidence exists that websites with high SEO rankings

Figure 2: Estimated monthly coe cients by SEO Ranking



(a) High SEO Ranking

(b) Low SEO Ranking

The fgures graph estimated monthly coefcients of a regression of the logarithm of daily entries in thousands plus one to a news site from Facebook relative to Twitter. The bands indicate the confdence intervals around each estimated coefcient. The vertical lines correspond to months when the algorithms changes were implemented|December for the demotion of engagement bait and February for the promotion of posts be correlated with a decline in the number of logo s from Facebook.

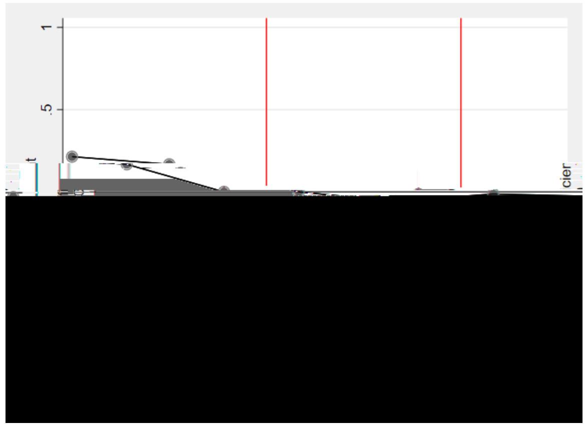
We collect data from comScore on the number of exits in each month in our sample from Facebook and Twitter. Note that a user has a choice when on Facebook to either navigate to other another site or to terminate their online session (logo).

Table 4 reports the summary statistics of the data on exits from social media sites. We compute the daily number of exits as the number of exits divided by the number of days in a month because months vary in length. Note that we have a small sample of 12 observations because we observe two social media sites (Facebook and Twitter) over a period of 6 months.

Then we run a regression similar to Equation (1). However this time we de ne our dependent variable as the logarithm of the number of exits from Facebook or Twitter.

Table 5 reports the results of the regression. The estimated coe cient on $PostBait \times Facebook$ is negative and statistically signi cant, suggesting that the algorithm change to demote clickbait was correlated with a decline in logo s. However, the estimated coe cient on $PostFriends \times Facebook$

Figure 3: Estimated monthly coe cients for logo s



The fgure graphs estimated monthly coefcients of logarithm of daily exits in thousands plus one to a news site from Facebook relative to Twitter. The bands indicate the confdence intervals around each estimated coefcient. The vertical lines correspond to months when the algorithms changes were implemented|December for the demotion of engagement bait and February for the promotion of posts from friends and family. Robust standard errors. *p< 0:1 and ** p< 0:05.

6 Conclusion

We examine the e ect of two algorithm changes on Facebook's News Feed and explore how referrals to news sites consequently shift relative to Twitter. Our di erence-in-di erences strategy indicates that when the algorithm demoted engagement bait, referrals to more recently established news sites declined. One possible explanation is that younger news sites might engage in more frequently in spammy tactics, perhaps in an e ort to accumulate readers and prominence. As a consequence, users appear to positively respond to the improved experience and were less likely to logo from Facebook relative to Twitter. By contrast, our results do not lend evidence that any changes in referrals to news sites or logo s occurred when the algorithm promoted content from friends and family (and thereby demoted content from news sites).

Our results also illustrate a contrast between the two algorithm changes. The rst algorithm change was an attempt to identify meaningful content through promoting posts from friends and family. The second algorithm change was an attempt to identify irrelevant content through demoting clickbait posts. Given that the second change was more impactful on user behavior, one possibility is that it may be easier for the algorithm to identify low-quality content, where there perhaps may be more agreement among users what constitutes click-bait, instead of high-quality content, where it may be harder to predict what is meaningful to users.

Our paper related directly to the increased attention from policymakers on the role of algorithms in shaping the news industry. Australia's recent amendment to its News Media Bargaining Code highlights these concerns. The amendment requires platforms to notify news organizations 28 days in advance of any algorithmic changes that could signi cantly impact referral tra c. The goal of the policy is to o er news outlets a chance to adapt to algorithm shifts that could otherwise undermine their tra c and revenue. Our results show that algorithm tweaks, such as Facebook's demotion of engagement bait, can have notable consequences for news sites, especially newer or less-established outlets.

Our paper has several caveats. First, our analysis focuses on news media sites and not other forms of online information. Second, we focus on referrals from social media and do not address direct navigation or other marketing channels. Notwithstanding these limits, our paper provides a useful step in understanding how algorithms may arbitrate online information and how news media may respond.

References

Athey, S., M. Mobius, and J. Pal (2017). The Impact of Aggregators on Internet News Consumption. *working paper*.

Barbaschow, A. (2021). Media Bargaining Code Amendments Include a More 'Streamlined' Algorithm Change Notice. *ZDNet*.

Calzada, J., N. Duch-Brown, and R. Gil (2024). Who Decides What News You Read: Search Engines and Media Markets. *working paper*.

Calzada, J. and R. Gil (2020). What Do News Aggregators Do? Evidence from Google News in Spain and Germany. *Marketing Science 39*(1), 134{167.

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

A Appendix

Table A-1: Top 40 news sites

	Avg Entries
9news.com	8203.6
bloomberg.com	9100.9
boston.com	7046.6
breitbart.com	22780.6
businessinsider.com	9516.3
cbsnews.com	19327.2
chicagotribune.com	8046.9
cnbc.com	10105.1
cnet.com	7017.6
cnn.com	128047.1
dailymail.co.uk	38420.7
dailywire.com	20307.1
forbes.com	29787.8
foxnews.com	99253.5
ibtimes.com	5620.8
independent.co.uk	7032.7
kiwireport.com	18657.6
latimes.com	8955.0
legacy.com	15055.3
marketwatch.com	8066.2
medium.com	6086.8
nbcnews.com	17309.0
ndtv.com	8639.9
npr.org	18465.4
nydailynews.com	11310.2
nypost.com	23701.9
nytimes.com	80372.7
patch.com	12616.3
politico.com	12116.9
reuters.com	6177.6
slate.com	6492.9
theatlantic.com	7207.0
theblaze.com	6098.1
theguardian.com	10686.1
thehill.com	18555.3
vox.com	5518.8
washingtonpost.com	47406.7
wsj.com	12586.5
wtop.com	10097.7
zerohedge.com	14662.6

Notes: This lists the top 40 news sites in our nal sample with the highest, average daily entries from Facebook or Twitter.

Table A-2: Demographic description of users

Measure	Facebook	Twitter
Male	48.4	59.8
Age 18-24	11.0	14.6
Age 25-34	18.4	20.2
Age 35-44	15.6	15.1
Age 45-54	17.9	16.9
Age 55+	30.4	23.5
Income < 25k	8.4	6.7
Income 25-60k	24.4	19.7
Income 60-100k	29.0	27.0
Income > 100k	38.1	46.6

Source: comScore

Note: This table reports the fraction of users within each demographic category for October 2017. Statistics are reported for users of Facebook and Twitter.