News Feeds and User Engagement: Evidence from the Reddit News Tab

Alex Moehring

April 1, 2022

Abstract

We study how the introduction of a new non-personalized news feed impacts user engagement quantity, quality, and diversity on Reddit. In June 2018, Reddit introduced the News tab on iOS devices that surfaces popular content from a curated list of news-related communities. We leverage this natural experiment to identify the causal effects of the News tab on iOS user engagement in a difference-in-differences design. We find that the News tab increases the share of iOS devices that engage with news-related content and there is a relatively larger increase in low-quality engagement, measured through voting on the platform. We also find that the diversity of engagement within news categories and within articles from publishers across the political spectrum increases as a result of the News tab. These results suggest that non-personalized feeds can be an important tool to mitigate algorithmic filter bubbles, and need not come at the expense of reduced user engagement.

1 Introduction

Social media platforms play an important role in everyday life, with the average person with access to the internet spending over two hours daily on social media [Kemp, 2020]. The prevalence of social media in modern life has important implications for individual well-being [Allcott et al., 2020], the quality of news people consume [Vosoughi et al., 2018], and political polarization [Levy, 2021]. Perhaps the most common way users interact with many of the most popular social media platforms is through a news feed that aggregates content from across the platform (e.g. Facebook News Feed and Twitter timeline). A growing body of research has demonstrated the importance of news feed algorithms on individual well-being [Kramer et al., 2014], platform engagement [Dujeancourt et al., 2021], and exposure to counter-attitudinal news [Bakshy et al., 2015, Levy, 2021, Huszár et al., 2021]. Here we study the effects of a large social media platform introducing a popularity-driven news feed on engagement quantity, quality, and diversity. We find that the introduction of the feed increases engagement in some, but not all, featured communities and the engagement induced by the feed is heterogeneous in quality, with many communities seeing a larger relative increase in low-quality engagement compared to high-quality engagement. We also find suggestive evidence that the communities that saw the largest increases in engagement were communities most often featured in the new feed. Moreover, the introduction of the feed increases individual engagement diversity, suggesting non-personalized news feeds can be an important tool in mitigating algorithmic filter bubbles [Pariser, 2011].

This study leverages a natural experiment on the Reddit platform where a News tab offering popular content from a curated list of communities was introduced on iOS devices but not desktop or Android devices. At the time, Reddit statements indicated the iOS only roll out was to test and improve the feed and the platform had plans to introduce this feature across all devices. Therefore, we view this as an exogenous change to the app and argue that, absent the introduction of the News tab, Reddit engagement trends would have been similar across Android and iOS users. This allows us to identify the causal effect of the News Tab using a difference-in-differences strategy.

We find that the introduction of the news feed induces a statistically significant increase in the probability of posting any content on a news-related community. This increase is concentrated in the Politics, Technology, Entertainment, and Business related communities that were featured in the News tab. The News tab caused the monthly probability of posting by iOS users to increase by 19.6% (0.69 percentage points) in the Politics community, 17.3% (0.50 percentage points) in Technology communities, 7.0% (0.35 percentage points) in Entertainment communities, and 106.0% (0.21 percentage points) in Business communities. However, this new engagement that was caused by the News Tab is of heterogeneous quality. We find a statistically significant and economically meaningful increase in low-quality engagement as measured by voting on the platform in the Politics (36.0%, 0.23 percentage points) and Technology (13.1%, 0.08 percentage points) communities, and a non-statistically significant decline in Science and US & World communities. While the probability of posting a high quality comment also increased in the Politics and Technology communities, the increase was smaller in relative terms. Notably, the non-personalized feed also increases individual engagement diversity as implemented through the Shannon Entropy [Shannon, 1948, Holtz et al., 2020] and Herfindahl–Hirschman Index [Rhoades, 1993, Claussen et al., 2019]. This is true both for the diversity of engagement across the communities included in the News tab as well as for the diversity of engagement across articles from publishers of different political slant.

These results have important managerial and policy implications. In particular, we highlight the effects of social media news feed algorithms on engagement quantity, quality, and diversity. Despite the increase in engagement, the relative rise in low-quality engagement has the potential to make the platform less valuable to existing members by increasing the costs of finding high quality discussion and information [Gu et al., 2007]. This presents a trade-off, as prior work suggests the increase in engagement diversity caused by the News tab may have positive implications for user retention [Oestreicher-Singer and Zalmanson, 2013, Anderson et al., 2020]. While the increase in engagement diversity across categories of news has important managerial implications, from a policy perspective, the increase in the diversity of engagement across publishers from various political viewpoints suggests that non-personalized feeds can be an important tool to mitigate algorithmic filter bubbles.

2 Related literature

In this paper we contribute to several streams of existing literature. First, we add to the literature studying the impacts of social media feed algorithms on users and society. This includes work studying the impact of social media feeds on individual well-being [Kramer et al., 2014, Allcott et al., 2020], media consumption [Bakshy et al., 2015, Allcott et al., 2020, Levy, 2021], exposure to content from politicians [Huszár et al., 2021], and user engagement [Dujeancourt et al., 2021]. Kramer et al. [2014] find that changes to the Facebook News Feed that promote (or suppress) posts containing positive expressions cause users to post more (less) positive posts and less (more) negative posts, highlighting the importance of the News Feed algorithm in determining the content that users interact with and downstream effects on user behavior. Similarly, Bakshy et al. [2015], Allcott et al. [2020], and Levy [2021] find that the news feed impacts the news people read, and in particular, exposure to counter-attitudinal sources. We contribute to this literature by further demonstrating the importance of news feed design on user behavior. In particular, the addition of the news feed we consider increases user engagement, and the diversity of user engagement, suggesting that non-personalized feeds can help mitigate filter bubbles that are often generated by social media feeds.

Second, we contribute to the growing literature studying the impacts of news feeds and algorithmic recommendations on consumer behavior. This has been studied in the context of product sales [Oestreicher-Singer and Sundararajan, 2012, Hosanagar et al., 2014, Lee and Hosanagar, 2019] as well as content consumption [Bakshy et al., 2015, Claussen et al., 2019, Holtz et al., 2020, Dujeancourt et al., 2021]. The existing work find that personalized algorithms increase content consumption relative to manually curated recommendations [Claussen et al., 2019, Holtz et al., 2020] and chronologically ordered news feeds [Dujeancourt et al., 2021]. In addition, Bakshy et al. [2015], Claussen et al. [2019] and Holtz et al. [2020] find (to varying extents) that personalized recommendations decrease individual consumption diversity, supporting the notion of algorithmic recommendations leading to filter bubbles. We contribute to this body of work by studying the causal effects of a new non-personalized feed on engagement and engagement diversity. In contrast to the work studying personalized feeds [Bakshy et al., 2015, Claussen et al., 2019, Holtz et al., 2020], we find the non-personalized news feed increases individual engagement diversity and this increase in diversity does not come at the expense of decreased engagement. Moreover, this paper is among the first to study the impacts of a new feed at a major social media outlet on engagement quantity, quality, and diversity.

Finally, we add to the literature studying motivations for user generated content. Past research has investigated numerous interventions to induce additional user generated content. These interventions include financial incentives with mixed results [Cabral and Li, 2015, Khern-am nuai et al., 2018, Burtch et al., 2018], successful social norm interventions [Chen et al., 2010, Burtch et al., 2018], and status or rewards [Goes et al., 2016, Restivo and Van De Rijt, 2012, Gallus, 2017, Burtch et al., 2021]. Of particular relevance to this work are the trade-offs faced in stimulating additional user generated content. For example, Khern-am nuai et al. [2018] find that financial incentives can increase the quantity of online reviews, though these incentives result in the marginal reviews being of lower quality. Gu et al. [2007] emphasize this tradeoff explicitly and study the competing positive network externalities, stemming from additional engagement providing more information, and negative externalities, if additional low-quality engagement distracts members of the community and increases costs of finding relevant information. We contribute to this literature by asking if additional news feeds can stimulate user-generated content, in addition to the impacts these have on the quality of this engagement.

3 Setting

Reddit is a popular social media platform founded in 2005 with over 52 million daily active users as of January 2020.[1] The platform consists of over 100,000 active communities called subreddits, which host user-generated content focused on a particular topic. Within a community, users can post new submissions or comment on others' submissions. By default, content is presented to users

[1] `https://www.redditinc.com/press`

using a proprietary algorithm that favors upvotes and fresher content.[2]

Voting on content is an important part of Reddit both practically, as it is a key driver of content promotion, and as a method of rewarding content that contribute to the community and demoting those that do not.[3] While norms vary within communities, Reddit guidelines are explicit that voting should reflect contributions to the community and conversation.[4] In particular, downvoting only because you disagree with the content is explicitly discouraged and downvoting should be reserved for content that is not contributing to the community's conversation. Therefore, in this study we will use voting data to infer post quality as judged by members of the community.

Reddit is accessible to users through web browsers or mobile apps. In April 2016 Reddit launched their official mobile apps for Android and iOS devices. Before the official Reddit apps were supported, there were a number of third party apps that allowed users to browse the site and many unofficial apps are still available today, though the official Reddit app is the dominant app in the market.[5] Users of the official mobile app are able to access three primary sections of the app through a navigation bar at the top of the screen: "Popular", "Home", and "News", though the latter News section is only available to users with iOS devices. The Popular tab aggregates popular content from across the site and the Home tab aggregates content from communities in which the user is a member. The News tab is the focus of this study and is discussed in detail in Section 3.1.

3.1 Natural Experiment

In June 2018, Reddit introduced an update to its mobile app on Apple (iOS) devices that introduced the News tab, which provided a feed of content from communities that focus on discussion and sharing of news related content. The tab is displayed prominently in the mobile app alongside the Home tab that shows submissions from communities a user is a member of and the Popular tab that shows popular content from across the platform (Figure 1). Within the News tab, users first view a feed containing posts from all news categories. Users may then select individual topics to view more focused feeds that display posts related to the selected topic. The communities that are displayed in the News tab are chosen to be those that most often engage with news, who are actively moderated and in compliance with Reddit policies on acceptable content and guidelines for healthy communities, and who require the title of posts linking to news articles to be an accurate reflection of the article title. These guidelines result in most posts in the News tab following a

[2] The exact details of this algorithm are not publicly available.

[3] Upvoting (downvoting) a user's post or comment impacts their Karma score, which is a publicly available number summarzing "how much good the user has done for the reddit community" (`https://www.reddit.com/wiki/faq`).

[4] "Vote. If you think something contributes to conversation, upvote it. If you think it does not contribute to the subreddit it is posted in or is off-topic in a particular community, downvote it." (`https://www.reddithelp.com/hc/en-us/articles/205926439`)

[5] While data on installs is not publicly available, the official Reddit applications have received more than 2 million reviews on both the Apple App Store and the Google Play Store. The next closest competitor has roughly 400,000 reviews.

common structure, where the post title is an article headline and the body links to the full article (Figure 1).

Figure 1: Screenshot of Reddit News tab

In Reddit's public comments at the time, they announced that the News tab was originally being released on iOS, but would eventually be available on most devices.[6] Our empirical strategy, which will be discussed in greater detail in Section 4.3, relies on the assumption that, absent the introduction of the News tab, engagement trends of iOS users in our sample would have followed engagement trends of Android users in our sample.[7] We provide evidence consistent with such an

[6] https://www.reddit.com/r/announcements/comments/8sth30/extra_extra_were_launching_a_news_tab_as_a_beta

[7] Our preferred results require a slightly weaker assumption that common trends hold conditional on observed pre-treatment engagement. This is because we use Coarsened Exact Matching weights in the analysis [Iacus et al., 2012], described in Appendix A.

assumption in Section 4.4, though the assumption cannot be explicitly tested empirically.

4 Data

The data for this study are based on a dataset of public Reddit submissions and comments described in Baumgartner et al. [2020].[8] We focus on posts between June 2017 and June 2019 which contain a total of 349 million submissions and 3.0 billion comments during this period. We restrict our sample to the subset of users for whom we can infer their mobile device, and this sample has 28.0 million total comments across all communities during the period. We also focus on comments rather than submissions, as comments make up the majority of posts and this is particularly evident in communities promoted on the News tab. In our sample of users, comments on communities featured in the News tab account for 97.89% of all posts. Table 1 below shows descriptive statistics for our sample.

4.1 Inferring device type

A drawback of the Baumgartner et al. [2020] data relative to the proprietary data collected by the platform is that we only observe publicly available information, which does not include the device a user was using when making a post. As a result, we must infer device type from posts on the platform. To do so, we consider the subset of users who have posted in the RedditMobile community, which is an official Reddit community for announcements, discussion, and feedback on the official Reddit mobile apps. When posting in this community, users are often posting feedback about the mobile apps and typically include explicit tags about the device and version of the mobile app they are giving feedback (Figure 2). We infer user device using the following procedure. First, if the user tagged a particular operating system in their post we assign their device accordingly. Second, if the user has tagged their operating system on the community through author 'flair' (for example, the first post in Figure 2 has tagged iOS 14 as their author flair), we assign them to that device. Finally, if both of these methods fail and the post is a comment, we assign the commenter the device of the post they are commenting on. This procedure allowed us to identify 18,274 Android users and 19,127 iOS users. There were an additional 1,579 users who authored posts that would have been classified as both Android and iOS devices, and these users are excluded from all analyses.

[8]Gaffney and Matias [2018] find evidence of missing data in this dataset. In particular, they find that less than 0.04% of comments and 0.65% of submissions are missing from the dataset in the early years. We believe missing data on this scale is unlikely to be driving our results for two reasons. First, as discussed in Baumgartner et al. [2020], the data collection process has improved as a result of the flaws highlighted in Gaffney and Matias [2018] which analyze data before the period studied here. Second, Gaffney and Matias [2018] find that heavy users are more likely to be impacted than light users. Our primary outcomes are indicators if a user posted *any* posts in a month. Therefore, for the outcome to be changed we would have to be missing all of a users posts in a given month which is less likely.

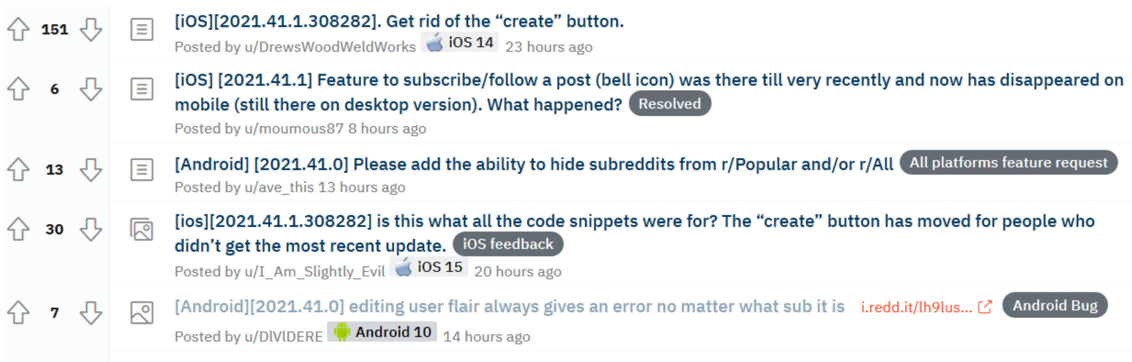


Figure 2: Screenshot of Reddit Mobile feed use to identify user devices

The process of inferring device operating systems limits our sample substantially. As a result, we include all users who have posted on the RedditMobile community rather than only those posting before the News tab was introduced. The primary risk of inferring device type from posts after the News tab was introduced is a bias if users select devices based on this intervention, but we believe this is unlikely to be driving our results.[9] A more likely issue is device switching that is independent of the Reddit News tab. This should bias our estimates of engagement toward zero as long as the News tab did not induce some users to engage less with news related communities.

4.2 Communities of interest

The News tab includes 54 communities (subreddits) that are structured into 8 higher-level categories of news.[10] In this study we aggregate engagement to the category level and focus on heterogeneity along news categories rather than individual communities. We do this because the News tab is structured around these categories and the smaller number of categories facilitates comparisons of heterogeneous effects.

We exclude a handful of communities due to concerns about the common trends assumption. First, we exclude technology communities that specifically reference Apple or Google as we are conditioning on users having an Apple or Android device and expect these users to have different

[9] When the News tab was announced, Reddit administrators were explicit that they intended to make this feature available across all devices which should mitigate switching immediately following the release. The announcement referenced specific plans for availability on desktop and a top comment from an administrator stated the intention to make this available across all platforms (`https://www.reddit.com/r/announcements/comments/8sth30/extra_extra_were_launching_a_news_tab_as_a_beta/e12auz7?utm_source=share&utm_medium=web2x&context=3`). As it became evident that this feature was not planned to be released on Android, selective switching is more plausible, though we find it unlikely that a substantial share of our sample is choosing a smartphone operating system because of this particular Reddit feature.

[10] When Reddit launched the News tab in 2018, the list of communities that were referenced were not made public. To approximate the list of communities that are promoted by the News tab we consider the communities that are present as of August 2021.

engagement patterns on these communities. Second, we exclude all communities in the Crypto category. This is because of the massive increase in traffic in late 2017 that resulted in differential engagement by iOS and Android users. In the end, we focus on the following 8 categories of news: US/World, Politics, Technology, Science, Sports, Business, Gaming, and Entertainment.

4.3 Empirical Strategy

We estimate the effect of the News tab on Reddit activity using a difference-in-differences design that makes a common trends assumptions. Formally, we model outcomes Y_{it} using a two-way fixed effects panel regression model

$$Y_{it} = \alpha_i + \lambda_t + \tau \mathrm{Post}_t D_i + \varepsilon_{it} \tag{1}$$

where Y_{it} is the outcome of interest, D_i is an indicator equal to one if user i has an iOS device, Post_t is an indicator for the post-treatment period, and α_i (λ_t) represent unobserved individual (time) fixed effects. Identification of this model comes from a common trends assumption that assumes, absent the introduction of the News tab, the average outcome for iOS and Android users would have had the same variation over time [Abadie and Cattaneo, 2018].

In addition to average effects, we study dynamic treatment effects by estimating event-study models of the form

$$Y_{it} = \alpha_i + \lambda_t + \tau_t D_{it} + \varepsilon_{it}. \tag{2}$$

We then can interpret τ_t as the average treatment effect on iOS users in period t to understand how the treatment effect varies over time. In addition, this specification forms the basis for our test of pre-trends discussed in more detail in the following section. All statistical inference on estimates of Equation 1 and 2 use cluster robust standard errors clustered at the user level [Liang and Zeger, 1986].

Before estimating Equations 1 and 2 we perform a Coarsened Exact Matching (CEM) procedure that accounts for imbalance in baseline engagement in the pre-period through re-weighting [Iacus et al., 2012, Gertler et al., 2016]. The matching procedure is explained in detail in Appendix A. This weakens the identification assumption required, as we now only need common trends to hold conditional on the covariates used in matching. Results of the analysis without using the CEM weights are shown in Appendix G and the results are largely consistent.

4.4 Testing for Pre-Trends

To have a causal interpretation, the empirical strategy described in Section 4.3 relies on a common trends assumption. To be explicit, our identifying assumption is that, absent the introduction of the News tab, iOS and Android engagement would have followed the same time trends conditional on pre-treatment covariates used in matching. While this cannot be tested empirically, we can test for common trends in the pre-treatment period that would be consistent with this assumption. To do so, we perform the joint test that pre-treatment coefficients in Equation 2 are equal to zero.

When considering if a user posted in any news related community, we fail to reject the null hypothesis of common pre-trends (p=0.36). Moreover, when looking at specific topics we fail to reject the null of common pre-trends at the 5% level in all cases except for entertainment related communities, which we can reject with a p-value of 0.05 (Figure 12). For the outcome of an indicator of any low-quality post, we fail to reject the null hypotheses of common pre-trends for all 8 categories of news (Figure 13) and the same is true for the high-quality post analysis (Figure 14).

5 Results

5.1 Effect on engagement quantity

To study the impact of the News tab on engagement with news related communities, we first estimate Equation 1 where the outcome is an indicator equal to one if a user posts on any community suggested by the News tab. In aggregate, we find the News tab increases the probability of posting on any news related community by 3.5% (0.61 percentage points percentage points, p <0.01).

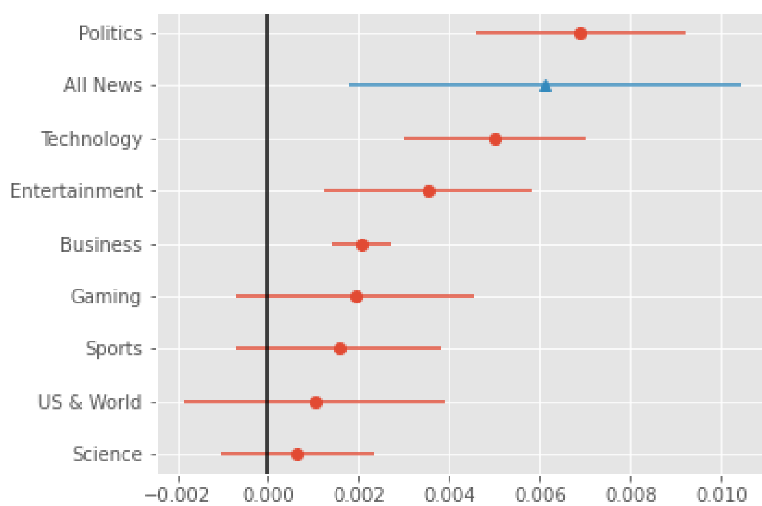


Figure 3: Treatment effect estimates on probability of posting

Note: Coefficients from estimates of Equation 1 on an indicator if a user posted in the community in a given month. Each point represents the estimated average treatment effect on iOS users on probability of engagement for the 8 categories of news. Bars represent 95% confidence intervals.

This aggregate view, however, masks substantial variation in effect sizes by the topic of news

(Figure 3). Recall there are 8 categories of news included in the News tab (US/World, Politics, Technology, Business, Science, Sports, Gaming, and Entertainment) and we can estimate the effect of the News tab on the monthly engagement probability for each category. Figure 3 plots the estimates of the average treatment effect on iOS users of the News tab on engagement with each of the 8 news categories. In all categories, the treatment effect point estimates are positive, though some are statistically indistinguishable from 0. This suggests the News tab increases total engagement, but this increase is concentrated in a subset of news categories. There is a statistically significant treatment effect for the Politics ($p<0.001$), Technology ($p<0.001$), Entertainment ($p<0.01$), and Business ($p<0.001$) categories that are also significant when correcting for multiple hypothesis testing [Holm, 1979].

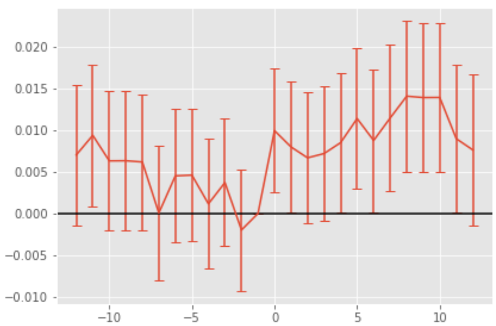
Recall these effect sizes represent the share of individuals induced by the News tab to post in each particular community in a given month. Baseline engagement rates are relatively low in this sample, with on average only 3.5% of iOS users posting in the Politics community in the year leading up to the introduction of the News tab. Therefore, a treatment effect of 0.69 percentage points represents a 19.6% increase in the *monthly* share of iOS users who post in the Politics community. There is also a 17.3% increase in the share posting in Technology communities, a 7.0% increase in the share posting in Entertainment communities, and a 106.0% increase in the share posting in Business communities.

The above analysis focuses on the extensive margin of engagement, showing that the News tab induces additional iOS users to post on news related communities relative to Android users. Next, we consider how the News tab impacts the intensive margin by estimating Equation 1 on a series of additional outcomes. Specifically, we estimate Equation 1 where the outcome is a series of indicators equal to one if total engagement in a category is above a threshold ranging from 0 to 50. This estimates the treatment effect for iOS users on the probability of a user having more posts in a month than the threshold.[11] The results of this analysis are shown in Figure 17 and Figure 18. There is a clear pattern, where the absolute treatment effects of the News tab are largest for lower thresholds suggesting the News tab induces new users to post a few times rather than inducing users to post more regularly.

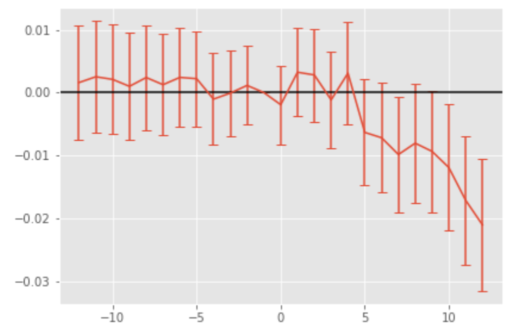
5.2 Effect on engagement quality

In addition to the quantity of engagement, we also observe the quality of engagement from upvotes and downvotes by participants in the communities. Recall that votes on Reddit are intended to be a mechanism to signal if a comment or submission contributes to the conversation or not, which we

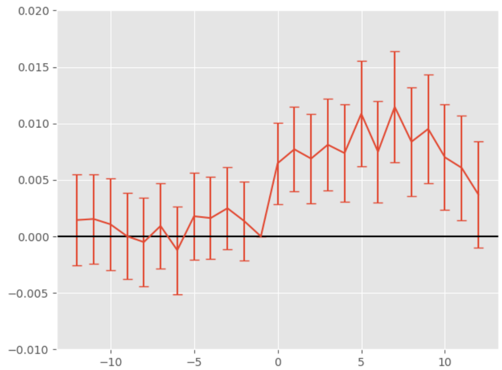
[11] A natural outcome for the extensive margin analysis would be the log-transformed number of posts, though the log-transformation suffers two pitfalls. First, given the sparsity of our dataset the arbitrary choice of how to handle zeros would be consequential. Second, and more importantly, common trends in the extensive margin (probability of posting) is inconsistent with common trends in the log-transformed outcome unless we make further strong assumptions that are rejected in the data.



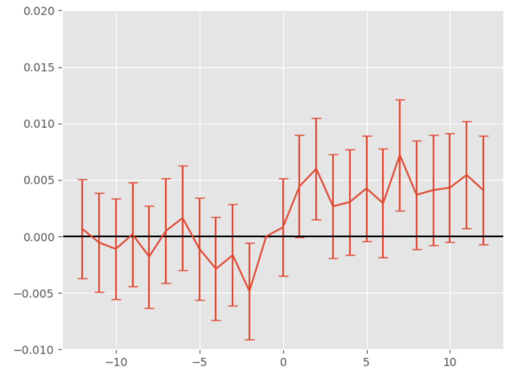
(a) All news communities



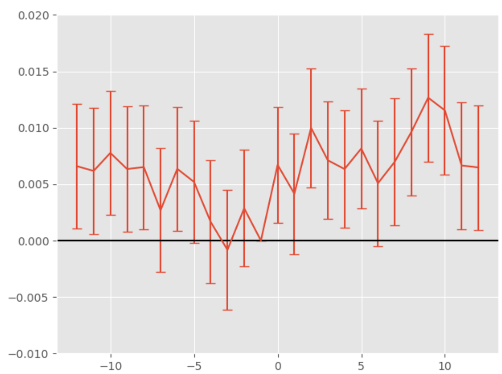
(b) All non-news communities



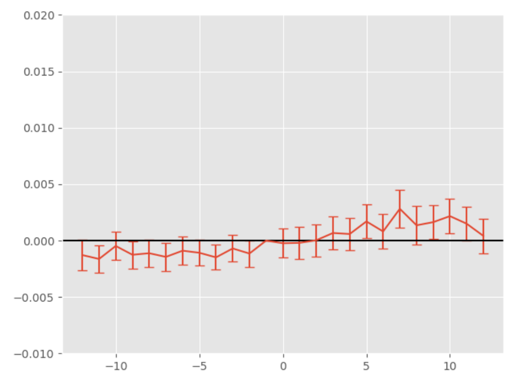
(c) Politics



(d) Technology



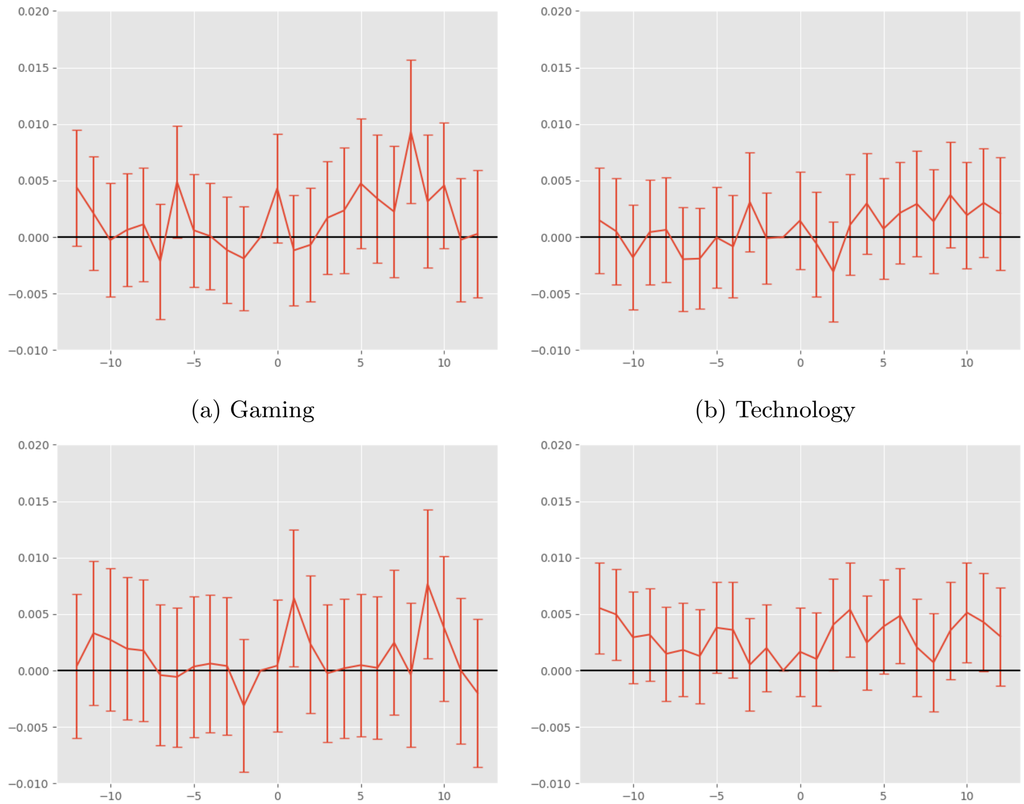
(e) Entertainment



(f) Business

Figure 4: Dynamic treatment effect estimates on probability of posting

Note: Dynamic treatment effect estimates on probability of posting in community, estimated using Equation 2. Bars represent 95% confidence intervals.



(a) Gaming (b) Technology

(c) US & World (d) Science

Figure 5: Dynamic treatment effect estimates on probability of posting

Note: Dynamic treatment effect estimates on probability of posting in community, estimated using Equation 2. Bars represent 95% confidence intervals.

are interpreting as a signal of post quality as judged by the participants in the community.

To study the impact of the News tab on engagement quality, we again estimate Equation 1 with an indicator if any of an author's posts were 'negative' or 'positive'. Here, we define a negative post to be a comment that received more downvotes than upvotes and a positive post to be a comment that received more upvotes than downvotes. Results of these regressions are plotted in Figure 6. We find that there is a heterogeneous effect on engagement quality by topic. For example, while the Politics and Technology communities see an increase in the number of users with both positive and negative posts, there is a relatively larger increase in the share of users with negative posts suggesting the marginal comments induced by the News tab are lower quality in these communities.

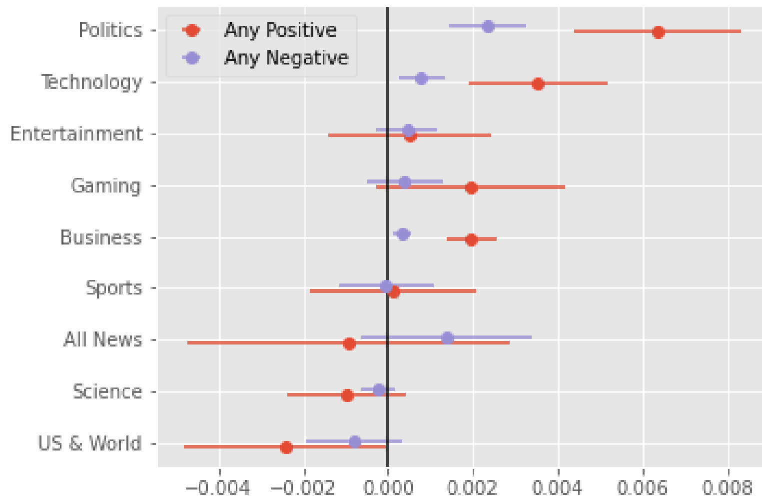


Figure 6: Treatment effect on probability of a post by post quality

Note: Coefficients from estimates of Equation 1 on an indicator if a user posted a positive (negative) post in a given month. Each point represents the estimated average treatment effect on iOS users on the probability of posting a positive or negative comment for the 8 categories of news. Bars represent 95% confidence intervals.

5.3 Effect by domain

Focusing on news categories we classify as "hard news", which includes US & World, Politics, and Business, we now investigate the effect of the News tab on the political slant of the news publishers that users engage with. In particular, we consider heterogeneous effects on engagement with news articles by publishers of varying political slant. Appendix B explains in detail how political slant is measured for each publisher.

The vast majority of threads in the news communities are started by someone sharing an article related to the community's topic. For this analysis, we drop the 4.2% of posts that link to Reddit (this is primarily general discussion threads) and YouTube (<0.1% of posts). We match publisher domains to the domain political slant measures of Robertson et al. [2018], who calculate domain level political slant of 19,022 of the most popular domains. Additional details about these data can be found in Appendix B. Over 94% of the remaining posts in hard news communities by our sample were on a thread started by a link to a publisher domain contained in the Robertson et al. [2018] slant data. We then partition the publishers into five equally sized bins based on their slant. The majority of engagement is on articles from left-leaning publishers, with less than 20% of posts on threads initiated by articles from right-leaning outlets and over half of posts on threads initiated by articles from left-leaning outlets (Figure 7).

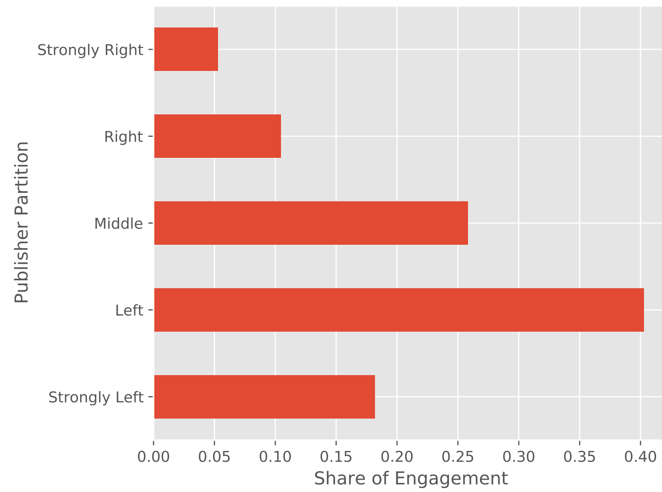


Figure 7: Distribution of posts in hard-news communities across domain slant partitions
For the set of hard news communities (US & World, Politics, Business), this figures plots the share of engagement on posts started by a publisher across the political spectrum.

Next we investigate how the News tab differentially impacts engagement on the various partitions. We find that the News tab increases engagement across the political spectrum, though point effects are largest among strongly left and moderate publishers. The increase among conservative-leaning publishers are statistically indistinguishable from zero (Figure 8).

5.4 Effect on individual engagement diversity

Turning toward the diversity of engagement, we find that the News tab induces individuals to engage with more diverse content. This is true both among the topics included in the News tab and among publisher slant partitions within "hard news" communities (Figure 9). Diversity here is operationalized as the Shannon Entropy [Shannon, 1948, Holtz et al., 2020] of engagement shares by news category and publisher slant partition, respectively. Full details of the diversity measure can be found in Appendix C where we also demonstrate the robustness of this result to other diversity measures, including the Herfindahl-Hirschman Index.

We find that the Shannon Entropy of diversity across categories of news included in the News tab increased by 12.0% and the News tab increased the Shannon Entropy of diversity across publishers from different political slants by 8.2%. To try and contextualize the magnitude of this increase in diversity, we calculated what share of the maximum possible increase in diversity we observed from the News tab. Specifically, for iOS users in the post-treatment period, we calculated the maximum possible Shannon Entropy of engagement given the total engagement amount and estimated Equa-

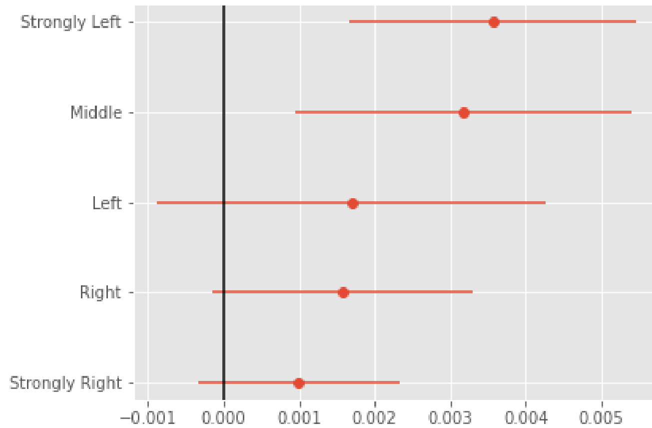


Figure 8: Treatment effects on probability of posting on a thread by publisher political slant

Note: This figure plots estimates of Equation 1 where the outcome is an indicator if the user posts on a thread in a hard news community discussing an article from a publisher within each slant partition.

tion 1 on this new outcome. This provides an upper bound on the potential increase in engagement diversity holding total engagement fixed. We find that the increase in engagement caused by the News tab represent 3.8% of the maximum increase for diversity across news categories and 7.7% of the maximum increase for diversity across publishers of varying political slant.

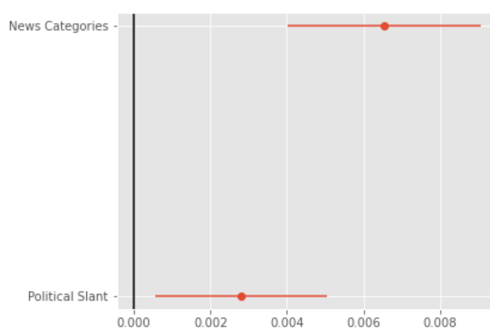
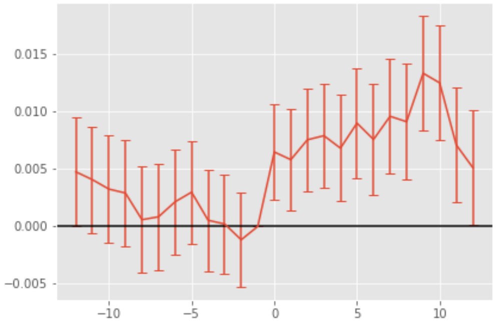
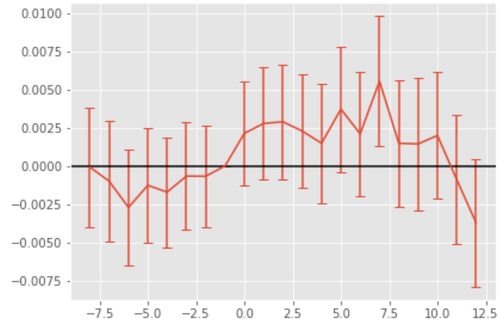


Figure 9: Effect of News tab on engagement diversity

Note: Treatment effect estimates of Equation 1 of individual engagement diversity by device, as measured through the Shannon entropy. The News Category estimate reflects the treatment effect on diversity of engagement across the 8 categories of communities on the news tab. The Political Slant estimate reflects the treatment effect on engagement diversity across political slant partitions within hard news communities.



(a) Across community engagement diversity



(b) Political slant engagement diversity

Figure 10: Individual engagement diversity time series

Note: These plot the estimates of Equation 2 of individual engagement diversity by device, as measured through the Shannon entropy. Figure 10a plots diversity of engagement across the 8 categories of communities on the news tab. Figure 10b plots engagement diversity across political slant partitions within hard news communities.

6 Heterogeneity and robustness

As shown in Section 5.1, there is substantial heterogeneity in the impact of the News tab on engagement with the various categories of news, with the Politics community seeing the largest increase in engagement. A potential explanation for this heterogeneity could be a result of the design of the News tab itself. While the News tab does have subsections for each of the 8 different categories of news content, the feed individuals first interact with aggregates popular content from across all categories of news content (Figure 1). It is plausible that users are more likely to engage with content promoted on this page. To evaluate this hypothesis, we create a monthly index of popularity for each news category and predict monthly treatment effect estimates with this index.[12] We find that this measure of popularity is correlated with the monthly treatment effect estimates on the probability of posting. While this cannot be interpreted causally, as popularity is endogenous, this evidence is consistent with the hypothesis that the heterogeneous effects of the News tab are a result of the underlying popularity of the different categories. An important implication of this hypothesis is that the choice of the popularity algorithm is critical. If this hypothesis were correct, it would mean promoting content from a different category of news would increase engagement with that category, More important, it would also suggest an algorithm that favored content slanted towards a particular political party would increase engagement with this content (e.g. Huszár et al.

[12] Unfortunately, historical data on which posts were promoted are unavailable. Therefore, we scrape a Reddit page that shows the most popular posts from a given day. See `https://www.reddit.com/r/changelog/comments/k663qy/introducing_rereddit_go_back_in_time_to_see_top/` for a description of this page. We then define the popularity index as the monthly average share of posts in the top 50 most popular news posts that came from each category of news.

[2021]), further highlighting the importance of the choice of algorithm in influencing the content individuals see and ultimately engage with.

A competing hypothesis that could explain why the largest treatment effects are in the Politics category is that the News tab was introduced in the months leading up to a U.S. midterm election. To address the concern that the our results are confounded by the election, we estimated the effect of the News tab on placebo communities that are focused on discussing politics but were not included in the News tab politics section. We find precise null effects for these communities suggesting that the effect we find for the Politics community included in the News tab is in fact a result of the News tab and not a result of the introduction of the News tab coinciding with an election cycle (Figure 16).

7 Discussion and conclusion

We find that the News tab increased engagement with news related content, though the marginal content induced by the News tab was of heterogeneous quality. In addition, the News tab increases the diversity of engagement across news categories and content from outlets with different political leanings.

This study, however, is not without limitations. First, our analysis relies on publicly observed information which requires us to limit our sample to individuals who reveal their device through the RedditMobile community. While we believe this does not impact the internal validity of our results, we must be cautious when generalizing these results as this sample of users may not be representative of the broader population. In addition, our data only contain engagement measured by posting to the communities and we do not observe content consumption.

Our findings have several important takeaways. First, we highlight the heterogeneous increase in engagement that was induced by the News tab. We found heterogeneity in two dimensions. Not all communities saw significant increases in engagement with only Politics, Technology, Entertainment, and Business related communities seeing significant increases. In addition, the increased engagement was of heterogeneous quality as some communities (namely, Politics), also saw a significant increase in the number of users posting low-quality content. This highlights an important trade-off in platform design. Existing work suggests design changes that increase engagement, in particular the diversity of engagement, in virtual communities may improve retention of the affected users [Anderson et al., 2020]. However, the new engagement may be inconsistent with the norms of the community which could have a negative externality on existing users [Gu et al., 2007].

In addition, our findings on the diversity of engagement have important policy and social implications. Previous work has found that personalized news feed algorithms and personalized recommendations have led to filter bubbles and decreased content consumption diversity [Bakshy et al., 2015, Claussen et al., 2019, Allcott et al., 2020, Levy, 2021, Holtz et al., 2020]. Moreover, these studies have found a tradeoff in engagement quantity and diversity that led Holtz et al. [2020] to introduce the concept of the "Engagement-Diversity Connection." We demonstrate that augmenting personalized feed algorithms with non-personalized feeds of news related content can increase both engagement and engagement diversity. This is true both for engagement among different categories of news as well as for the diversity of engagement among publishers of different political slant. From a managerial perspective, more diverse engagement has been shown to positively impact user retention [Anderson et al., 2020]. From a policy perspective, users consuming and engaging with a more diverse set of political viewpoints has important positive implications for civic life [Sunstein, 2003]. This suggests that non-personalized feeds can be an important tool in mitigating algorithmic filter bubbles and should be studied further to assess the robustness of these findings.

	Any Post	Number of Posts			Any Neg Score		Any Pos Score	
	Mean	Mean	Min	Max	Mean	P(any|post)	Mean	P(any|post)
All Posts	0.749	749.643	0	93,953	0.575	0.769	0.748	0.999
All News	0.466	60.053	0	15,140	0.216	0.464	0.460	0.985
US/World	0.268	8.627	0	4,786	0.099	0.370	0.261	0.972
Politics	0.150	9.193	0	9,759	0.052	0.348	0.144	0.960
Technology	0.179	1.919	0	933	0.041	0.230	0.173	0.968
Science	0.152	1.606	0	2,569	0.027	0.174	0.148	0.969
Sports	0.170	25.753	0	14,616	0.062	0.366	0.166	0.976
Gaming	0.231	7.911	0	5,799	0.064	0.279	0.226	0.979
Entertainment	0.218	4.751	0	4,585	0.058	0.265	0.212	0.971
Business	0.022	0.292	0	2,774	0.004	0.189	0.021	0.950

Table 1: Sample summary statistics

	(1)	(2)	(3)
Popularity	0.036	0.019	0.052
	(0.010)	(0.008)	(0.055)
p-value	<0.001	0.012	0.344
Category FEs		X	X
Category Interactions (centered)			X
Obs	104	104	104
Adj. R^2	0.112	0.603	0.579
F-stat	13.641	28.026	15.333

Table 2: Correlation in Community Popularity and Treatment Effects

Note: Coefficients from the regression of monthly treatment effect estimate on community popularity. Model (2) adds community fixed effects. Model (3) adds community fixed effects and the interaction of centered community fixed effects with popularity. This results is equivalent to averaging the estimates.

References

A. Abadie and M. D. Cattaneo. Econometric methods for program evaluation. *Annual Review of Economics*, 10:465–503, 2018.

H. Allcott, L. Braghieri, S. Eichmeyer, and M. Gentzkow. The Welfare Effects of Social Media. *American Economic Review*, 110, 2020. ISSN 0002-8282. doi: 10.1257/aer.20190658.

A. Anderson, L. Maystre, I. Anderson, R. Mehrotra, and M. Lalmas. Algorithmic effects on the diversity of consumption on spotify. In *Proceedings of The Web Conference 2020*, pages 2155–2165, 2020.

E. Bakshy, S. Messing, and L. A. Adamic. Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239):1130–1132, June 2015. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aaa1160. URL `https://science-sciencemag-org.libproxy.mit.edu/content/348/6239/1130`. Publisher: American Association for the Advancement of Science Section: Report.

J. Baumgartner, S. Zannettou, B. Keegan, M. Squire, and J. Blackburn. The pushshift reddit dataset. In *Proceedings of the international AAAI conference on web and social media*, volume 14, pages 830–839, 2020.

C. Budak, S. Goel, and J. M. Rao. Fair and balanced? quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, 80(S1):250–271, 2016.

G. Burtch, Y. Hong, R. Bapna, and V. Griskevicius. Stimulating online reviews by combining financial incentives and social norms. *Management Science*, 64(5):2065–2082, 2018.

G. Burtch, Q. He, Y. Hong, and D. Lee. How do peer awards motivate creative content? experimental evidence from reddit. *Management Science*, 2021.

L. Cabral and L. Li. A dollar for your thoughts: Feedback-conditional rebates on ebay. *Management Science*, 61(9):2052–2063, 2015.

Y. Chen, F. M. Harper, J. Konstan, and S. X. Li. Social comparisons and contributions to online communities: A field experiment on movielens. *American Economic Review*, 100(4):1358–98, 2010.

J. Claussen, C. Peukert, and A. Sen. The editor vs. the algorithm: Returns to data and externalities in online news. *Working Paper*, 2019.

E. Dujeancourt, M. Garz, A. Ghose, J. Hagen, J. Lischka, M. Nordin, J. Rickardsson, and M. Schwarz. The effects of algorithmic content selection on user engagement with news on twitter. *Working Paper*, 2021.

D. Gaffney and J. N. Matias. Caveat emptor, computational social science: Large-scale missing data in a widely-published reddit corpus. *PloS one*, 13(7):e0200162, 2018.

J. Gallus. Fostering public good contributions with symbolic awards: A large-scale natural field experiment at wikipedia. *Management Science*, 63(12):3999–4015, 2017.

P. J. Gertler, S. Martinez, P. Premand, L. B. Rawlings, and C. M. Vermeersch. *Impact evaluation in practice*. World Bank Publications, 2016.

P. B. Goes, C. Guo, and M. Lin. Do incentive hierarchies induce user effort? evidence from an online knowledge exchange. *Information Systems Research*, 27(3):497–516, 2016.

B. Gu, P. Konana, B. Rajagopalan, and H.-W. M. Chen. Competition among virtual communities and user valuation: The case of investing-related communities. *Information systems research*, 18 (1):68–85, 2007.

S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pages 65–70, 1979.

D. Holtz, B. Carterette, P. Chandar, Z. Nazari, H. Cramer, and S. Aral. The engagement-diversity connection: Evidence from a field experiment on spotify. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 75–76, 2020.

K. Hosanagar, D. Fleder, D. Lee, and A. Buja. Will the global village fracture into tribes? recommender systems and their effects on consumer fragmentation. *Management Science*, 60(4): 805–823, 2014.

F. Huszár, S. I. Ktena, C. O'Brien, L. Belli, A. Schlaikjer, and M. Hardt. Algorithmic amplification of politics on twitter. *arXiv preprint arXiv:2110.11010*, 2021.

S. M. Iacus, G. King, and G. Porro. Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1):1–24, 2012.

S. Kemp. Digital 2020 reports. *"wearesocial.com/digital-2020*, 2020.

W. Khern-am nuai, K. Kannan, and H. Ghasemkhani. Extrinsic versus intrinsic rewards for contributing reviews in an online platform. *Information Systems Research*, 29(4):871–892, 2018.

A. D. Kramer, J. E. Guillory, and J. T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24): 8788–8790, 2014.

D. Lee and K. Hosanagar. How do recommender systems affect sales diversity? a cross-category investigation via randomized field experiment. *Information Systems Research*, 30(1):239–259, 2019.

R. Levy. Social media, news consumption, and polarization: Evidence from a field experiment. *American economic review*, 111(3):831–70, 2021.

K.-Y. Liang and S. L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.

G. Oestreicher-Singer and A. Sundararajan. Recommendation networks and the long tail of electronic commerce. *Mis quarterly*, pages 65–83, 2012.

G. Oestreicher-Singer and L. Zalmanson. Content or community? a digital business strategy for content providers in the social age. *MIS quarterly*, pages 591–616, 2013.

E. Pariser. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

M. Restivo and A. Van De Rijt. Experimental study of informal rewards in peer production. *PloS one*, 7(3):e34358, 2012.

S. A. Rhoades. The herfindahl-hirschman index. *Fed. Res. Bull.*, 79:188, 1993.

R. E. Robertson, S. Jiang, K. Joseph, L. Friedland, D. Lazer, and C. Wilson. Auditing partisan audience bias within google search. *Proceedings of the ACM on Human-Computer Interaction*, 2 (CSCW):1–22, 2018.

C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27 (3):379–423, 1948.

C. Sunstein. R. 2007. republic. com 2.0, 2003.

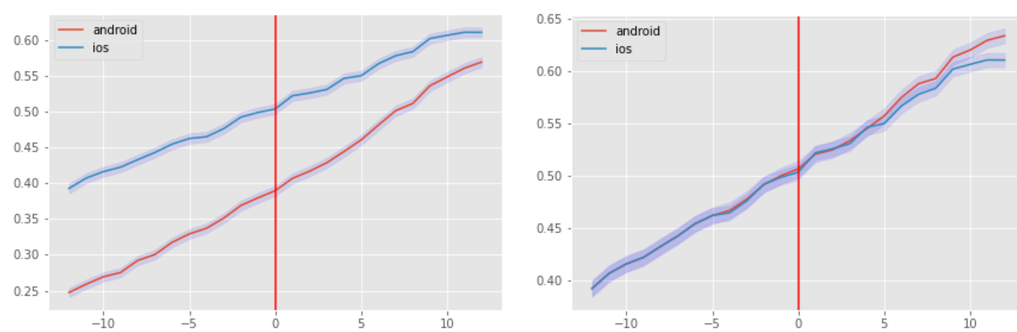
S. Vosoughi, D. Roy, and S. Aral. The Spread of True and False News Online. *Science*, 359(6380): 1146–1151, 2018.

Appendix A Matching Procedure

Our analysis of the impact of the News tab relies on a common trends assumption to have a causal interpretation. Non-weighted engagement trends, i.e. those estimating Equation 2 without the Coarsened Exact Matching (CEM) procedure are largely consistent with this hypothesis with one notable exception where the probability of posting on any non-news community shows a clear pre-trend (Figure 21 and Figure 22). We use the CEM procedure outlined in Iacus et al. [2012] to mitigate this imbalance.

There are 12 pre-treatment periods in our analysis and we match on the first 6 of these pre-treatment periods, holding out the following 6 to provide further evidence our common trends assumption is plausible. We define $X_i = (x_{i,-12}, \ldots, x_{i,-7})'$ as a vector of dummy variables $x_{i,t}$ equal to one if user i made any posts (on any community, not just news) in period t. We then generate weights from CEM following Iacus et al. [2012]. Given the relatively low-dimensionality of this matching exercise we are able to find exact matches and this is equivalent to subclassification. We then use these weights in our difference-in-differences framework, which requires the assumption of common trends conditional on the covariates used in matching. In other words, we assume that iOS engagement trends would have followed Android engagement trends absent the introduction of the News tab, *conditional* on X_i.

To demonstrate the effectiveness of the matching procedure, we can plot the unweighted and weighted probability of posting any post on Reddit over time by treatment group (Figure 11). While difficult to see in the raw time series figure below, there is clear evidence of differential pre-trends between iOS and Android users (Figure 21). Re-weighting eliminates this imbalance and it does so even in the periods before treatment that are not used in weighting.



(a) Unweighted (b) Weighted

Figure 11: Probability of any post by treatment group

Appendix B Data on domain political slant

We obtain political slant by domain using Robertson et al. [2018]. To generate this score, Robertson et al. [2018] collect recent tweets containing links from known Democrats and Republicans. The slant measure is then calculated as the difference in the probability of a sharing a domain conditioned on being republican who has shared at least one domain less the same conditional probability for democrats, normalized to be between -1 and 1. Formally, the slant measure is define as

$$\text{bias-score}(i) = \frac{\frac{r_i}{\sum_{j \in I} r_j} - \frac{d_i}{\sum_{j \in I} d_j}}{\frac{r_i}{\sum_{j \in I} r_j} + \frac{d_i}{\sum_{j \in I} d_j}}$$

where r_j (d_j) is the number of unique Republicans (Democrats) who shared domain j and I is the set of all domains. This measure is equal to 0 if shared by equal shares of Republicans and Democrats and equal to 1 (-1) if it was shared only by Republicans (Democrats). Robertson et al. [2018] demonstrate their measure of domain slant agrees with several existing measures of publisher slant [Bakshy et al., 2015, Budak et al., 2016]

Appendix C Measures of information diversity

Our measures of engagement diversity first partitions engagement into K bins based on either the category of news community or the political slant of the publisher within hard-news communities as calculated in Robertson et al. [2018]. We then operationalize the diversity of engagement following Holtz et al. [2020], measuring individual-level engagement diversity using Shannon entropy [Shannon, 1948]. The Shannon entropy of user i's engagement is defined as

$$id_i = -\sum_{k=1}^{K} s_{ki} \ln(s_{ki}), \tag{3}$$

where s_{ki} is the share of user i's posts on a thread based on an article from publishers in partition $k \in \{1, \ldots, K\}$. If $s_{ki} = 0$, the partition's contribution to the Shannon entropy is zero which implies users who do not engage with any posts have $id_i = 0$ [Holtz et al., 2020].

We also operationalize engagement diversity using the Herfindahl–Hirschman Index (HHI) [Rhoades, 1993]. This measure is defined as the sume of squared engagement shares:

$$\mathrm{HHI}_i = \sum_{k=1}^{K} s_{ki}^2.$$

When an individual has no consumption in a period, we define the HHI to be equal to one which is equivalent to engaging entirely with content from a single category.

Appendix D Testing for common pre-trends

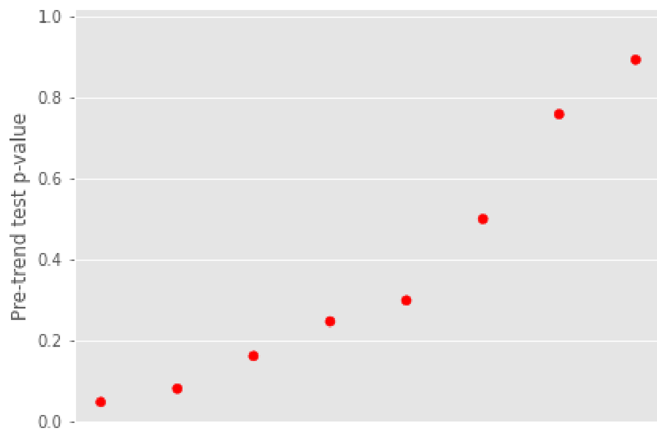


Figure 12: Common pre-trends p-values for indicator of any post

Plot of p-values of test of common pre-trends after CEM matching. The test is the joint test that all pre-period treatment effect coefficients in Equation 2 are equal to 0.

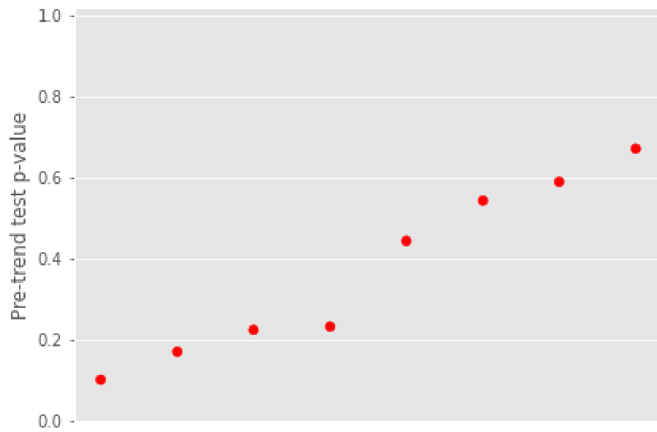


Figure 13: Common pre-trends p-values for indicator of low quality post

Plot of p-values of test of common pre-trends after CEM matching. The test is the joint test that all pre-period treatment effect coefficients in Equation 2 are equal to 0.

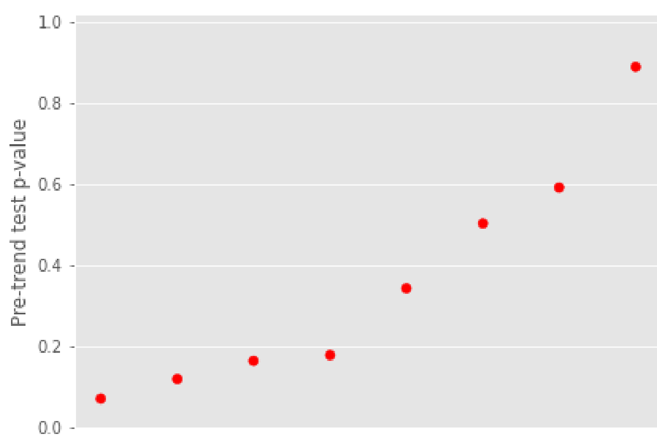


Figure 14: Common pre-trends p-values for indicator of high quality post

Plot of p-values of test of common pre-trends after CEM matching. The test is the joint test that all pre-period treatment effect coefficients in Equation 2 are equal to 0.

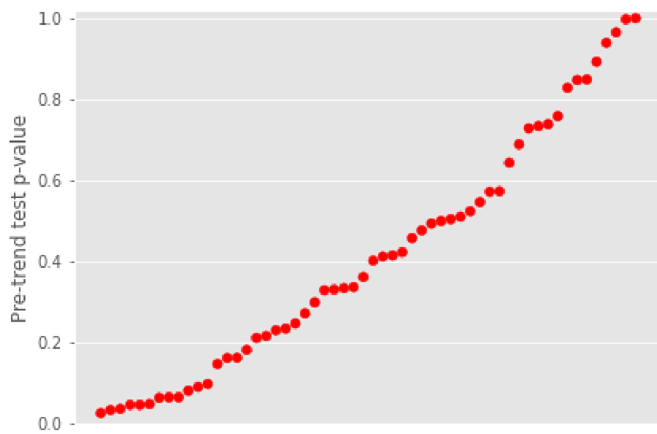


Figure 15: Common pre-trends p-values for intensive margin analysis

Plot of p-values of test of common pre-trends after CEM matching. Each point represents the p-value of a joint test that all pre-period treatment effect coefficients in Equation 2 are equal to 0. The figure plots the p-values for all thresholds shown in Figure 17 and Figure 18 across all of the categories included in the News tab.

Appendix E Robustness checks

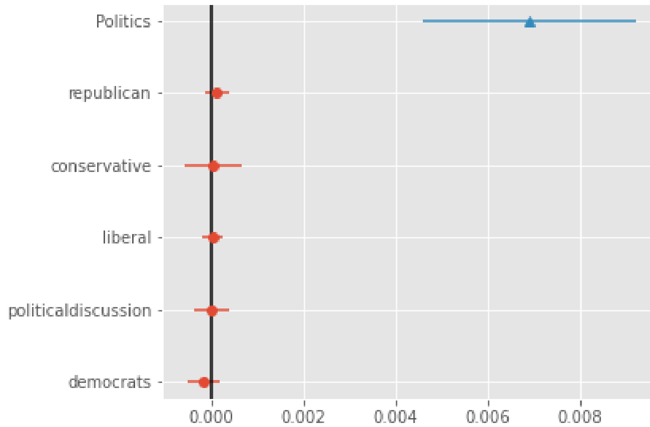
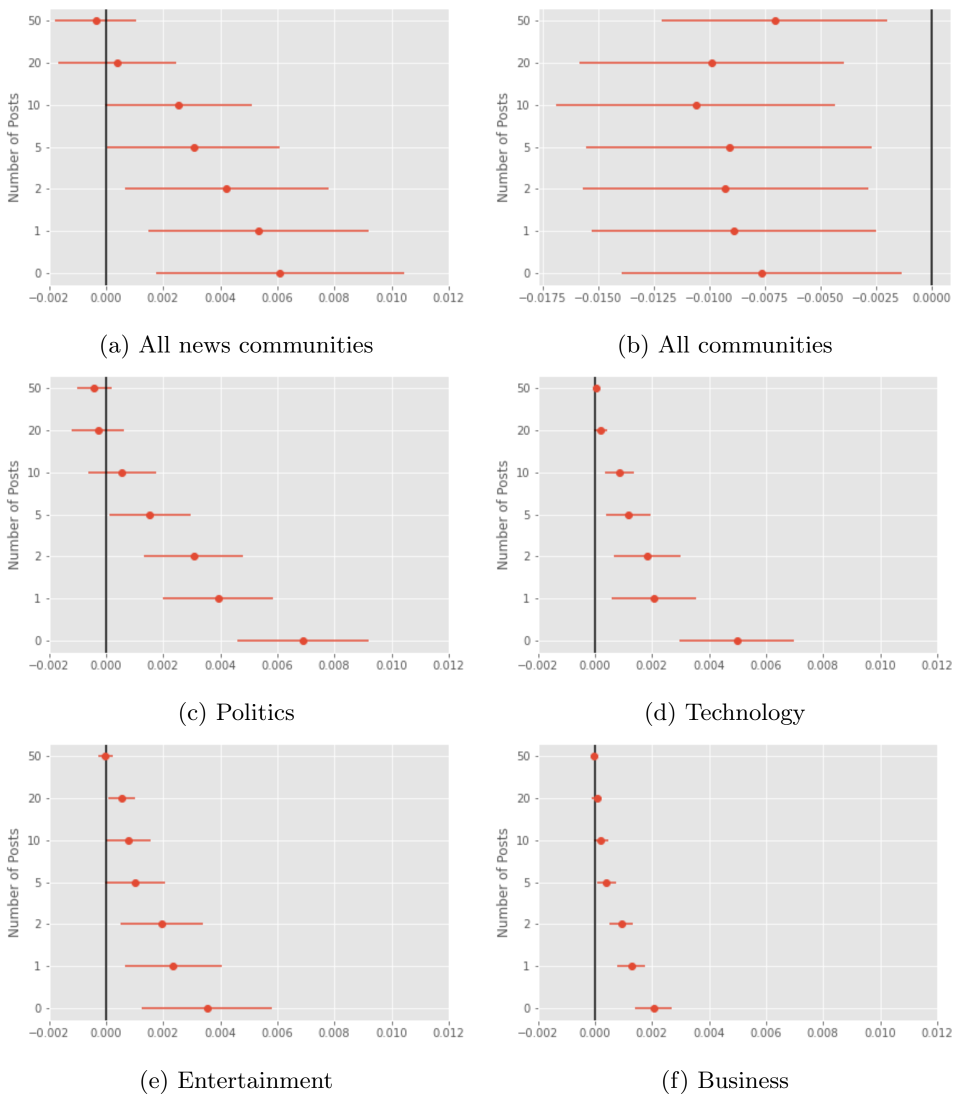


Figure 16: Election Placebo Tests

Here we show results of a placebo test on political communities not included in the News tab (red) compared to the Politics community included in the News tab (blue). We find no treatment effect on communities excluded from the News tab suggesting the effect is not confounded by the 2018 midterm election.

Appendix F Effect on intensive margin of engagement



(a) All news communities

(b) All communities

(c) Politics

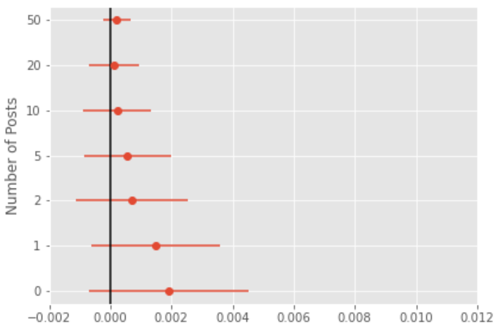
(d) Technology

(e) Entertainment

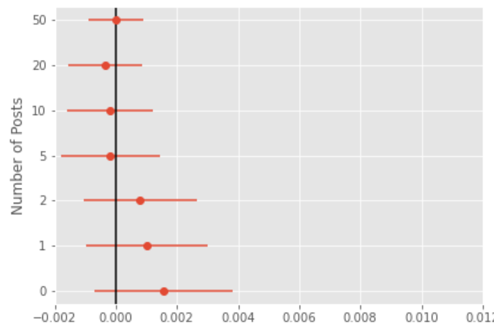
(f) Business

Figure 17: Treatment effect estimates on probability of posting more than k posts

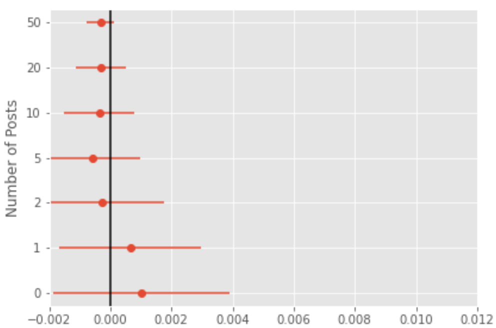
Each plot shows the treatment effect estimate from Equation 1 where the outcome is an indicator if the user posted more than a threshold posts in a month (within each community). The thresholds are shown on the y-axis. Horizontal bars represent 95% confidence intervals



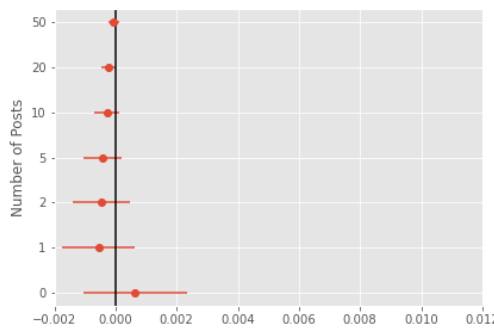
(a) Gaming



(b) Sports



(c) US & World



(d) Science

Figure 18: Treatment effect estimates on probability of posting more than k posts

Each plot shows the treatment effect estimate from Equation 1 where the outcome is an indicator if the user posted more than a threshold posts in a month (within each community). The thresholds are shown on the y-axis. Horizontal bars represent 95% confidence intervals

Appendix G Analyses without CEM weights

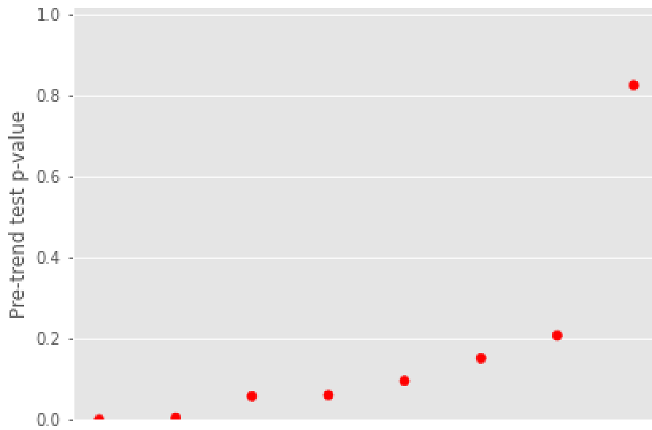


Figure 19: Common pre-trends p-values

Plot of p-values of test of common pre-trends without CEM matching. The test is the joint test that all pre-period treatment effect coefficients in Equation 2 are equal to 0.

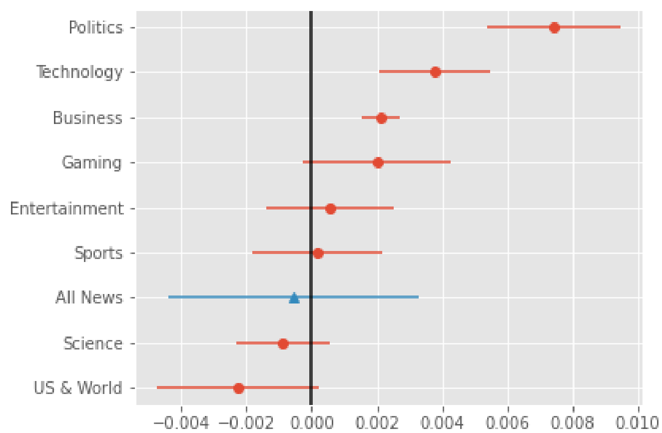
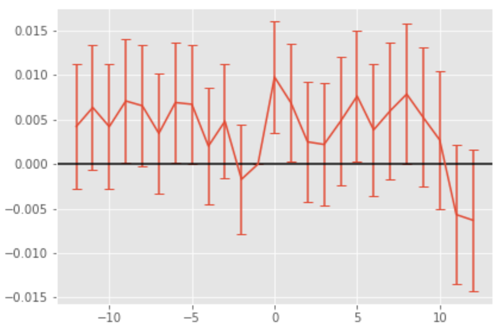
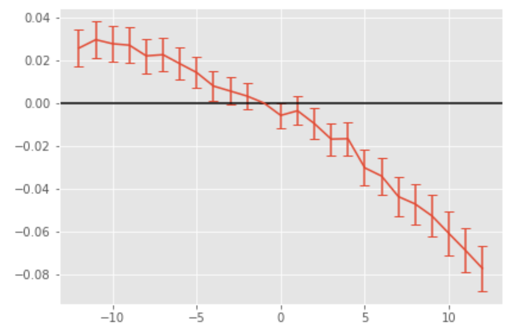


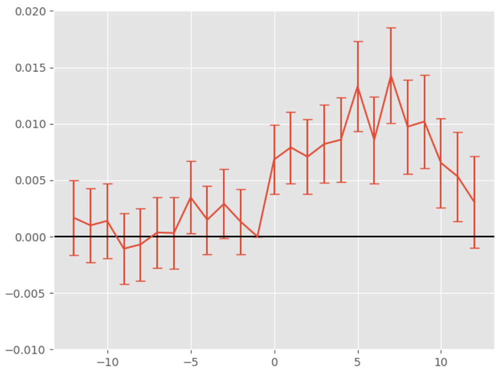
Figure 20: Treatment effect estimates on probability of posting



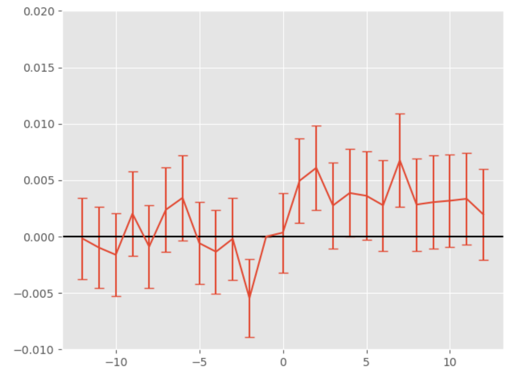
(a) All news communities



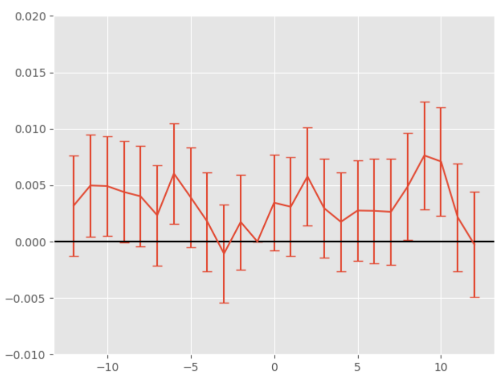
(b) All non-news communities



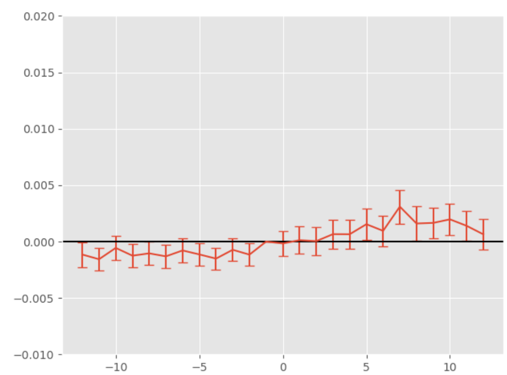
(c) Politics



(d) Technology



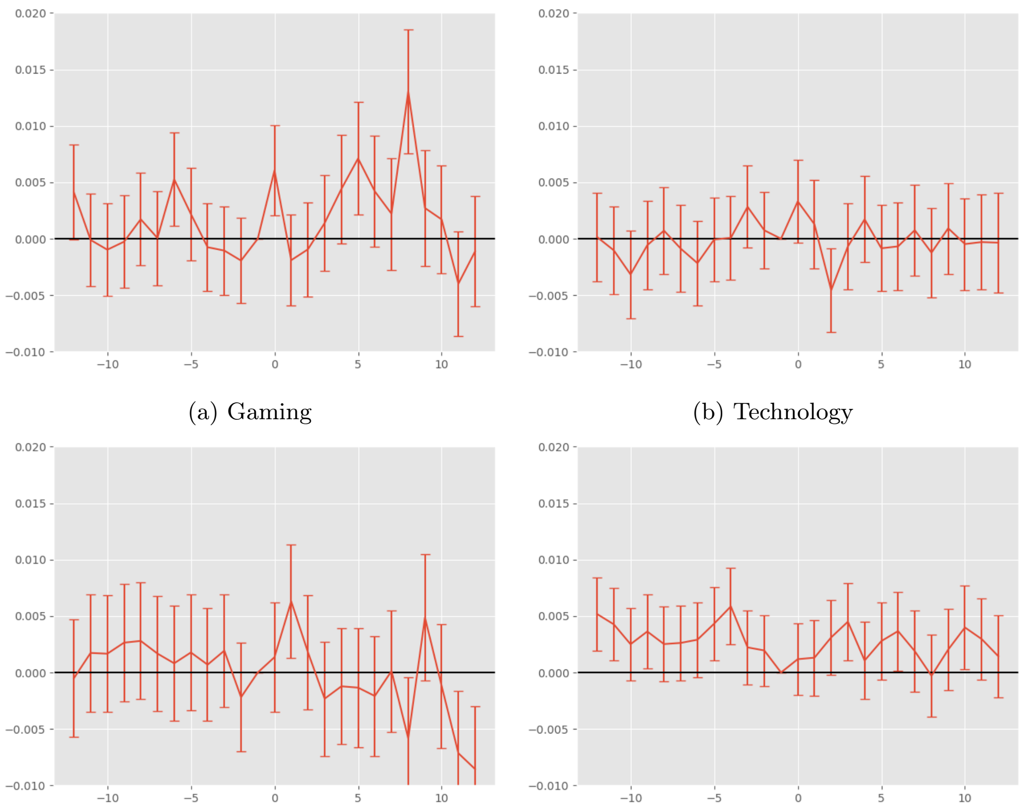
(e) Entertainment



(f) Business

Figure 21: Dynamic treatment effect estimates on probability of posting

Dynamic treatment effect estimates on probability of posting in community, estimated using Equation 1 without the Coarsened Exact Matching weights.



(a) Gaming

(b) Technology

(c) US & World

(d) Science

Figure 22: Dynamic treatment effect estimates on probability of posting

Dynamic treatment effect estimates on probability of posting in community, estimated using Equation 1 without the Coarsened Exact Matching weights.