Platforms as Editors: How Algorithmic Curation Shapes Online News Supply

Yikang Shen *

Abstract

Recommendation algorithms shape not only how content is distributed but also what producers choose to supply. This paper provides empirical evidence of such supply responses and quantifies their implications for platform design. Combining matched print and online headlines from major newspapers with over 3 million Facebook posts, I document that online headlines are more ideologically slanted than their print counterparts. Exploiting Facebook's 2025 restoration of political content, I find that the online–print slant gap for political relative to non-political news increases by about 81%, revealing a substantial supply response to the change in algorithmic incentives. I then develop a structural model of user demand, algorithmic recommendations, and editorial choices. The model shows that media outlets adjust slant primarily to attract algorithmic exposure, with supplied slant accounting for approximately 80% of the extra slant in viewed content. Counterfactuals show that removing algorithmic amplification eliminates roughly 90% of this supply-induced slant. These findings show that evaluating platform algorithms requires accounting for their effects on content supply, rather than treating the available content as fixed.

Keywords: platform algorithms, endogenous content supply, media economics.

*Carnegie Mellon University (Tepper School of Business) · yikangs@andrew.cmu.edu. I am grateful to Jesse Shapiro and Ro'ee Levy for generously sharing research materials. I thank Maryam Saeedi, Ananya Sen, Ali Shourideh, Ariel Zetlin-Jones, Gregory Martin, Lara Marie Berger, Julia Cagé, Gérard Padró i Miquel, Nicolas Longuet-Marx, and Alexey Makarin and participants at the NBER Political Economy Meeting, ZEW Conference on the Economics of ICT, Econometric Society North American Summer Meeting, and ESIF Economics and AI+Machine Learning Meeting for helpful comments and discussions.

1 Introduction

Recommendation algorithms play a central role in determining which products and information consumers encounter on digital platforms. Much of the existing empirical evidence evaluates how these algorithms affect exposure and consumption, taking the available content as given. Yet content producers have incentives to adjust their products in response to how algorithms allocate attention. In traditional markets, for example, prior work has shown that outlets adjust their editorial slant in response to consumer demand for ideologically aligned content (Gentzkow and Shapiro, 2010). As a result, changes in recommendation algorithms can affect not only which content is shown to users, but also the characteristics of the content available for recommendation. Understanding this supply response is essential for evaluating the equilibrium effects of platform design. This paper provides empirical evidence of this mechanism in the online news market. News outlets compete for attention on platforms such as Facebook, where recommendation algorithms determine the visibility of their content. Unlike traditional media markets, where editorial decisions primarily reflect consumer demand, online news producers must also account for the incentives created by algorithmic recommendations. I ask whether changes in these incentives affect the ideological framing of news and how much these supply responses contribute to the information users ultimately encounter.

The answers to these questions are central to understanding how social media and online news markets should be regulated—domains that have become integral to the functioning of modern democracies. It has been well documented empirically that news content can have significant effects on political attitudes and outcomes (Strömberg, 2004, Gentzkow, 2006). In addition, slanted news has been shown to have persuasive effects on ideologies and voting behaviors (DellaVigna and Kaplan, 2007, Martin and Yurukoglu, 2017, Alesina et al., 2020, Djourelova, 2023). Particularly on social media platforms, more exposure to content expressing antidemocratic attitudes or partisan animosity has been shown to increase affective polarization and elevate users' negative emotions (Piccardi et al., 2025). When such

content is more prevalent on the platform, users are more likely to encounter and consume it, amplifying risks of polarization and psychological distress. These patterns underscore the importance of understanding how news supply responds to algorithmic design.

The central findings of this paper are twofold. First, online news headlines are, on average, more ideologically slanted and express more negative emotion than their print counterparts across major newspapers. Using Facebook's 2025 restoration of political content, I show that this divergence responds to platform incentives: after the policy change, outlets substantially increase the extent to which online headlines depart from their print counterparts for political relative to non-political news. Second, users display a clear preference for ideologically aligned and unexpectedly slanted content, while the recommendation algorithm amplifies these forces when allocating attention. Media outlets respond strategically by increasing the slant of their online content to compete for platform-driven viewership. Counterfactual simulations show that removing algorithmic amplification would eliminate approximately 90 percent of the supply-induced slant generated by these incentives.

To generate these results, I assemble a new dataset by matching print and online headlines from eleven major newspapers between 2017 and 2025, and link them to the corresponding Facebook posts and measures of user consumption. Then, I measure each headline's ideological slant using a phrase-frequency measure based on the *Congressional Record* (following Gentzkow and Shapiro, 2010), and its emotional level as sentiment using a pretrained machine learning model. The matched data reveal persistent differences in how the same news stories are presented online and in print. Online headlines are systematically more negative and, on average, more ideologically pronounced than their print counterparts. At the same time, the slant of both versions varies substantially with the underlying news environment. This motivates my focus on the within-article online–print gap rather than the level of online slant alone: comparing the two versions of the same story helps separate changes in the underlying news being covered from changes in how that news is presented online. I focus on headlines and posts rather than the full articles they link to because these are themselves

important news products consumed on social media. Sundar et al. (2025), for example, find that roughly 75% of Facebook posts linking to news are shared without users first clicking through to the linked web page, with this behavior especially pronounced for extreme and ideologically aligned political content.

I next provide direct evidence that media outlets adjust their online news supply in response to algorithmic incentives. The descriptive online–print gap alone cannot distinguish platform incentives from differences in audience demand. To isolate the role of the platform, I exploit Facebook's 2025 restoration of political content. Because the policy directly altered the distribution of political content, I compare changes in the absolute online–print slant gap for political and non-political stories before and after the implementation period. Following the restoration, the online–print slant gap increases substantially more for political than for non-political stories. Using my preferred political-content classifier, the estimated increase is about 0.021, approximately 81% of the pre-policy difference. The result is similar using both political-content classifiers, while there is little corresponding change in sentiment divergence. Together, these results provide direct evidence that platform recommendation incentives affect the ideological content supplied by news outlets online.

The regression analysis establishes that platform incentives affect how much online headlines depart from their print counterparts, but it does not explain how outlets choose the direction of those changes. To study this editorial choice, I develop a structural model that jointly estimates user preferences, the platform's recommendation rule, and media outlets' dynamic strategies for slant provision. With the model, I show that readers prefer both like-minded and unexpected news. They are more likely to engage with posts whose slant matches their own ideological orientation, consistent with Nyhan et al. (2023) and Braghieri et al. (2024). At the same time, I find that users respond positively to news that departs from an outlet's usual ideological tone. For example, engagement is higher when the New York Times publishes a more conservative post than usual, or when Fox News publishes a more liberal one. I interpret this pattern as a reduced-form response to unexpected slant,

which may arise through several mechanisms. First, deviations from an outlet's typical position may be more salient or surprising and therefore attract greater attention, consistent with Bordalo et al. (2013) and Bordalo et al. (2022). Second, cross-ideological content may appear more credible or informative because it runs against the outlet's perceived ideological incentives, related to the mechanism in Chiang and Knight (2011). Third, users may make strategic inferences from such deviations: Huang et al. (2026) show that readers react more positively to news that is inconsistent with a media outlet's usual stance, consistent with interpreting such reports as containing stronger underlying information. Treating the platform algorithm as an abstract decision maker that selects which news to display, I estimate that Facebook's recommendation system introduces substantial additional incentives for politically slanted content relative to what would arise from user preferences alone. On the supply side, newspapers choose the ideological slant of their social-media posts to maximize the probability of being recommended to users, with smaller outlets placing greater weight on visibility than larger ones. These supply responses account for about 80% of slant in user exposure. Counterfactual simulations show that removing the algorithm's amplification of ideological signals would eliminate roughly 90% of the extra polarization between online and print headlines. More importantly, such a policy substantially flattens the distribution of content exposure across users with different ideologies while generating only modest welfare losses. In contrast, an intervention similar to Facebook's algorithm update produces comparable welfare losses but has limited effects on exposure diversity. Together, these results show that recommendation systems do not merely mediate information—they actively shape the equilibrium supply of news.

Methodologically, this paper extends tools from structural industrial organization to the domain of media economics. Because the market includes many competing outlets, I adopt the concept of an oblivious equilibrium (Weintraub et al., 2008), which allows each outlet to optimize against aggregate market conditions without tracking rivals individually—a tractable and behaviorally realistic framework for large digital platforms. In traditional me-

dia markets, like cable news or newspapers in a given geographic area, typically there are two to three major outlets. However, in the online market like Facebook, all outlets–cable, broadcast, newspapers, influencers–are competing with each other, none of which has a dominant share in the market. For estimation, I extend the share inversion method of Berry et al. (1995) to a setting with algorithmic decision and two demand layers: one for recommendation and one for engagement. Then I combine it with a simulated conditional choice probability approach (Hotz et al., 1994) for the dynamic editorial problem. Together, this framework provides the first quantitative decomposition of how much of online polarization arises from user demand versus algorithmic design.

This paper contributes to four strands of literature. First, the paper contributes to a growing literature on how platform design affects the incentives and supply decisions of content producers (Bhargava, 2022, Qian and Jain, 2024, Saeedi et al., 2024). Most of this literature is theoretical, while empirical evidence on supply responses to platform design remains limited. One recent exception is Zeng et al. (2025), who use a randomized field experiment on a video-sharing platform to show that recommendation quality affects the quantity of content users produce. This evidence primarily concerns the extensive margin of content creation. Much less is known about the intensive margin—whether platform incentives change the characteristics of the content that producers choose to supply. I study this margin in the news market, showing how recommendation incentives affect the ideological framing of a given underlying story. By jointly modeling users, the platform, and media outlets, I quantify how these supply responses feed back into equilibrium news exposure.

Second, this paper contributes to the literature on news production, editorial strategy, and media competition. Prior work shows that media firms tailor content in response to consumer demand, advertising incentives, competitive pressures, and measures of online audience attention (Gentzkow et al., 2014, Sen and Yildirim, 2016, Allcott and Gentzkow, 2017, Cagé et al., 2019, Beattie et al., 2021, Leung and Strumpf, 2023). Recent work also highlights the influence of social media on mainstream media coverage decisions (Cagé et al., 2024),

showing that events receiving more attention on social media are more likely to be covered by traditional media. In contrast, this paper focuses on the intensive margin of editorial choice. Conditional on covering a given event, I show that media outlets systematically adjust the language and ideological slant of content to capture attention from platform algorithms and online users. To study this, I develop a dynamic, article-level model of editorial choice in which outlets optimally choose slant for each piece of content. Editorial decisions trade off algorithmic visibility, user engagement, and the evolution of outlet reputation over time. This framework allows me to quantify the equilibrium effects of platform algorithms on both the supply and distribution of news, and to evaluate counterfactual platform designs. In this sense, platform algorithms become a key input into newsroom decision-making, analogous to prices, advertising demand, or competition in earlier media environments.

Third, the paper contributes to the literature on algorithms, social media, and news dissemination. Existing work shows that ranking and recommendation systems shape exposure, engagement, and polarization by determining which content users see (Bakshy et al., 2015, Levy, 2021, Calzada and Gil, 2020, Moehring, 2024). This paper shifts the focus from exposure to production. Exploiting Facebook's 2025 restoration of political content as a quasi-experimental shock, I provide direct evidence that media outlets adjust the slant of their online headlines in response to changes in platform incentives. I then develop a structural model that separately identifies user preferences, algorithmic amplification, and outlet responses. This allows me to quantify how recommendation algorithms affect not just which articles are seen, but how articles are presented.

Finally, it relates to a large literature on the demand for slanted media, which shows that consumers prefer ideologically congenial information and that media firms respond to these preferences in equilibrium (Mullainathan and Shleifer, 2005, Gentzkow and Shapiro, 2006, Gentzkow and Shapiro, 2010, Martin and Yurukoglu, 2017, Chiang and Knight, 2011, Guess, 2021). Recent work using article-level data emphasizes that much of the variation in slant arises within outlets rather than across them, highlighting the importance of studying

content at the level of individual articles (Braghieri et al., 2024). Relatedly, Huang et al. (2026) show that users respond not only to ideological alignment but also to whether news is consistent with a media outlet's usual stance. Building on these insights, I allow user engagement to depend separately on ideological alignment and deviations from an outlet's typical positioning. More importantly, I demonstrate that these demand-side forces do not only shape consumption, but also feed back into the equilibrium supply of news slant through algorithmic curation.

Taken together, the evidence suggests that social-media algorithms play a quantitatively important role in shaping the supply of news. By altering the returns to ideologically slanted content, platform design changes both the distribution and the production of news. Evaluating platform algorithms and related policies therefore requires accounting for equilibrium supply responses, rather than treating the set of available content as fixed.

2 Institutional Overview: News on Social Media

This section summarizes how news distribution and consumption on social media evolved from 2017 to 2025 and how major newspapers adapted organizationally by hiring search engine optimization (SEO) editors and social media editors.

2.1 From Web & Print to Social Distribution

In 2017, social media was already central to news access: two-thirds of U.S. adults reported getting at least some news from social platforms, with 20% doing so often.[1] Platform composition changed over time. Using a 2023 Pew survey reported in 2024, about 30% of U.S. adults said they regularly get news on Facebook, with smaller shares on Instagram (16%), TikTok (14%), and X/Twitter (12%).[2] By late 2024, Pew additionally documents the rise

[1] Pew Research Center (Sept. 7, 2017), "News Use Across Social Media Platforms 2017."

[2] Pew Research Center (June 12, 2024), "How Americans Get News on TikTok, X, Facebook and Instagram."

of "news influencers": about 21% of U.S. adults (and 37% of those aged 18–29) report regularly getting news from influencers on social media.[3] These trends underscore a structural shift: audience discovery increasingly occurs in feeds governed by platform algorithms, not on outlet homepages.

2.2 Organizational Responses in Legacy Newsrooms

Major newspapers responded by creating and professionalizing roles dedicated to platform discovery (SEO) and off-platform packaging and distribution (social media editing). While precise first-hire dates are not always publicly archived, the pattern is consistent: SEO and audience roles diffused widely in the mid-to-late 2010s and expanded again with short-form video in the early 2020s. The typcial job descriptions include optimizing headlines and metadata for search intent, identifying trending, and monitoring ranking performance.

The New York Times named its first Social Media Editor in May 2009 (Jennifer Preston), formalizing a newsroom role to use social tools to find sources, shape coverage, and engage audiences.[4] By 2023, the Times advertised an *Associate SEO Editor* to "maintain visibility on the biggest news stories of the day," provide search recommendations, and plan around anticipated events—indicating a mature editorial search function.[5] I also document newsroom changes of other major newspapers in Section A.

Between 2017 and 2024, news discovery shifted toward feeds and influencers, with platform composition diversifying (Facebook remained largest; TikTok rose sharply among younger adults). These facts rationalize newsroom investments in SEO/social roles and help explain the supply-side adjustments modeled in this paper. As platforms changed recommendation policies, newsrooms adapted content packaging to preserve visibility—consistent with the mechanisms I estimate in the structural model.

[3] Pew Research Center (Nov. 18, 2024), "America's News Influencers."
[4] The Guardian (May 26, 2009), "New York Times names first social media editor."
[5] Talking Biz News (Aug. 10, 2023), "The New York Times seeks an associate SEO editor."

2.3 Platform Evolution of Facebook

Since 2017, Facebook's algorithm and its role as a news distributor have undergone substantial transformations—both in how content is ranked and in how many Americans rely on the platform for news.

In the years following 2017, Facebook's News Feed algorithm increasingly prioritized posts that generated engagement—such as comments, shares, and reactions—and those that fostered "meaningful social interactions" among users. In January 2018, Facebook formally announced a major algorithm update aimed at shifting attention away from passive consumption of public content toward posts from friends, family, and community groups.[6] This adjustment substantially reduced the visibility of news content, particularly from outlets that relied on organic distribution through the platform.

Beginning in 2021, Facebook initiated a series of experiments designed to "reduce the amount of political content" shown in user feeds.[7] These tests were expanded and concluded in 2022, when Meta reported that "placing less emphasis on shares and comments for political content is an effective way to reduce the amount of political content shown to users." As a result, engagement-driven amplification of political news fell sharply, and many outlets observed measurable declines in reach for politically oriented posts. For news organizations, these changes implied that visibility was increasingly determined by platform algorithms rather than editorial priorities—effectively tying exposure to engagement metrics. This evolution illustrates Facebook's ongoing experimentation with balancing engagement optimization, user satisfaction, and public scrutiny over algorithmic bias.

In January 2025, Meta announced a partial reversal of this policy. Reflecting on the reduction of civic content since 2021, Meta described the prior policy as a "pretty blunt approach" and announced that it would begin phasing political and civic content back into Facebook, Instagram, and Threads using more personalized ranking.[8] Rather than broadly

[6] *TechCrunch* (Jan. 28, 2018), "How Publishers Will Survive Facebook's News Feed Change."
[7] Meta (Feb. 10, 2021), "Political Content in Feeds."
[8] Meta (Jan. 7, 2025), "More Speech and Fewer Mistakes."

demoting political content based on its content type, Facebook would increasingly rank such content using explicit signals, such as likes, and implicit signals, such as viewing behavior, to predict whether it is relevant to a particular user. Meta also stated that it would recommend more political content when these personalized signals indicated user interest. Thus, the 2025 announcement marked a shift from broad content-based demotion toward a more user-specific treatment of political news. I use these changes below to examine whether the way newspapers frame the same news online responds to changes in platform incentives.

Facebook remained an important gateway to news throughout this period. According to the Pew Research Center, about 30% of U.S. adults regularly received news from Facebook in 2023.[9] This share has declined from earlier years: in 2017, roughly 45% of Americans reported getting news from Facebook.[10] At the same time, as of 2024, approximately 54% of U.S. adults say they sometimes get news from social media platforms more broadly.[11] These statistics indicate that while Facebook remains a major gateway to news, its relative importance has declined as users have diversified toward other platforms such as YouTube, X (formerly Twitter), and TikTok.

Together, these policy changes provide a useful setting for studying how news supply responds to platform incentives. Facebook first reduced the algorithmic distribution of political content and later moved toward reintroducing such content based on personalized user signals. In Section 4, I exploit these changes to test whether online headline framing adjusts in the corresponding direction.

3 Data

This section describes the construction of the dataset that links print and online newspaper headlines to social media posts and user engagement measures. The final dataset combines

[9] Pew Research Center (June 12, 2024), "How Americans Get News on TikTok, X, Facebook and Instagram."

[10] Pew Research Center (Sept. 7, 2017), "News Use Across Social Media Platforms 2017."

[11] Pew Research Center (June 12, 2024), "Social Media and News Fact Sheet."

multiple large-scale sources—ProQuest, the New York Times Developers API, Meta's Content Library, and the U.S. Congressional Record—to analyze how algorithms influence news production and online presentation. The sample spans April 2017 to December 2025 and includes major U.S. news outlets and prominent news influencers.

3.1 Newspaper Headlines: ProQuest and NYT Developers

The core of the dataset consists of matched print and online headlines for major newspapers (including all 7 newspapers with more than 300,000 combining circulation and online subscriptions in 2023). For the New York Times, through the *NYT Developers API*, precise metadata linking online publication time, URL, and the corresponding full print version information is provided. For other newspapers, both online and print headlines are obtained from the *ProQuest TDM Database*, which provides text and metadata for newspaper articles, including publication date, section, and full article text. Together, I acquire 2.7 million online headlines under the name of each outlet's official website and 1.8 million print headlines separately under each newspaper itself.

To construct a one-to-one match between print and online versions of the same article, I match records by the first 300 characters of article body text after removing boilerplate and punctuation. In the case of multiple headlines for the same news article, I use the latest one as it reflects the editor's final decision. The resulting dataset allows direct comparison between the print and the online headlines of each article. This alignment makes it possible to measure how editorial tone and ideological framing differ between print and online dissemination.

3.2 Facebook Posts: Meta Content Library

The social media component of the dataset comes from Meta's *Content Library*, which provides comprehensive post-level data on all publicly available Facebook pages. From this source, I collect approximately 3 million posts from 27 major media outlets and 11 top news influencers with more than three million followers. For each post, the dataset includes

the posting date, textual content, URL link to the associated article, and engagement metrics—number of views, shares, comments and reactions.

The outlets represented in the sample span all major media segments:

- **National Newspapers:** New York Times, Washington Post, USA Today, Wall Street Journal, The Guardian;
- **Regional Newspapers:** Boston Globe, Chicago Tribune, Los Angeles Times, Orlando Sentinel, Tampa Bay Times, Baltimore Sun;
- **Digital Media:** Al Jazeera, BuzzFeed, HuffPost, Voice of America, POLITICO, Bloomberg, NPR, Reuters, Yahoo Finance;
- **Broadcast:** ABC News, CBS News, NBC News, BBC News;
- **Cable:** Fox News, CNN, MSNBC;
- **Influencers:** Heather Cox Richardson, Tomi Lahren, Ben Shapiro, Glenn Beck, Robert Reich, Charlie Kirk, Dan Bongino, Sean Hannity, Graham Allen, Terrence K. Williams, Jay Sekulow.

Each post is matched to the corresponding article in the newspaper dataset using the URL link when possible.

3.3 Congressional Record and Political Slant

To measure ideological slant, I extend the dictionary-based measure from Gentzkow and Shapiro (2010). The original index quantifies how likely each phrase is to be spoken by Democrats versus Republicans in the U.S. Congress. I reconstruct the measure using *Congressional Record transcripts*[12] from 2017 to 2025 following the method of phrase extraction in Gentzkow et al. (2019) and update the phrase-level log-odds ratios for each year to reflect contemporary political language.

[12]Files obtained from https://www.congress.gov/congressional-record.

Each article or post is assigned a slant score computed as the mean of phrase-level weights. Positive values indicate a right-leaning language, negative values indicate a left-leaning language, and values near zero correspond to neutral phrasing. Because both print and online headlines are processed through the same metric, the difference between them for the same article directly captures the shift in ideological framing between platforms.

This measure serves as the key input for both the difference-in-differences regressions and the structural estimation of user and algorithmic preferences.

3.4 Topics, Political Content, and Sentiment

To characterize the content and tone of each headline and Facebook post, I construct several measures using pretrained transformer-based language models. These measures identify news-related content, distinguish political news from other stories, and measure the emotional tone of the text.

- **Sentiment:** I measure sentiment using the `Twitter-RoBERTa-base` model,[13] a transformer pretrained on approximately 124 million tweets and finetuned for sentiment classification. The model is designed for short text and is therefore well suited to news headlines and social-media posts. I construct a continuous sentiment score ranging from -1 to 1, where negative values indicate negative tone, positive values indicate positive tone, and values near zero indicate relatively neutral language.

- **Political-content classification:** Because political content is conceptually distinct from ideological slant, I separately classify whether each Facebook post concerns politics. I construct two binary political-content measures using independently developed text classifiers.

  My preferred measure uses the FeedKit RoBERTa classifier from Piccardi et al. (2025). FeedKit is a recent task-specific classifier developed to identify political content in

[13] See https://huggingface.co/cardiffnlp/twitter-roberta-base-sentiment-latest.

social-media text and was used in a closely related study of algorithmic feed design. Its training domain and intended application are therefore particularly well matched to the Facebook posts studied here.

As an alternative measure, I use the EXPO BERT classifier from Wojcieszak et al. (2023), which was developed in the political-science literature and is fine-tuned on manually coded news headlines to distinguish political from non-political content.

Both political-content measures are constructed independently of the Congressional Record-based slant measure. I use FeedKit as the primary classification in the regression analysis and report corresponding results using EXPO to assess the sensitivity of the findings to alternative definitions of political content. Appendix B provides additional details on the classifiers and their pairwise agreement.

- **Topic classification:** To identify whether Facebook posts contain news-related content, I use the `tweet-topic-21-multi` classifier,[14] a multi-label transformer model finetuned for topic classification. I use the resulting topic labels to distinguish news and social-concern content from categories such as entertainment, sports, and promotional material. The main Facebook sample is restricted to news-related posts.

Together, these measures distinguish the ideological framing, emotional tone, topic, and political content of the news. The political content indicators are used in the regression analysis to measure differential exposure to Facebook's algorithm changes, while sentiment provides a comparison outcome for whether those changes affect ideological framing or headline tone more broadly.

3.5 User Ideology: Pew Surveys

The Meta Content Library provides post-level views and engagement but does not identify the ideological preferences of individual users. To characterize the distribution of user ide-

[14] See https://huggingface.co/cardiffnlp/tweet-topic-21-multi.

ology, I use survey data from the Pew Research Center's American Trends Panel (ATP). For each survey wave, I restrict the sample to respondents who report using Facebook and compute survey-weighted shares across five ideological categories: very liberal, liberal, moderate, conservative, and very conservative. When multiple waves are available within a year, I average the wave-specific distributions.

The resulting distributions provide the annual share of Facebook users in each of the five ideological groups, which I treat as fixed within each year. To connect these survey categories to the continuous slant measure, I assign equally spaced reference positions along the ideological spectrum. The structural model then estimates the location and dispersion of these ideological preferences, while the relative size of each group is determined by the survey data. This approach avoids inferring the composition of Facebook users from observed engagement, which is itself influenced by user preferences and the recommendation algorithm.

3.6 Combined Dataset

The merged dataset integrates print and online newspaper headlines for leading U.S. outlets with Facebook posts and engagement measures from the Meta Content Library. I link Facebook posts to the corresponding online articles through their URLs and then match these articles to their print counterparts. Posts with matched print headlines account for approximately 12.2% of the full Facebook sample. Additional summary statistics for the full sample and the matched print-headline subsample are reported in Appendix Table 7.

Because the headline displayed with a link on Facebook is the content directly exposed to the platform's recommendation system and users, I use this link title as the main measure of the online headline throughout the analysis. These titles are generally very similar to the headlines displayed on the corresponding newspaper webpages. Importantly, the link title is associated with the article's URL through the webpage's link-preview metadata rather than being personalized on Facebook for individual users. Thus, the same title can generally accompany the article when the link is shared in other contexts as well. The distinction

is therefore between the print presentation of the article and its online presentation, rather than between a Facebook-specific headline and a newspaper headline.

Table 1 provides an example of an online headline shown on Facebook, and the print headline referring to the same New York Times article. These differences can translate into different measured sentiment and ideological slant.

Format	Headline of the Same Article	Sentiment	Slant
Online	Trump Begins Third Day of Twitter Attacks on "Weak" Border Laws	-0.8	0
Print	Trump Foresees National Guard Securing Border	0	1.3

Table 1: Online and Print Headlines for the Same New York Times Article

The example also illustrates the construction of the slant measure. Phrases such as "National Guard" and "securing border" receive positive, Republican-leaning weights in the Congressional Record-based dictionary. Their presence therefore leads the print headline to receive positive slant scores, while the online headline posted on Facebook, which does not contain these phrases, receives a slant score of zero. The print headline contains more strongly Republican-associated language according to the dictionary and consequently receives a somewhat larger positive slant score than the online headline. This example highlights that the measure captures ideological differences in the language used to describe the same underlying news event.

3.7 Descriptive Patterns

I first document how online and print presentation differ in the matched sample. Throughout the analysis, I refer to the title displayed with the article link on Facebook as the online headline. This is the headline users observe on Facebook when encountering the linked article and need not coincide exactly with the headline displayed after clicking through to the newspaper's website.

The distinction between online and print presentation is also institutionally meaningful. Online headlines can be revised after publication in response to audience feedback and social-

media performance. Leung and Strumpf, 2023 document that newspapers adjust online headlines in response to feedback from social media. By contrast, the print headline is fixed once the newspaper is produced and distributed. News organizations also commonly use distinct print and digital production workflows, with dedicated digital, audience-development, SEO, and social-media functions involved in online presentation.

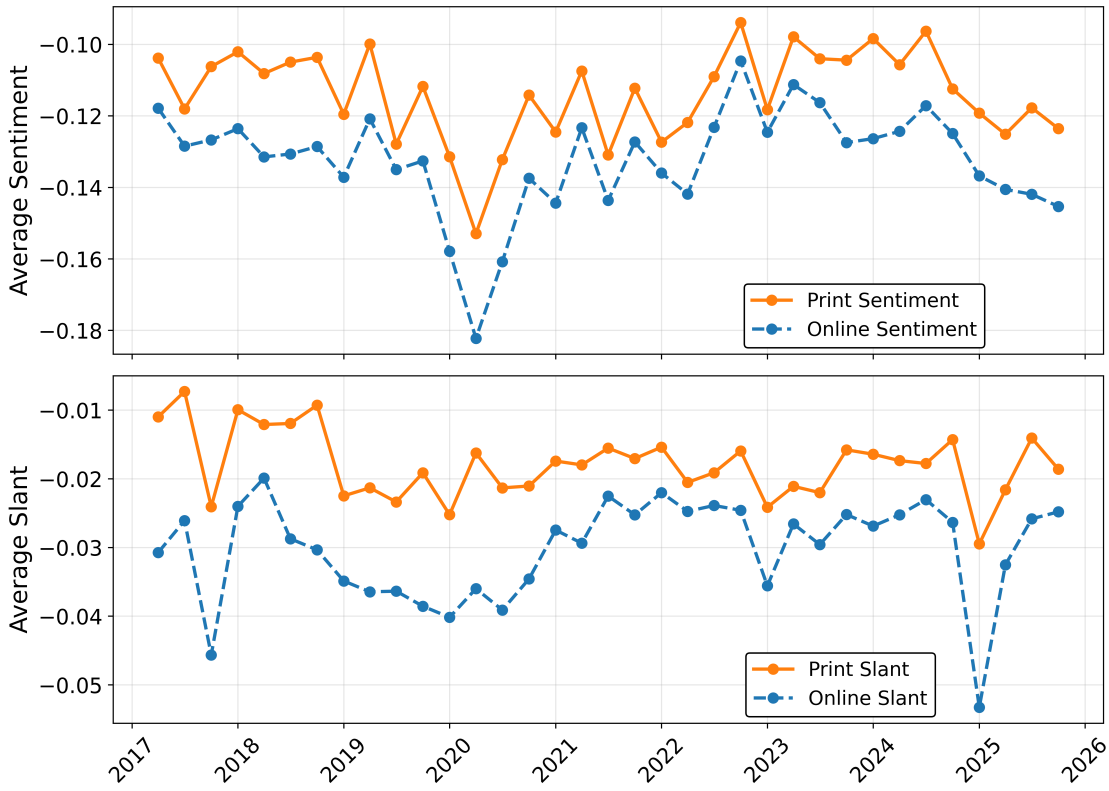


Figure 1: Quarterly Average Sentiment and Slant in Online and Print Headlines

Notes: The figure reports quarterly averages for matched articles from the 11 major newspapers in the sample. "Online" refers to the title displayed with the article link on Facebook, while "Print" refers to the corresponding print headline of the same article. The top panel compares average sentiment and the bottom panel compares average ideological slant.

Figure 1 plots quarterly average sentiment and ideological slant for the online and print headlines of matched articles. The two versions of the same articles exhibit persistent differences in presentation. Online headlines are, on average, more negative than their print counterparts throughout the sample period, consistent with evidence that negative content attracts greater attention in online news markets (Armona et al., 2024). The two series

also move together substantially over time, suggesting that both respond to changes in the underlying news environment.

A similar pattern appears for ideological slant. Online headlines generally use more left-leaning language than their corresponding print headlines. Because most newspapers in the matched sample are themselves left-leaning, this pattern is consistent with online presentation using more ideologically pronounced language than print presentation. Importantly, however, both print and online slant vary considerably over time. In early 2025, for example, average online slant moves sharply in the leftward direction before partially reversing, while print slant also changes over the same period.

This time variation illustrates why the empirical analysis below focuses on the online–print difference rather than on online slant in levels. Changes in the political news environment can alter the ideological content of the stories being covered and therefore move both online and print headlines. This concern is particularly relevant around periods of major political transition, such as the beginning of a new presidential administration. By comparing the Facebook link title with the print headline of the same article, the within-article gap differences out much of this common variation in the underlying story and isolates changes in how that story is presented online.

The persistence and systematic direction of the online–print differences also provide reassurance that the headline-level slant measure captures meaningful variation rather than purely idiosyncratic measurement noise. If the measured differences between online and print headlines were generated primarily by mean-zero phrase-level noise, one would not expect such persistent separation between the two series over time. This descriptive evidence does not by itself validate every article-level slant estimate, but it suggests that the measure captures systematic differences in how the same news content is presented across formats.

These descriptive differences do not, however, establish that recommendation algorithms cause newspapers to alter their online presentation. Newspapers may tailor online headlines to a different audience even in the absence of platform recommendation incentives,

and online and print readers may differ systematically in their preferences. Although the institutional evidence above shows that news organizations invest in SEO, social-media, and audience-development functions to optimize online visibility, separating platform incentives from differences in audience demand requires additional variation. The next section exploits Facebook's 2025 change in the treatment of political content to provide causal evidence on whether platform incentives contribute to the observed online–print divergence.

4 Regression Analysis

The descriptive evidence above establishes systematic differences between the online and print presentation of the same news stories, but it cannot determine whether those differences are caused by platform incentives or by differences in the audiences consuming news through the two channels. To isolate the role of the recommendation algorithm, I exploit Facebook's 2025 reversal of its policy toward political content.

4.1 Empirical Strategy and Main Results

Meta announced on January 7, 2025 that it would begin restoring political and civic content to users' feeds, reversing the broad demotion policy introduced in 2021. Implementation occurred gradually and was completed on March 28, 2025. The change provides useful identifying variation because it altered the algorithmic incentives associated specifically with political content while leaving non-political news comparatively less affected.

I therefore compare changes in the online–print headline divergence of political and non-political stories around the implementation of the policy. Political content is identified using the FeedKit classifier described in Section 3.4, with the EXPO classifier used as alternative measures. If recommendation incentives contribute to online editorial choices, the restoration of political content should produce a differential change in the way political stories are presented online relative to their print counterparts.

The central outcome is the absolute difference between the Facebook post and the corresponding print headline. For post n from outlet j on date t, I define

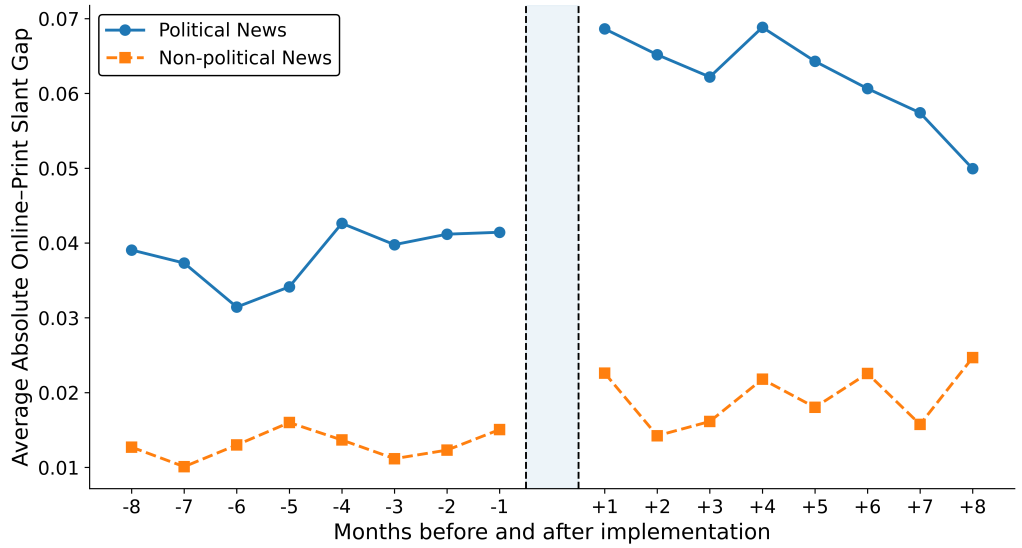
$$Y_{j,t}^{n} = \left| s_{j,t}^{n} - p_{j,t}^{n} \right|,$$

where $s_{j,t}^{n}$ denotes the slant of the Facebook post and $p_{j,t}^{n}$ denotes the slant of the matched print headline. I interpret this absolute gap as the extent to which the online presentation of a story departs from its print counterpart. A larger value indicates a greater difference in ideological language between the two versions, without imposing a direction on that difference.

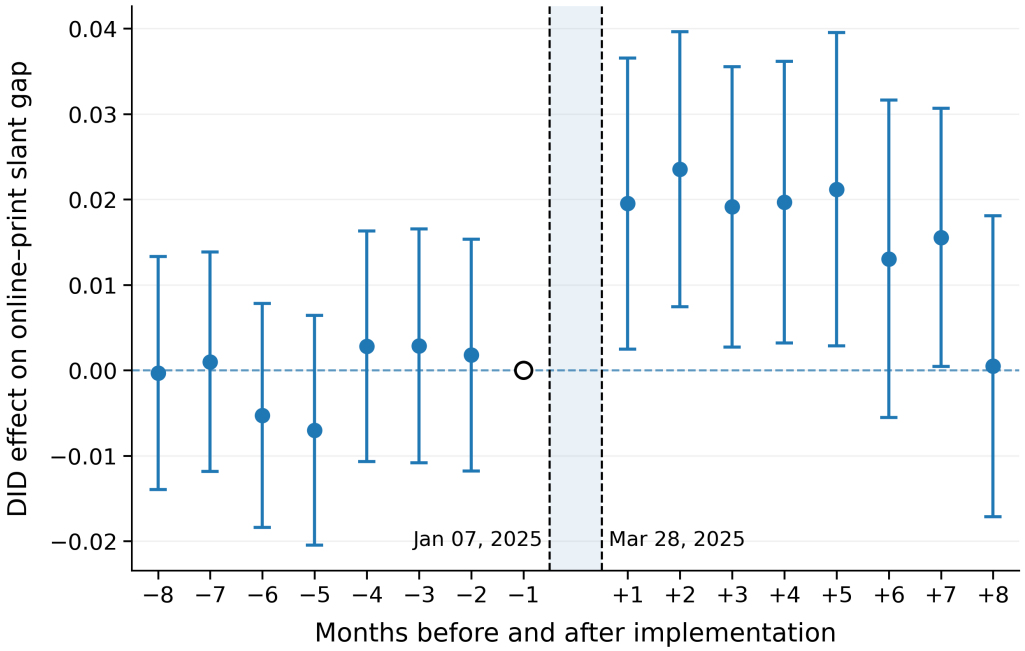
Figure 2 presents the evolution of this gap for political and non-political news around Facebook's 2025 restoration of political content. Panel (a) reports the average absolute online–print slant gap for each group in 31-day bins. Panel (b) reports dynamic difference-in-differences estimates from a regression with outlet and date fixed effects, interacting political-content status with event-time indicators. The month immediately preceding the implementation window is the omitted reference period.

Panel (a) shows that political news has a consistently larger online–print slant gap than non-political news even before the policy change. However, the difference between the two groups remains relatively stable over the pre-implementation period. Following implementation, the gap increases substantially for political news, while the corresponding increase for non-political news is considerably smaller.

Panel (b) formalizes this comparison using a dynamic DID specification. The pre-implementation coefficients fluctuate around zero without a clear systematic trend, providing support for the parallel-trends assumption. Following implementation, the political–non-political difference increases sharply, with estimated effects of approximately 0.02 during the first several months. The effect remains positive through the medium run, although its magnitude and precision decline toward the end of the sample.



(a) Average online–print slant gap



(b) Dynamic DID estimates

Figure 2: Political Content and Online–Print Slant Divergence around the 2025 Policy Change

Notes: Political content is identified using the FeedKit classifier. Panel (a) plots the average absolute online–print slant gap for political and non-political news in 31-day bins. Panel (b) reports the corresponding event-study DID coefficients, estimated with outlet and date fixed effects and normalized relative to the last pre-implementation bin. The shaded region marks the implementation period from January 7 to March 28, 2025, which is excluded from estimation. Error bars in Panel (b) represent 95% confidence intervals.

Together, the two panels indicate that the restoration of political content was followed by a substantial increase in the extent to which outlets edited political headlines away from their print counterparts, relative to non-political stories. The timing and differential nature of this response are consistent with media outlets adjusting their online presentation in response to changes in platform incentives.

The results are robust to an alternative measure of political content. Appendix Figure 11 repeats both analyses using the EXPO classifier. The resulting patterns are similar: the pre-implementation political–non-political difference is relatively stable, while a pronounced differential increase emerges after implementation. This consistency across independently developed classifiers suggests that the findings are not driven by the particular definition of political content.

To quantify this change, I estimate the following difference-in-differences specification:

$$\left|s^n_{j,t} - p^n_{j,t}\right| = \beta_0 Political^{n,m}_{j,t} + \beta_1 \left(After_t \times Political^{n,m}_{j,t}\right) + \gamma_j + \delta_t + \varepsilon^n_{j,t}, \tag{1}$$

where $Political^{n,m}_{j,t}$ is an indicator equal to one if post n is classified as political under classifier $m \in \{\text{FeedKit}, \text{EXPO}\}$. The indicator $After_t$ equals one for observations after the completion of the implementation period. Observations between January 7 and March 28, 2025 are excluded. All specifications include outlet fixed effects γ_j and date fixed effects δ_t. The coefficient of interest is β_1, which measures whether the online–print divergence of political stories changed relative to non-political stories following the policy change. Here I focus on a five-month window on either side of the implementation period, while the event-study figures use an extended eight-month window to examine pre-trends and the persistence of the effect.

Table 2 reports the results. Using the preferred FeedKit classification, political stories exhibit a 0.0253 larger online–print slant gap than non-political stories before the policy change. Following Facebook's restoration of political content, this difference increases by an

Table 2: Effect of Restoring Political Content on Online–Print Headline Divergence

	Partisanship Gap		Sentiment Gap	
	FeedKit	EXPO	FeedKit	EXPO
Political	0.0253***	0.0250***	-0.0313***	-0.0302***
	(0.0047)	(0.0046)	(0.0094)	(0.0083)
After × *Political*	0.0205***	0.0218***	0.0061	0.0046
	(0.0028)	(0.0037)	(0.0069)	(0.0088)
Media fixed effects	Yes	Yes	Yes	Yes
Date fixed effects	Yes	Yes	Yes	Yes
Observations	37,742	37,742	37,742	37,742
Media clusters	11	11	11	11
R^2	0.0250	0.0250	0.0325	0.0324

Notes: The table reports difference-in-differences estimates around Facebook's 2025 restoration of political content. The implementation period from January 7 through March 28, 2025 is excluded, and the estimation sample contains 155 days before and after the implementation period. The dependent variables are the absolute online–print differences in partisanship and sentiment, respectively. *Political* is defined separately using the FeedKit and EXPO political-content classifiers. All specifications include media-outlet and date fixed effects. Standard errors, reported in parentheses, are clustered by media outlet. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

additional 0.0205, significant at the 1% level. Thus, the political–non-political differential rises from 0.0253 to 0.0458, an increase of approximately 81%. The estimates are similar when political content is instead classified using EXPO, with interaction coefficients of 0.0218. Thus, the result is not sensitive to the particular definition of political content.

In contrast, the corresponding effects on the absolute sentiment gap are small and generally statistically insignificant. This suggests that the policy change primarily affected ideological framing rather than generating a broad increase in differences in headline tone.

It is important that the dependent variable measures the magnitude, rather than the direction, of editorial adjustment. At the individual-post level, outlets do not uniformly move headlines in a more liberal or more conservative direction. Some online versions become more left-leaning than their print counterparts, others become more right-leaning, and others remain largely unchanged. These heterogeneous directional responses can offset in signed averages even when substantial editing occurs. Therefore, the event study evidence

establishes that changes in Facebook's ranking incentives affect how much outlets alter the ideological presentation of political stories online. By itself, however, it does not explain why a particular outlet chooses a particular direction of adjustment, nor how those choices depend on audience preferences and the recommendation algorithm. The structural model in the next section addresses these questions.

4.2 Identification and Alternative Explanations

A central identification concern is that the 2025 policy change occurred during a period in which the underlying political news environment was also changing. For example, the beginning of a new presidential administration may increase the amount and salience of political news independently of Facebook's recommendation policy. The within-article design helps address this concern because the Facebook post is compared with the print headline of the same story. Changes in the underlying event being covered are therefore largely common to both versions, while the outcome captures how much the Facebook presentation departs from its print counterpart.

The print headline is particularly useful as a within-article benchmark because it is largely predetermined relative to short-run changes in online presentation. Once the print edition has been produced, its headline cannot respond contemporaneously to subsequent social-media feedback, whereas online headlines and Facebook presentation can continue to be adjusted. Consistent with this distinction, Leung and Strumpf, 2023 show that newspapers revise online headlines in response to feedback from social media. The use of distinct print and digital production workflows further limits the scope for immediate spillovers from short-run online feedback into the already-produced print version. Accordingly, over the relatively short horizons used in the empirical analysis, the print headline provides a natural control for the underlying story while the online–print gap captures subsequent differences in presentation.

I next examine changes in exposure, engagement conditional on exposure, and posting activity around the 2025 implementation period. These outcomes help distinguish the policy-

induced change in distribution from alternative explanations based on a contemporaneous increase in users' revealed demand for political content.

Table 3: Exposure, Engagement, and Posting Activity Around the 2025 Policy Change

Sample	Measure	Pre-period Mean	Post − Pre	S.E.	N
All Newspapers Posts	Views	33,795	4,753***	1,229	110,410
	Engagement / Views	0.00657	0.00015**	0.00007	110,410
	Posts (Daily)	362.3	-11.62***	4.07	310
FeedKit Political	Views	22,799	13,966***	1,362	67,382
	Engagement / Views	0.00919	-0.00019*	0.00010	67,382
	Posts (Daily)	205.1	24.79***	3.73	310
EXPO Political	Views	24,645	13,813***	1,386	69,364
	Engagement / Views	0.00896	-0.00025**	0.00010	69,364
	Posts (Daily)	210.5	26.85***	3.81	310

Notes: The table reports changes in exposure, engagement conditional on exposure, and posting activity around Facebook's 2025 restoration of political content. The implementation period from January 7 through March 28, 2025 is excluded. The pre- and post-periods each contain 155 days. Political posts are classified separately using the FeedKit and EXPO classifier. Views and engagement per view are measured at the post level, while posting activity is aggregated at the daily level. Standard errors are reported in the fifth column. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3 reports changes in exposure, engagement conditional on exposure, and posting activity around the policy implementation. Across both political-content classifications, political posts receive substantially more views following the restoration of political content. The increase ranges from approximately 14,000 views per post and is highly statistically significant. This sharp increase in exposure is consistent with the intended change in Facebook's distribution of political content.

At the same time, engagement conditional on exposure does not increase for political posts. Instead, engagement per view declines under both classifications, although the magnitude and statistical significance vary across classifiers. This pattern does not support the simple explanation that the increase in political exposure and editorial adjustment is driven by a contemporaneous increase in users' propensity to engage with political content. For all posts combined, engagement per view increases slightly, highlighting that the decline is specific to the political-content.

Political posting also increases substantially following the implementation period: the number of political posts rises by roughly 25–27 posts per day under both classifiers. I do not interpret this extensive-margin change as clean evidence of an algorithmic effect. The policy change coincides with a changing political news environment, including the beginning of a new presidential administration, which may independently increase the number of political stories available to outlets. This concern is one reason the main outcome is constructed as a within-article online–print difference rather than the level or quantity of political content. The print headline provides a benchmark for the underlying story being covered, so changes in the political news environment that affect the content of the story itself are largely common to both versions. The empirical analysis instead asks whether outlets alter the Facebook presentation of a given story relative to its print presentation. This distinction also defines the scope of the structural model. The model takes the set of available stories in a period as given and studies the intensive editorial margin—how outlets choose to frame those stories. Changes in topic selection or posting frequency, including the extensive-margin response documented in Table 3, therefore lie outside the scope of the model.

As an additional falsification exercise, I also repeat the main difference-in-differences specification using a set of placebo dates at which Facebook made no corresponding political-content policy change. I find no comparable systematic effects around these placebo dates. Appendix Table 8 reports the results.

4.3 Complementary Evidence from the 2022 Policy Change

The direction of the 2025 result is also consistent with the earlier period in which Facebook moved in the opposite policy direction. Beginning in 2021, Facebook tested reductions in the distribution of political content and gradually expanded these changes before announcing on July 19, 2022 that the revised treatment of political content had been implemented globally. Meta's own descriptions emphasize that this was a gradual sequence of tests and expansions rather than a single discrete intervention.

As complementary evidence, I estimate the same DID specification around the July 19, 2022 announcement. Appendix Table 9 shows that the political–non-political differential in the online–print slant gap decreases by approximately 15% using FeedKit and 14% using EXPO. The effect is statistically significant at the 5% level for FeedKit but not for EXPO.

The negative coefficients are consistent with a reduction in online–print slant divergence for political content relative to non-political content, opposite to the increase observed following the 2025 restoration. However, the 2022 policy was introduced gradually through tests and expansions lasting more than a year. Consequently, the July 19 announcement does not provide a comparably clean event date for identifying the full policy effect, and the estimates should be interpreted as descriptive changes around the announcement rather than causal effects of the initial rollout.

Taken together, the two episodes exhibit directional consistency: online–print slant divergence decreases around the completion of Facebook's political-content reduction and increases following its subsequent restoration. The main causal analysis therefore focuses on the more clearly delimited 2025 implementation period.

5 Model

I develop a dynamic structural model of online news markets to study how media outlets choose the ideological presentation of their news in response to user preferences and platform recommendation incentives. Three types of agents interact in the model: (i) a unit mass of *users* who consume news recommended by the platform, (ii) a *recommendation algorithm* that allocates users' attention across available posts, and (iii) J *media outlets* that strategically choose the slant of their online news posts.

The model focuses on the intensive editorial margin studied in the empirical analysis. In each period $t \in \{0, 1, 2, \ldots\}$, each outlet receives an exogenous set of news stories and chooses how much ideological slant to add to the online presentation of each story. The model

therefore takes story selection and the number of available stories as given and studies how outlets edit the ideological content of those stories.

After outlets make their editorial choices, the recommendation algorithm selects which available post to show to a user as a function of the user's ideology and the characteristics of the posts. Conditional on viewing a post, the user then decides whether to engage with it. These engagement decisions affect the incentives facing both the platform and media outlets, generating equilibrium feedback between user preferences, algorithmic recommendations, and outlets' editorial choices.

Here I introduce the model by working backward through this sequence of decisions: I first describe user engagement, then the platform's recommendation problem, and finally the media outlets' dynamic slant choices.

5.1 Users

In each period t, there is a continuum of users indexed by $i \in [0, 1]$. Each user i is characterized by an ideology $\mu_{i,t}$ that determines their preferences over politically slanted content. User ideology follows a distribution F_t, which may vary over time.

If a post $\{n_{j,t}\}$ from media outlets $j \in \{1, \ldots, J\}$ is recommended to user i, denote $R^n_{ij,t} = 1$. Conditional on being recommended, the user decides whether to engage based on their utility. Denote the engagement decision as $I^n_{ij,t} \in \{0, 1\}$.

The utility of user i from engaging with post $n_{j,t}$ is:

$$U^n_{ij,t}(I^n_{ij,t} = 1 \mid R^n_{ij,t} = 1) = W^n_{ij,t} + \varepsilon^n_{ij,t} \tag{2}$$

where $\varepsilon^n_{ij,t} \sim \text{Logistic}(0, 1)$ and the mean utility for engagement is

$$W^n_{ij,t} = g(X^n_{j,t}) + u^n_{j,t} + \alpha_t(s^n_{j,t} - \lambda_{j,t-1})^2 - \beta_t(s^n_{j,t} - \mu_{i,t})^2 \tag{3}$$

Here, $s^n_{j,t}$ is the slant of post $n_{j,t}$, and $\lambda_{j,t-1}$ is the expected slant of outlet j based on pre-

vious periods. The term $g(X_{j,t}^n)$ captures the predictable component of content value, where $X_{j,t}^n$ includes observable story and post characteristics, as well as outlet and publication-time characteristics. These characteristics are available to the media outlet when choosing its online headline and to the algorithm when recommending the post.

The term $u_{j,t}^n$ captures residual post-level content value not explained by these observed characteristics. I assume this residual is realized after the outlet chooses its headline, so the outlet makes its editing decision based on observable content characteristics and expected demand rather than the realization of this shock.

The term $-\beta_t(s_{j,t}^n - \mu_{i,t})^2$ captures the distaste for reading a news post whose slant $s_{j,t}^n$ deviates from the reader's preferred slant $\mu_{i,t}$, as proposed in Mullainathan and Shleifer (2005) and applied in Gentzkow and Shapiro (2010) and Martin and Yurukoglu (2017). The term $\alpha_t(s_{j,t}^n - \lambda_{j,t-1})^2$ captures the reduced-form response to news that deviates from an outlet's typical ideological position. Such a response may arise because atypical content is more salient or surprising (Bordalo et al., 2013, 2022), because cross-ideological reporting is perceived as more credible or informative (Chiang and Knight, 2011), or because readers make strategic inferences about the strength of the underlying information when an outlet reports news inconsistent with its usual stance (Huang et al., 2026).

The two slant-related terms capture competing forces in user preferences. While ideological alignment increases utility when $\beta_t > 0$, unexpected slant increases utility when $\alpha_t > 0$. Their relative magnitudes determine whether users prefer content close to their ideological positions or more extreme content that deviates from an outlet's typical slant.

Given the logistic error, the probability that a user engages after a recommendation is

$$\Pr(I_{ij,t}^n = 1 \mid R_{ij,t}^n = 1) = \frac{\exp(W_{ij,t}^n)}{1 + \exp(W_{ij,t}^n)} \tag{4}$$

The expected utility from the engagement decision, including the option not to engage, is $\log(1 + \exp(W_{ij,t}^n))$. This expected surplus enters the algorithm's recommendation objective,

allowing recommendations to respond to users' preferences over news content.

Notice that there is no price in the users' utility function. Although several newspapers in the data have their articles behind the paywall, the main product to be consumed here is the post or the headline itself, instead of the article linked to it. In fact, Sundar et al. (2025) finds that around 75% of the Facebook posts with links to news were shared without clicking into the linked web pages. Such phenomenon is more serve for extreme and user-aligned political content. Therefore, I focus on the posts as the product, which does not require a payment to be consumed.

5.2 Algorithm

I model the recommendation algorithm as selecting among available posts to maximize a reduced-form recommendation objective. The algorithm observes the characteristics of all available posts, including both the predictable content values and the residual post-level shocks that are not observed by media outlets when choosing their headlines. It also accounts for users' ideological preferences and predicts the utility they would obtain from engaging with each post. Its predicted utility for recommending post $n_{j,t}$ is:

$$\widehat{U}^n_{ij,t}(R^n_{ij,t}=1) = \widehat{W}^n_{ij,t} + \widehat{\varepsilon}^n_{ij,t} \tag{5}$$

where $\widehat{\varepsilon}^n_{ij,t} \sim$ i.i.d. EV1$(0,1)$ and mean utility for recommendation is

$$\widehat{W}^n_{ij,t} = \widehat{g}(X^n_{j,t}) + \widehat{u}^n_{j,t} + \widehat{\alpha}_t(s^n_{j,t} - \lambda_{j,t-1})^2 - \widehat{\beta}_t(s^n_{j,t} - \mu_{i,t})^2 + \zeta_t \log(1+\exp(W^n_{ij,t})) \tag{6}$$

The term $\widehat{g}(X^n_{j,t})$ captures the predictable component of the algorithm's additional valuation of post characteristics, beyond their contribution to user utility. The residual $\widehat{u}^n_{j,t}$ captures additional post-specific information used by the algorithm in ranking. Both $g(X^n_{j,t})$ and $\widehat{g}(X^n_{j,t})$ are observable to media outlets when choosing headlines. In contrast, $u^n_{j,t}$ and $\widehat{u}^n_{j,t}$ are not observed by outlets at that stage, but are observed by the algorithm when recom-

mendations are made.

The parameter ζ_t measures the weight the algorithm places on expected user utility from engagement. The parameters $\widehat{\alpha}_t$ and $\widehat{\beta}_t$ capture the algorithm's additional preferences over unexpected slant and ideological alignment, respectively, beyond their effects on user utility. These preferences may reflect platform objectives that differ from maximizing user engagement, or imperfect predictions of users' preferences.

I summarize non-news content and other recommendation alternatives through an outside option whose systematic utility is normalized to zero. Under the i.i.d. extreme-value assumption, the probability of recommending a particular news post is

$$\Pr(R^n_{ij,t} = 1) = \frac{\exp(\widehat{W}^n_{ij,t})}{1 + \sum_{m,h} \exp(\widehat{W}^m_{ih,t})} \tag{7}$$

Then the unconditional probability that a post is both recommended and engaged is

$$\Pr(I^n_{ij,t} = 1, R^n_{ij,t} = 1) = \frac{\exp(W^n_{ij,t})}{1 + \exp(W^n_{ij,t})} \frac{\exp(\widehat{W}^n_{ij,t})}{1 + \sum_{m,h} \exp(\widehat{W}^m_{ih,t})} \tag{8}$$

Aggregating over user types yields the population-level viewing and engagement shares:

$$\Pr(R^n_{j,t} = 1) = \int_i \frac{\exp(\widehat{W}^n_{ij,t})}{1 + \sum_{m,h} \exp(\widehat{W}^m_{ih,t})} dF_t(\mu_{i,t}) \tag{9}$$

$$\Pr(I^n_{j,t} = 1, R^n_{j,t} = 1) = \int_i \frac{\exp(W^n_{ij,t})}{1 + \exp(W^n_{ij,t})} \frac{\exp(\widehat{W}^n_{ij,t})}{1 + \sum_{m,h} \exp(\widehat{W}^m_{ih,t})} dF_t(\mu_{i,t}) \tag{10}$$

There are two channels through which the algorithm affects the exposure of slanted news. The direct channel operates through $\widehat{\alpha}_t$ and $\widehat{\beta}_t$, which place additional weight on unexpected slant and ideological alignment in the recommendation decision. The indirect channel operates through ζ_t: when the algorithm places positive weight on expected user utility, it is more likely to recommend content that users are predicted to find engaging.

Importantly, recommendations depend on both ideological preferences and content value.

A post that is poorly aligned with a user's ideology may nevertheless be recommended if its content value is sufficiently high. Consequently, the composition of users exposed to each post depends on the algorithm's recommendation decisions.

I model users and the algorithm as making static decisions within each period, while allowing their preferences and ranking parameters to change over time. This captures the evolving composition of Facebook users and frequent changes to the recommendation system, while keeping the focus on media outlets' dynamic decisions over news slant.

5.3 Media Outlets

Media outlets make decisions dynamically. In each period t, outlet j is endowed with $N_{j,t}$ news stories. Each story $n_{j,t} \in \{1, \ldots, N_{j,t}\}$ is characterized by its print headline slant $p^n_{j,t}$ and a vector of observable content characteristics $X^n_{j,t}$. These characteristics determine the predictable components of user content value $g(X^n_{j,t})$ and algorithmic content value $\widehat{g}(X^n_{j,t})$, both of which are known to the outlet when making its editing decision. The residual content-value shocks $u^n_{j,t}$ and $\widehat{u}^n_{j,t}$, in contrast, are not observed by the outlet but are observed by the algorithm when recommendations are made.

I describe the dynamic decision problem for a given outlet j in period t and suppress the j, t subscripts for notational simplicity. Let $\omega^n = \{p^n, X^n\}$ denote the observable characteristics of story n. The editors of the outlet decide whether to adjust the slant from the print version of post n by choosing $a^n \in \{-1, 0, 1\}$, which gives the slant as

$$s^n(a^n) = p^n + a^n \epsilon^n \tag{11}$$

where $\epsilon^n > 0$ represents the magnitude of the slant adjustment. Its distribution is known to the outlet, but its realization is uncertain when the editing decision is made. Thus, the outlet chooses the direction of the adjustment before observing its exact magnitude. Since the outlet also does not observe u^n or $\widehat{u}^n$ when making this decision, it evaluates each action

using expected views and engagement conditional on the observable story characteristics and market states.

Once all news stories have been edited and posted, the expected slant of the outlet $\lambda_{j,t}$ evolves as a moving average of slants from its past posts,

$$\lambda_t = (1-\rho)\lambda_{t-1} + \rho\frac{1}{N}\sum_n s^n. \tag{12}$$

The outlet's expected slant therefore evolves gradually with its published content. Producing more right-leaning (left-leaning) headlines shifts its expected slant toward the right (left), which in turn affects the demand for future posts through the unexpected-slant term in user and algorithmic utility.

Let θ denote the exogenous market and platform states, and let $\Omega = \{\omega^n\}_{n=1}^N$ collect the observable characteristics of the outlet's news stories. The timing of decisions is illustrated in Figure 3. At the beginning of each period, the outlet observes its current stories Ω, its previous expected slant λ_-, and the market states θ. It then chooses how to adjust the slant of each headline. Other outlets make their editorial decisions during the same period. After these decisions, the slant adjustments are realized and all posts are published. The algorithm subsequently observes the realized content values, including u^n and $\widehat{u}^n$, and selects personalized recommendations from the available posts. Users decide whether to engage conditional on receiving a recommendation. At the end of the period, the outlet's expected slant is updated to λ, and the economy transitions to the next period with exogenous states θ' drawn according to $G(\theta' \mid \theta)$.

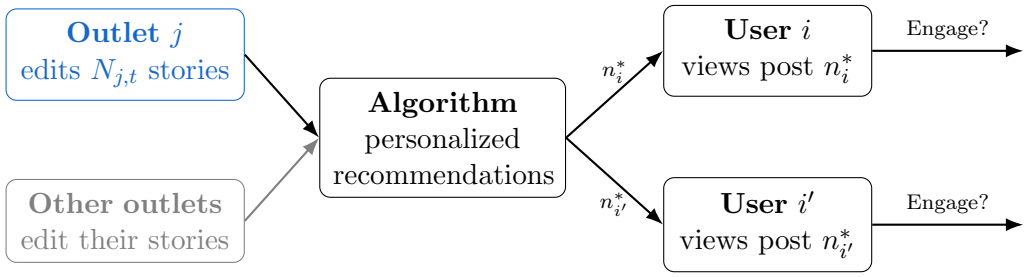


Figure 3: Timing in each period

Denote $A = \{a^n\}_{n=1}^N$ as the decisions on all news from the outlet in the period. The value function of the outlet j can be written as

$$V_j(\Omega; \lambda_-, \theta) = \mathbb{E}_\upsilon \Bigg[\max_{A \in \{-1,0,1\}^N} \Big\{ \tau_j^V U^V(A) + \tau_j^E U^E(A) - c_j(A) + \upsilon^A + \beta_j \mathbb{E}\left[V_j(\Omega'; \lambda, \theta') \mid \Omega, \lambda_-, \theta, A \right] \Big\} \Bigg], \tag{13}$$

where υ^A follows an i.i.d. type-I extreme-value distribution across action profiles. The action-specific editing cost is

$$c_j(A) = c_{j,-1} \sum_{n=1}^N \mathbb{1}\{a^n = -1\} + c_{j,+1} \sum_{n=1}^N \mathbb{1}\{a^n = 1\}. \tag{14}$$

Let $S_n^V(A)$ and $S_n^E(A)$ denote the population-level recommendation and joint recommendation–engagement probabilities defined in the previous subsection, evaluated under the outlet's editing decisions A and the realized market conditions. The expected flow payoffs are $U^V(A) = \mathbb{E}[\sum_{n=1}^N S_n^V(A) \mid \Omega, \lambda_-, \theta, A]$ and $U^E(A) = \mathbb{E}[\sum_{n=1}^N S_n^E(A) \mid \Omega, \lambda_-, \theta, A]$. These expectations integrate over the unobserved content-value shocks, the realized magnitudes of headline adjustments, and the equilibrium behavior of competing outlets. While the algorithm observes the realized content values when allocating recommendations, outlets choose their editing actions based on expected views and engagement. Because each user receives at most one recommendation, the recommendation events for different posts are mutually exclusive. Their probabilities can nevertheless be summed to obtain the outlet's expected total exposure. The same aggregation applies to engagement.

The parameters τ_j^V and τ_j^E measure the outlet's marginal value of expected exposure and engagement, respectively. Larger values imply stronger incentives to adjust headline slant in response to the recommendation algorithm and user demand. The parameters $c_{j,-1}$ and $c_{j,+1}$ capture direction-specific editing costs, allowing outlets to differ in their willingness to adopt more left-leaning or right-leaning headlines. Together with the dynamic effect of editing on

future outlet slant, these parameters determine the outlet's optimal supply response.

5.4 Oblivious Equilibrium

The above environment describes a dynamic game wherein media outlets add slants dynamically and strategically. In other words, an outlet's past choices affect its current choice through their effect on the outlet's expected slant $\lambda_{j,t-1}$. In addition, other outlets' current and past choices affect a given outlet's choice, as the probability of each post being viewed or engaged will change accordingly. The literature pioneered by Pakes and McGuire (2001) often uses Markov Perfect Equilibrium (MPE) to analyze players' behavior in such dynamic games. However, in my setting with more than thirty media outlets and hundreds of news posts in each period, the state variable for the game becomes the slant of each post as well as the expected slant of each media outlet which is a vector of dimension easily more than a thousand. This is the curse of dimensionality, which is particularly a problem in dynamic settings with large state space, as discussed in Pakes and McGuire (2001).

In order to curb this problem, I use the oblivious equilibrium as developed by Weintraub et al. (2008). Additionally, I use the adapted version that accounts for time-varying state space as in Saeedi (2019). Under this equilibrium concept, it is assumed that the total utility of all posts are the state variables; that is, all histories of the game that lead to the same level of total utility of all posts must lead to the same strategies. This is in contrast with MPE, in which variations in the past individual states affect equilibrium strategies beyond their effect on average industry variables. This equilibrium concept is particularly applicable to markets that are comprised of many news suppliers. Note that with many smaller outlets, an approximate law of large numbers implies that the distributions of individual states are somewhat invariant, which would then mean that industry average becomes the relevant state variable in the game. In my sample, there are 38 media outlets, each with dozens of posts in a period. Among the 3 million Facebook posts, less than 10 posts have more than 5% of the views of that period, and most posts have less than 1% of view share; therefore,

oblivious equilibrium can be a good approximation.

The application of the oblivious equilibrium has two key implications: first, the only relevant endogenous state variable is the expected slant of the given media outlet; second, each outlet is small and, as a result, an outlet's choice does not affect the state of the industry. In other words, outlets are oblivious to their knowledge of other outlets' state.

Given this discussion, I define the aggregate demand and platform states as

$$\theta_t = \{\alpha_t, \beta_t, \widehat{\alpha}_t, \widehat{\beta}_t, \zeta_t, g_t, \widehat{g}_t, F_t\},$$

where g_t and $\widehat{g}_t$ represent the predictable content-value functions for users and the algorithm, respectively, and F_t is the distribution of the five ideological user types.

For each ideological type g, let $M_{g,t} = 1 + \sum_{h=1}^{J} \sum_{m=1}^{N_{h,t}} \exp\left(\widehat{W}^m_{gh,t}\right)$ denote the market denominator in the recommendation probability, where the first term represents the non-news outside option. In the oblivious equilibrium, each media outlet takes the aggregate recommendation environment $\mathcal{M}_t = \{M_{g,t}\}_{g=1}^{5}$ as given when evaluating its editorial choices, rather than tracking the individual decisions of competing outlets. These aggregate quantities are equilibrium objects, determined by the supply of content and the algorithm's recommendation utility.

An oblivious equilibrium consists of editorial policy functions $A^*_j(\Omega_j; \lambda_{j,t-1}, \theta_t, \mathcal{M}_t)$ and a consistent aggregate recommendation environment such that:

1. Recommendation and engagement probabilities are determined by the user and algorithmic decision problems.

2. Each outlet's editorial policy maximizes its expected dynamic value, taking its own states and the aggregate recommendation environment as given.

3. The aggregate recommendation environment is consistent with the content supplied by all outlets and the resulting algorithmic utilities.

6 Identification and Estimation

This section describes the identification and estimation of the model parameters introduced above. Estimation proceeds in two stages. First, I estimate the demand-side parameters that govern user behavior ($\{\alpha_t, \beta_t\}$) and algorithmic recommendation ($\{\widehat{\alpha}_t, \widehat{\beta}_t, \zeta_t\}$) using post-level view and engagement data. Estimation is performed separately for each period t to allow for time variation in the platform and user environment. Second, I estimate the supply-side parameters that govern media outlet behavior ($\tau_j^V, \tau_j^E, c_{a,j}, \beta_j$) using observed editorial slant choices and the recovered conditional choice probabilities (CCPs).

6.1 Users and Algorithms (Demand Side)

For each post $n_{j,t}$ published by outlet j in period t, I observe the number of views $V_{j,t}^n$ and the number of engagements (comments plus shares) $E_{j,t}^n$. I also observe the post slant $s_{j,t}^n$ and the number of posts from outlet j in period t, $N_{j,t}$.

Before estimating the parameters, two state variables are estimated separately. First, I construct the outlet's expected slant by the equation

$$\lambda_{j,t} = (1-\rho)\lambda_{j,t-1} + \rho \frac{1}{N_{j,t}} \sum_n s_{j,t}^n, \tag{15}$$

where the initial $\lambda_{j,0}$ is the average slant of posts in several periods preceding the estimation window, and the decay parameter ρ is calibrated.

The second state variable is the distribution of user ideology. I distinguish five ideological groups corresponding to the Pew survey categories: very liberal, liberal, moderate, conservative, and very conservative. Let $q_g \in \{-1, -0.5, 0, 0.5, 1\}$ denote the normalized position of group g, and let $p_{g,t}$ denote its population share among Facebook users. The population shares are calibrated using the annual Pew survey distributions described in Section 3.5 and held fixed within each year. To map the survey categories into the continuous slant measure,

I parameterize the ideological position of each group as

$$\mu_{g,t} = a_t + b_t q_g, \qquad b_t > 0. \tag{16}$$

Here, a_t determines the center of the ideological distribution, while b_t determines the dispersion of ideological preferences. Unlike the group shares, these two parameters are estimated jointly with the demand-side parameters in each estimation period.

This approach uses external survey information to determine the relative size of ideological groups, rather than inferring user composition from observed engagement, which is itself affected by personalized recommendations.

Given the calibrated ideological group shares and the outlet slant state, the demand-side estimation recovers the parameters governing user preferences and algorithmic recommendations. The key structural parameters are α_t and β_t, which govern users' preferences over unexpected slant and ideological alignment; $\widehat{\alpha}_t$ and $\widehat{\beta}_t$, which govern the algorithm's additional preferences over these dimensions; ζ_t, which measures the weight on expected user utility; and a_t and b_t, which map survey ideology into ideological preferences. The coefficients on observed content characteristics are estimated jointly with these parameters.

I rewrite the user and algorithm mean utilities as

$$W^n_{ij,t} = \delta^n_{j,t} - \beta_t (s^n_{j,t} - \mu_{i,t})^2, \tag{17}$$

$$\widehat{W}^n_{ij,t} = \widehat{\delta}^n_{j,t} - \widehat{\beta}_t (s^n_{j,t} - \mu_{i,t})^2 + \zeta_t \log(1 + \exp(W^n_{ij,t})), \tag{18}$$

where the post-specific mean utilities are

$$\delta^n_{j,t} = \alpha_t (s^n_{j,t} - \lambda_{j,t-1})^2 + g(X^n_{j,t}) + u^n_{j,t}, \tag{19}$$

$$\widehat{\delta}^n_{j,t} = \widehat{\alpha}_t (s^n_{j,t} - \lambda_{j,t-1})^2 + \widehat{g}(X^n_{j,t}) + \widehat{u}^n_{j,t}. \tag{20}$$

The functions $g(X^n_{j,t})$ and $\widehat{g}(X^n_{j,t})$ capture observed determinants of user engagement and

algorithmic recommendations, respectively. I include media-outlet, calendar-date, and hour-of-day fixed effects, together with the political-content indicator, text slant, sentiment, and text length. These characteristics capture systematic differences in news content and posting conditions that may affect views and engagement independently of headline slant.

To control for additional content heterogeneity, I incorporate semantic representations of the accompanying Facebook post text, following the use of pretrained embeddings in demand estimation by Compiani et al. (2026). Specifically, I encode the post's `text` field using the frozen pretrained model `BAAI/bge-small-en-v1.5`, which produces a 384-dimensional embedding. The online headline is excluded from this representation, so the embeddings describe the accompanying post text rather than the editorial choice being modeled.

I apply principal component analysis to the pooled embeddings and retain the leading components. To distinguish the additional semantic information from the interpretable content measures, I residualize each component with respect to the observed text characteristics, political-content indicator, and publication-time controls. The resulting residual components enter both $g(X^n_{j,t})$ and $\widehat{g}(X^n_{j,t})$ as additional controls for post-level content value.

Specifically, for principal component k, I estimate

$$PC^{n,k}_{j,t} = L^{n\prime}_{j,t}\pi_{k,t} + r^{n,k}_{j,t}, \tag{21}$$

where $L^n_{j,t}$ contains the interpretable content measures and publication-time controls. I then use $r^{n,k}_{j,t}$ as the additional semantic control. I retain the first 100 principal components in the baseline specification. The explanatory power of the embedding controls increases substantially with the number of components, but the incremental improvements become relatively small beyond 50 components. I therefore use 100 components to capture additional semantic heterogeneity while allowing for a sufficiently flexible specification of content value.

These controls explain a substantial share of the variation in recovered mean utilities, particularly for the recommendation algorithm. Figure 4 reports the adjusted R^2 of the

demand-side regressions over time. The explanatory power for the algorithm's mean utility is generally between 0.65 and 0.85, suggesting that observed content characteristics capture much of the systematic variation in the algorithm's valuation of news posts. The corresponding explanatory power for user engagement utility is lower but remains substantial.

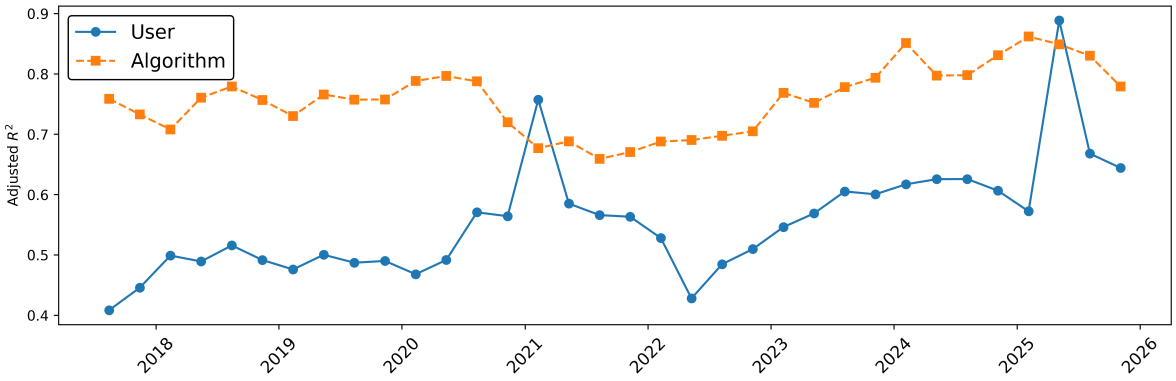


Figure 4: Model Fit for User Engagement and Algorithmic Recommendation Utility

Notes: The figure reports adjusted R^2 from period-specific regressions of recovered user engagement and algorithmic recommendation mean utilities. Each regression includes the structural slant terms, media-outlet, date, and hour-of-day fixed effects, political-content indicators, post-text slant, sentiment, and length, together with 100 residualized principal components of pretrained Facebook post-text embeddings. The embeddings exclude online headlines. Adjusted R^2 measures the explanatory power of the mean-utility regressions rather than the prediction accuracy of individual recommendation or engagement decisions.

This high explanatory power supports the use of rich content controls to account for differences in underlying newsworthiness that might otherwise be correlated with headline slant. Together with the model's timing assumptions, these controls make the conditional orthogonality restrictions on residual content value more plausible, although these restrictions remain maintained identification assumptions.

Identification of user preferences and algorithmic recommendation incentives relies on the joint distribution of views and engagement across posts with different characteristics and ideological slants. For a given post n, let $r_n(\mu)$ denote the probability of recommendation to a user with ideology μ, and let $e_n(\mu)$ denote the probability of engagement conditional on

recommendation. The model implies

$$S_n^V = \int r_n(\mu)\, dF_t(\mu), \tag{22}$$

$$S_n^E = \int e_n(\mu) r_n(\mu)\, dF_t(\mu). \tag{23}$$

Consequently, engagement conditional on viewing satisfies

$$\frac{S_n^E}{S_n^V} = \frac{\int e_n(\mu) r_n(\mu)\, dF_t(\mu)}{\int r_n(\mu)\, dF_t(\mu)}. \tag{24}$$

Thus, observed engagement per view reflects not only user preferences but also the composition of users selected to view each post by the recommendation algorithm. The model explicitly accounts for this selection through the personalized recommendation probabilities.

The externally calibrated ideological distribution provides the relative population masses of different user types, while variation in views and engagement across post slants provides information about their ideological preferences and the algorithm's ranking incentives. The model jointly fits both outcomes rather than treating engagement conditional on exposure as an independent measure of user preferences. Conditional on candidate nonlinear parameters and the ideological distribution, the two-layer share inversion recovers the post-specific mean utilities. Identification of the structural parameters additionally relies on restrictions on residual content value, as discussed below.

With the number of views $V_{j,t}^n$ and the number of engagements (like comments or shares) $E_{j,t}^n$ for each post, the demand-side parameters are estimated in two loops:

Inner loop (inversion of shares): Given candidate values $(\beta_t, \widehat{\beta}_t, \zeta_t, a_t, b_t)$ and the calibrated ideological group shares, I invert the predicted view and engagement shares to recover $\delta_{j,t}^n$ and $\widehat{\delta}_{j,t}^n$.

(i) Initialize $\delta_{j,t}^{n,(0)}$ and $\widehat{\delta}_{j,t}^{n,(0)}$.

(ii) Compute predicted shares under the model based on equation 9:

$$\widehat{S}^{n,E}_{j,t} = \Pr(I^n_{j,t} = 1, R^n_{j,t} = 1 \mid \beta_t, \widehat{\beta}_t, \zeta_t, \delta^{n,(0)}_{j,t}, \widehat{\delta}^{n,(0)}_{j,t})$$
$$\widehat{S}^{n,V}_{j,t} = \Pr(R^n_{j,t} = 1 \mid \beta_t, \widehat{\beta}_t, \zeta_t, \delta^{n,(0)}_{j,t}, \widehat{\delta}^{n,(0)}_{j,t})$$

(iii) Update using share inversion:

$$\delta^{n,(d+1)}_{j,t} = \delta^{n,(d)}_{j,t} + \log S^{n,E}_{j,t} - \log \widehat{S}^{n,E}_{j,t}$$
$$\widehat{\delta}^{n,(d+1)}_{j,t} = \widehat{\delta}^{n,(d)}_{j,t} + \log S^{n,V}_{j,t} - \log \widehat{S}^{n,V}_{j,t}$$

where $S^{n,E}_{j,t} = E^n_{j,t} / \sum_k \sum_m V^m_{k,t}$ and $S^{n,V}_{j,t} = V^n_{j,t} / \sum_k \sum_m V^m_{k,t}$ are the observed share of engagement and view. The denominator $\sum_k \sum_m V^m_{k,t}$ represents the number of views from all posts in period t, including non-news posts. Non-news content constitutes the outside option in the recommendation model, so the observed shares of the news posts included in estimation sum to less than one. See a proof that the above operation is a contraction mapping in Appendix C.

(iv) Repeat (ii) and (iii) until convergence.

Outer loop (structural estimation): Given the recovered mean utilities, I estimate

$$\delta^n_{j,t} = \alpha_t (s^n_{j,t} - \lambda_{j,t-1})^2 + L^{n\prime}_{j,t} \gamma_t + \sum_{k=1}^{K} \pi_{k,t} r^{n,k}_{j,t} + u^n_{j,t}, \tag{25}$$

$$\widehat{\delta}^n_{j,t} = \widehat{\alpha}_t (s^n_{j,t} - \lambda_{j,t-1})^2 + L^{n\prime}_{j,t} \widehat{\gamma}_t + \sum_{k=1}^{K} \widehat{\pi}_{k,t} r^{n,k}_{j,t} + \widehat{u}^n_{j,t}. \tag{26}$$

Here, $L^n_{j,t}$ contains the interpretable content characteristics and fixed effects, while $r^{n,k}_{j,t}$ denotes the residualized embedding components. The coefficients on these controls, together with α_t and $\widehat{\alpha}_t$, are estimated by least squares conditional on the nonlinear parameters.

I then select the nonlinear parameters to minimize the combined residual sum of squares:

$$\min_{\beta_t,\widehat{\beta}_t,\zeta_t,a_t,b_t} \sum_j \sum_n \left[(u_{j,t}^n)^2 + (\widehat{u}_{j,t}^n)^2\right]. \tag{27}$$

The procedure therefore combines a two-layer inversion of observed shares with minimum-distance estimation of user preferences, ideological locations, and algorithmic ranking parameters. I allow users with different ideologies to be recommended with posts with different probabilities, i.e., personalized recommendation, which is the core concept in algorithm design.

A central identification concern is that headline slant may be correlated with the underlying attractiveness of a news story. I address this concern by conditioning on interpretable content characteristics and high-dimensional semantic controls constructed from the accompanying Facebook post text, excluding the online headline. These controls capture substantial variation in content value, particularly in the algorithm's recommendation utility. Identification requires that, conditional on the observed content characteristics and other model states, the residual demand shocks are orthogonal to the outlet's headline choices. This restriction is consistent with the model's information structure: outlets observe the predictable content values $g(X_n)$ and $\widehat{g}(X_n)$ when choosing headlines, but not the residual shocks u_n and $\widehat{u}_n$, which are subsequently observed by the algorithm. The rich content controls reduce the scope for unobserved newsworthiness to confound the estimated slant preferences, while the two-layer demand system jointly accounts for algorithmic selection and users' engagement decisions.

6.2 Media Outlets (Supply Side)

On the supply side, each media outlet chooses slant adjustments $a_{j,t}^n \in \{-1, 0, 1\}$ conditional on the state variables $\theta_{j,t}$ and observable post characteristics $\omega_{j,t}^n = \{p_{j,t}^n, X_{j,t}^n\}$. Each action results in a different headline slant $s_{j,t}^n$ according to Equation 11. When evaluating its choices,

the outlet uses the estimated demand model to predict expected views and engagement, integrating over the residual content-value shocks $u_{j,t}^n$ and $\widehat{u}_{j,t}^n$ that are unobserved when the editing decisions are made.

The outlet-level parameters to estimate are τ_j^V, τ_j^E, $c_{-1,j}$, $c_{+1,j}$, and β_j. The parameters τ_j^V and τ_j^E measure the outlet's value of views and engagement, respectively, while β_j is its discount factor. The cost of no slant adjustment is normalized to zero, with $c_{-1,j}$ and $c_{+1,j}$ representing the costs of shifting headline slant to the left and right, respectively.

The structural model assumes that outlets jointly choose the slant of all N posts within a period, resulting in an action space of size 3^N. Solving this problem directly is computationally intractable for the number of posts observed in the data. For estimation, I therefore approximate the joint editing problem with a sequential decision process in which editors choose the slant of one post at a time. Each period corresponds to a day, and the outlet knows the number of stories N to be published at the beginning of the period. The sequential formulation reduces each decision to three alternatives while allowing current editing choices to affect future payoffs through the evolution of the outlet's expected slant. The approximation preserves the dynamic reputation mechanism but does not require solving jointly over all possible combinations of headline adjustments. Then the value function of media outlet j becomes

$$\begin{aligned} V_j\Big(\omega^n, \hat{\lambda}, \lambda_-, N, \theta\Big) = \max_{a^n \in \{-1,0,1\}} & \ \tau_j^V U^V(a^n) + \tau_j^E U^E(a^n) - c_{a^n} \\ & + v_a^n + \mathbb{E}\left[V_j\left(\omega^{n+1}, \hat{\lambda} + \frac{s^n(a^n)}{N}, \lambda_-, N, \theta\right)\right] \end{aligned}$$

where $v_a^n \sim i.i.d. EV1\,(0,1)$ and $U^V(a^n), U^E(a^n)$ are the expected probabilities of view and engagement, which depends on $\omega^n, \lambda_-, \theta$ and the distribution of s^n conditional on a^n. $\hat{\lambda}$ is used for updating the expected slant at the end of the period. In the special case of the last

post in the period, the value function becomes

$$V_j\Big(\omega^N, \hat{\lambda}, \lambda_-, N, \theta\Big) = \max_{a^N \in \{-1,0,1\}} \tau_j^V U^V(a^N) + \tau_j^E U^E(a^N) - c_{a^N} + \upsilon_a^N + \beta_j \mathbb{E}\Big[V_j\Big(\omega^{1'}, 0, \lambda, N', \theta'\Big)\Big]$$

where $\lambda = (1-\rho)\lambda_- + \rho(\hat{\lambda} + \frac{s^n(a^n)}{N})$ is how the expected slant being updated and $\omega^{1'}$ is the property of the first news in next period. Once the game enters the next period, a new N' is drawn from the empirical distribution and the next period state θ' is drawn from the known transition probabilities $G(\theta'|\theta)$.

With the sequential approximation, I estimate the outlet's dynamic incentives using a two-step conditional choice probability (CCP) approach following Hotz and Miller (1993). The first step estimates editorial choice probabilities conditional on observed states. The second step uses these probabilities and simulated continuation values to recover the structural payoff parameters.

Let $\Theta = \{n, p^n, X^n, \widehat{\lambda}, \lambda_-, N, \theta\}$ collect the observable state variables. Since editing decisions are made sequentially within each day, continuation values are not discounted between stories within the same day, while future payoffs are discounted by β_j when transitioning to the next day. The deterministic current payoff from choosing action $a \in \{-1, 0, 1\}$ is

$$\begin{aligned} Q_j(a \mid \Theta) = \tau_j^V U^V(a \mid \Theta) + \tau_j^E U^E(a \mid \Theta) \\ - c_{-1,j}\mathbf{1}\{a = -1\} - c_{+1,j}\mathbf{1}\{a = +1\} \end{aligned} \tag{28}$$

where $U^V(a \mid \Theta)$ and $U^E(a \mid \Theta)$ are the expected view and engagement probabilities computed from the estimated demand model. These expectations integrate over the content-value shocks unobserved by the outlet when choosing its headline.

Let $C_j(a \mid \Theta)$ denote the expected discounted continuation value following action a. Under the i.i.d. type-I extreme-value assumption for private action shocks, the conditional

choice probabilities satisfy

$$P_j(a \mid \Theta) = \frac{\exp\left(Q_j(a \mid \Theta) + C_j(a \mid \Theta)\right)}{\sum_{a' \in \{-1,0,1\}} \exp\left(Q_j(a' \mid \Theta) + C_j(a' \mid \Theta)\right)} \tag{29}$$

For a given discount factor β_j, I approximate the continuation values by simulating future states and editorial choices using the estimated CCPs, following Hotz et al. (1994). The continuation value can be decomposed into expected future views, engagement, and editing costs, which are linear in the structural payoff parameters. Let $\phi(a \mid \Theta)$ and $\psi(a \mid \Theta)$ denote the expected discounted sums of future views and engagement following action a, respectively. Similarly, let $\kappa_{-1}(a \mid \Theta)$ and $\kappa_{+1}(a \mid \Theta)$ denote the expected discounted numbers of future leftward and rightward headline adjustments.

The continuation value must also account for future action-specific utility shocks. Although these shocks are independent across actions, the expected shock associated with the optimally chosen action is generally nonzero. Under the extreme-value assumption, this conditional expectation is $\gamma - \log P_j(a \mid \Theta)$, where γ is Euler's constant. I denote the expected discounted contribution of these future selected action shocks by $H_j(a \mid \Theta)$, which can be calculated from the estimated CCPs. The continuation value can therefore be expressed as

$$\begin{aligned} C_j(a \mid \Theta) = \tau_j^V \phi(a \mid \Theta) + \tau_j^E \psi(a \mid \Theta) \\ - c_{-1,j}\kappa_{-1}(a \mid \Theta) - c_{+1,j}\kappa_{+1}(a \mid \Theta) + H_j(a \mid \Theta) \end{aligned} \tag{30}$$

All continuation statistics incorporate the appropriate discounting across periods and are computed by forward simulation conditional on the current action and state. For a given β_j, they are functions of the estimated CCPs, demand-side primitives, and state-transition process, rather than the structural payoff parameters being estimated.

Let $a_0 = 0$ denote the reference action, and define $\Delta f(a \mid \Theta) = f(a \mid \Theta) - f(a_0 \mid \Theta)$ for any action-specific function f. The logit structure implies that the difference in choice-specific values equals the log ratio of the corresponding CCPs. I construct the following

regressors:

$$\begin{aligned}
\Phi(a \mid \Theta) &= \Delta U^V(a \mid \Theta) + \Delta\phi(a \mid \Theta), \\
\Psi(a \mid \Theta) &= \Delta U^E(a \mid \Theta) + \Delta\psi(a \mid \Theta), \\
K_i(a \mid \Theta) &= \mathbf{1}\{a = i\} + \Delta\kappa_i(a \mid \Theta), \quad i \in \{-1, +1\}
\end{aligned} \tag{31}$$

Substituting the continuation-value decomposition into the log-CCP equation yields

$$\begin{aligned}
\log \frac{P_j(a \mid \Theta)}{P_j(a_0 \mid \Theta)} - \Delta H_j(a \mid \Theta) = \tau_j^V \Phi(a \mid \Theta) + \tau_j^E \Psi(a \mid \Theta) \\
- c_{-1,j} K_{-1}(a \mid \Theta) - c_{+1,j} K_{+1}(a \mid \Theta) + \varsigma_j(a \mid \Theta)
\end{aligned} \tag{32}$$

where $\varsigma_j(a \mid \Theta)$ is the econometric residual. For each candidate discount factor β_j, I construct the continuation statistics through forward simulation. Conditional on these objects, Equation 32 is linear in τ_j^V, τ_j^E, $c_{-1,j}$, and $c_{+1,j}$, which I estimate by least squares. I select $\widehat{\beta}_j$ to minimize the resulting residual sum of squares.

This approach combines the CCP representation of Hotz and Miller (1993) with simulation-based evaluation of dynamic continuation values, as in Hotz et al. (1994). It is also related to the two-step estimation of dynamic games developed by Bajari et al. (2007). In this application, the outlet's flow payoffs are determined by the estimated two-layer demand system, allowing the model to quantify how algorithmic recommendation incentives affect dynamic editorial decisions.

7 Empirical Results

This section presents the empirical results from both the demand (user and algorithm) and supply (media outlet) sides of the model. The estimation combines revealed-preference data on engagement and visibility from 3 million Facebook posts with matched print and online

headlines from eleven major newspapers. The results quantify: (i) how users respond to ideological features of news content; (ii) how algorithms amplify those preferences; and (iii) how media outlets dynamically adjust their headline slant to maximize performance under algorithmic feedback.

7.1 Algorithmic Amplification and User Preferences

Figure 5 plots the estimated coefficients on political content in user utility and in the algorithm's recommendation utility. The user-side coefficient is positive throughout the sample and generally rises over time, indicating that, conditional on the other content controls, users tend to derive higher utility from political content than from non-political content. This pattern is broadly consistent with the high engagement rates observed for political news on social media.

The algorithm-side coefficient follows a markedly different pattern. Before Facebook's effort to reduce political content, the estimated coefficient is close to zero and occasionally positive, suggesting that the recommendation system did not systematically demote political content relative to other posts. Beginning with the 2021–2022 rollout of the policy reducing political content, however, the algorithm-side coefficient becomes sharply negative and remains so through much of 2022–2024. This implies that, holding predicted user utility and other observable content characteristics fixed, the platform directly penalized political content in ranking during this period. Around the 2025 restoration, the coefficient rises substantially toward zero and becomes slightly positive by the end of the sample. Thus, the structural estimates suggest that Facebook's changing treatment of political content operated in part through a direct shift in how the algorithm ranked political posts.

Figure 6 complements this result by plotting the estimated coefficient ζ_t, which governs the extent to which the recommendation algorithm places additional weight on predicted user utility. Before the 2022 policy change, ζ_t is generally positive and often large, indicating that the platform gave substantial additional weight to posts predicted to generate high user

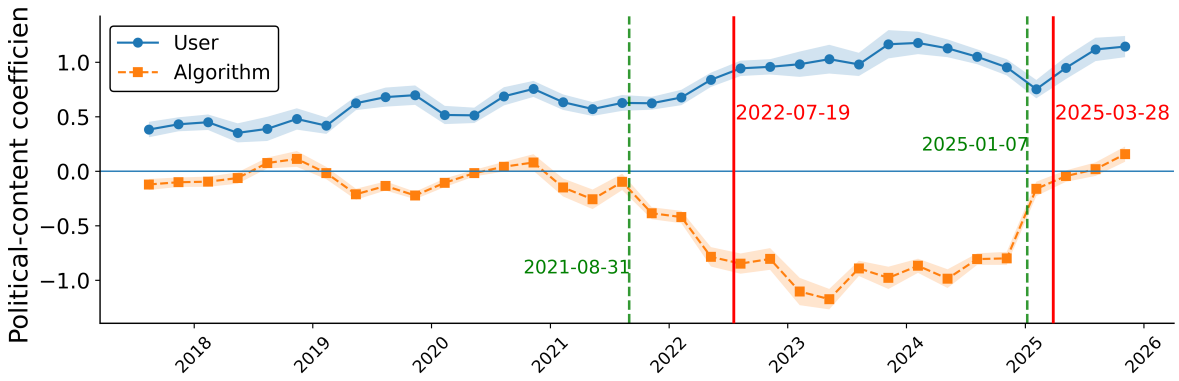


Figure 5: Estimated Political-Content Coefficients in User and Algorithm Utility

Notes: The figure reports estimates of the coefficients on political content in the user-engagement utility and in the algorithm's recommendation utility, using 91-day intervals. The blue solid line shows the user-side coefficient, while the orange dashed line shows the algorithm-side coefficient. Shaded bands indicate confidence intervals. The green dashed lines mark August 31, 2021 and January 7, 2025, corresponding to the beginning of Facebook's reduction in political content and the announcement of its restoration. The red solid lines mark July 19, 2022 and March 28, 2025, corresponding to later stages of implementation of the two policy changes.

utility beyond users' direct preferences over content. During the period in which Facebook reduced political content, ζ_t declines sharply, reaching values near zero or below zero for much of 2022–2024. This pattern is consistent with Meta's description that the platform would place less emphasis on engagement-related signals for political content. In the model, lower values of ζ_t mean that the algorithm is less willing to amplify content simply because it is predicted to generate high engagement utility.

The pattern reverses around Facebook's 2025 restoration of political content. Following the January 7 announcement and the subsequent implementation period, ζ_t rises again from the low levels observed in the preceding years. Although the estimates remain more volatile than in the pre-2022 period, the rebound suggests that renewed reliance on predicted user utility was one mechanism through which political content was brought back into users' feeds.

Taken together, Figures 5 and 6 point to two distinct channels through which Facebook's recommendation system changed over time. First, the platform directly altered its treatment of political content, moving from active demotion during 2022–2024 to a much more neutral

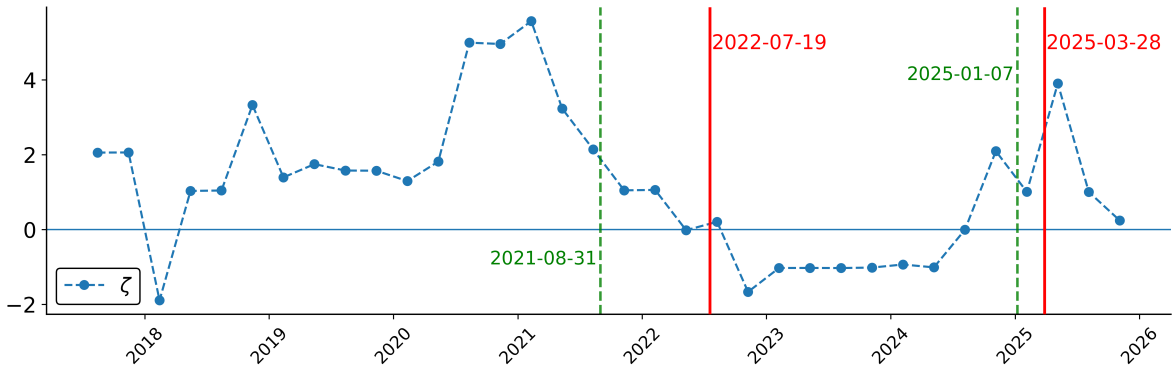


Figure 6: Evolution of Algorithmic Emphasis ζ_t

Notes: The figure reports estimates of ζ_t in 91-day intervals, where ζ_t governs the additional weight placed by the recommendation algorithm on predicted user utility. The green dashed lines mark the beginning of Facebook's reduction in political content on August 31, 2021, and the announcement of its restoration on January 7, 2025. The red solid lines mark July 19, 2022 and March 28, 2025, corresponding to the later stages of implementation of the two policy changes. The estimates decline substantially during the reduction of political content and rise again around its 2025 restoration.

or slightly favorable stance after the 2025 restoration. Second, it changed the extent to which ranking responded to predicted user utility, reducing that emphasis during the period of political-content suppression and raising it again once political content was restored. These structural results align with the reduced-form evidence: the reduction of political content is associated with weaker algorithmic amplification of political material, while the 2025 restoration is associated with a renewed expansion of political content and a larger online–print slant gap for political stories.

Table 4 reports the estimated demand-side coefficients prior to the algorithmic change. The parameters α_t and $\hat{\alpha}_t$ capture the salience effect from surprise or the trustworthiness from the unexpected slant by the users and the algorithm; the parameters β_t and $\hat{\beta}_t$ capture how strongly users prefer like-minded news. ζ_t represents the corresponding algorithmic weights on users engagements. The results show that the algorithm amplifies both dimensions of user preference through two channels: the estimated ζ_t is averaged around 2, implying that recommendation design roughly doubles the effective incentive toward like-minded and salient content through the extra weights on engagements. On the other hand, $\hat{\alpha}_t$ and

$\hat{\beta}_t$ are almost at the same level with α_t and $\hat{\beta}_t$, suggesting that the algorithm has extra consideration on the two slant related terms. In total, the algorithm amplification on slant related effect is around three times as much as the users original preferences. For the supply-side estimation, I use the sum of comments and shares as the engagement measure, since these actions are typically more costly and reflect stronger user intent than simple reactions. Moreover, as shown in Aldous et al. (2025), fewer than 1% of combined engagements on social media (reactions, shares, comments) come from the same user, suggesting that these metrics capture aggregate audience responses rather than repeated actions by a small set of users.

Table 4: Demand and Algorithm Parameters Before the Algorithmic Change

Engagements	Comments+Shares	Comments	Reactions
α	0.229 (0.027)	0.229 (0.030)	0.118 (0.031)
β	0.329 (0.034)	0.347 (0.037)	0.179 (0.039)
$\hat{\alpha}$	0.190 (0.023)	0.193 (0.029)	0.162 (0.032)
$\hat{\beta}$	0.244 (0.028)	0.261 (0.034)	0.222 (0.036)
ζ	2.229 (0.177)	2.166 (0.305)	1.255 (0.206)

Notes: Parameters α and β capture user-level preferences for surprise and ideological alignment; $\hat{\alpha}$ and $\hat{\beta}$ capture the corresponding algorithmic preferences; ζ measures how much the algorithm values user utility relative to engagement prediction. Standard errors in parentheses.

With the estimated model, I simulate the composition of slanted posts that are recommended to (viewed by) and engaged with by users. Figure 7 reports the percentage change in the share of liberal-leaning and conservative-leaning posts at the viewed and engaged stages relative to their baseline supply levels.

Figure 7 shows that ideological asymmetries arise both in exposure and in engagement behavior. Although news influencers, who are mostly conservative-leaning, are included in the dataset, most large media outlets are liberal-leaning and post more frequently. As a result, the supply of posts is tilted toward liberal-leaning slant. Comparing the supplied and viewed panels, the model implies that users are more likely to be exposed to ideologically

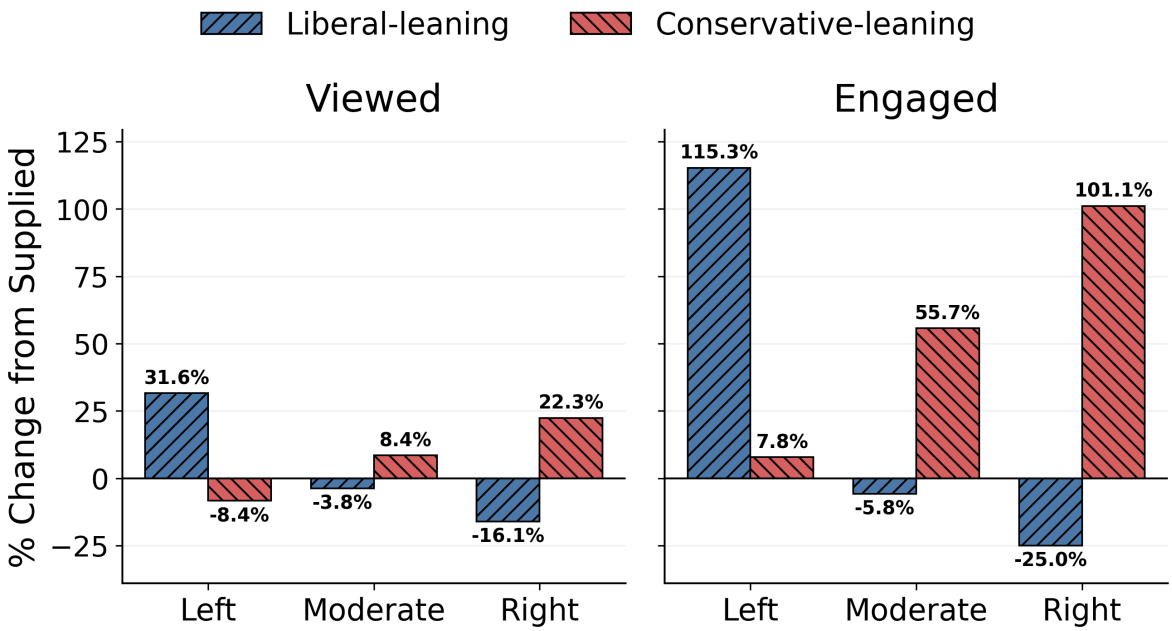


Figure 7: Percentage change in the share of ideologically slanted posts relative to supply. The figure reports the change in the share of liberal-leaning (negative slant) and conservative-leaning (positive slant) posts at the viewed and engaged stages relative to their baseline supply shares. The shares are simulated using the estimated model, with engagement measured by Comments + Shares. For presentation, the five ideological types used in estimation are aggregated into three groups: Left combines very liberal and liberal users, Moderate represents moderate users, and Right combines conservative and very conservative users. Posts are classified as liberal-leaning, conservative-leaning, or neutral based on their estimated slant. Positive values indicate that a given type of content is overrepresented relative to supply, while negative values indicate underrepresentation. The figure highlights how algorithmic curation and user engagement behavior jointly distort the composition of content away from the underlying supply.

aligned posts, consistent with the estimated results. In the engaged panel, all user groups are more likely to engage with conservative-leaning posts relative to their viewed (recommended) share. In contrast, only left-leaning users exhibit a higher engagement rate for liberal-leaning posts, while moderate and right-leaning users do not display a comparable tendency. This pattern may reflect the role of unexpected slant: since most large media outlets are liberal-leaning, their conservative-leaning posts may attract more attention and generate higher engagement.

Overall, these results confirm two key insights. First, readers exhibit strong preferences for both like-minded news and unexpected slant—consistent with selective exposure and

curiosity for ideological deviation. Second, Facebook's algorithm prior to mid-2022 substantially amplified these preferences, especially for politically slanted content.

7.2 Media Outlets' Objective and Slant Choice

Turning to the supply side, the estimated parameters $(\tau_j^V, \tau_j^E, c_{-1,j}, c_{+1,j}, \beta_j)$ reveal how outlets value views and engagements relative to the cost of headline adjustment. Figure 8 plots the estimated τ_j^V and τ_j^E coefficients by outlet, scaled by their follower counts. Since views and engagements enter linearly in my model, it implicitly assumes that outlets care about the first view or engagement the same as when they have thousands of them. However, one should expect those with fewer views or engagements have much higher marginal value on the metrics. In fact, the raw estimation shows that smaller regional newspapers have much higher τ_j^V and τ_j^E than bigger national newspapers. Therefore, I report the estimation scaled by their follower counts on Facebook, which bring the numbers to the same magnitude. However, this does not change how significant the estimates are.

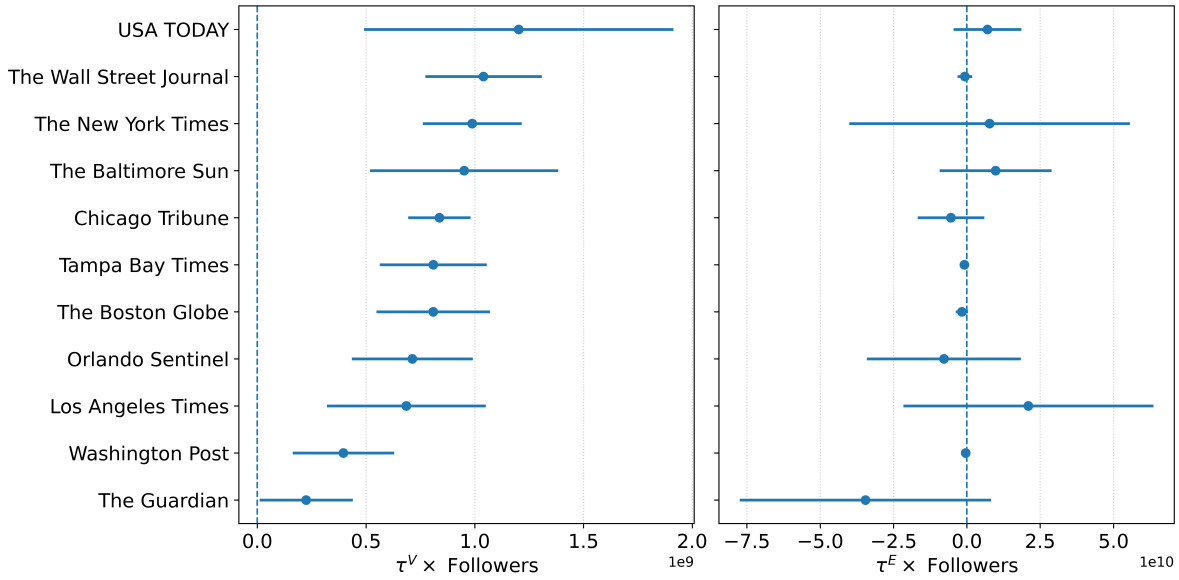


Figure 8: Estimated media-outlet weights on views and engagement (τ_j^V, τ_j^E), scaled by social-media follower counts. All outlets place significantly positive weight on view traffic but not engagement metrics.

The estimates indicate that all outlets prioritize maximizing views but not engagements,

suggesting that algorithms' visibility mechanisms dominate editorial incentives. Even though the algorithm takes users' preference into consideration when it recommends posts, as the decision on which post is viewed by the user is solely made by the algorithm and media outlets only care about views, the algorithm becomes the most critical component in shaping online news slant.

Figure 9 reports the estimated costs of left- and right-leaning headline adjustments, $c_{-1,j}$ and $c_{+1,j}$. These costs capture editorial frictions on adding extra slant to the existing print headlines. When $c_{+1,j} > c_{-1,j}$, it is more costly for the outlet to add right-leaning slant, which suggests a more liberal leaning editorial preference. And vice versa. The estimated cost asymmetries are modest on average, except Los Angeles Times exhibits a strong liberal leaning preference as its $c_{+1,j} > c_{-1,j}$ by a big margin.

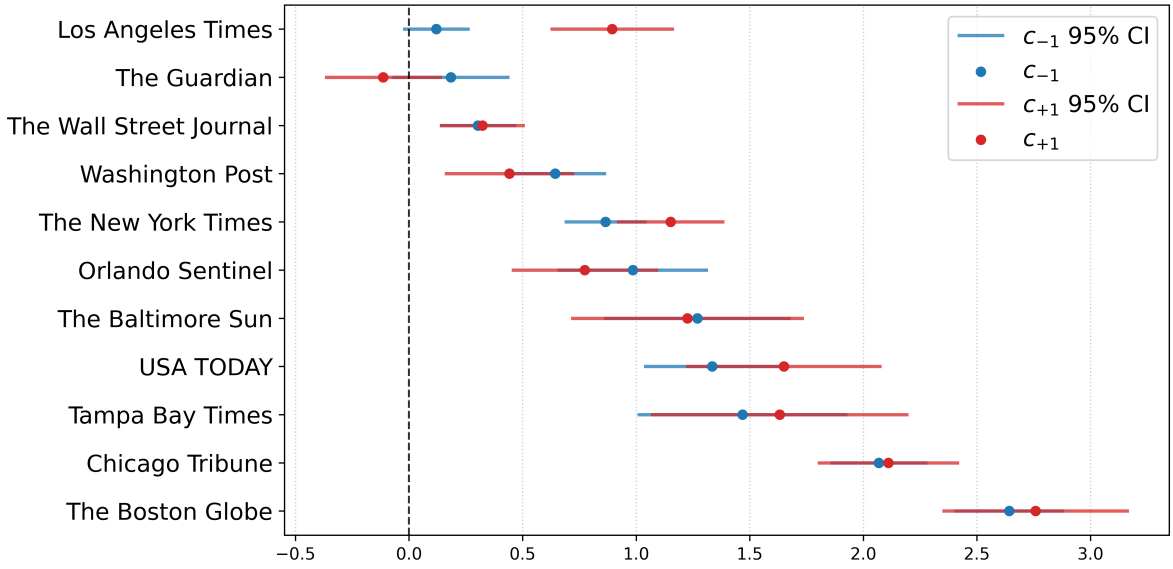


Figure 9: Estimated costs of left- and right-slant adjustments ($c_{-1,j}$ and $c_{+1,j}$). Costs capture reputational and editorial frictions associated with slant changes. Right-leaning deviations are on average more costly, reflecting both outlet branding and reader expectations.

The combination of demand and supply estimates provides a coherent quantitative picture of how social-media algorithms shape the production of news. Before 2022, Facebook's algorithm introduced substantial additional incentives for slanted content relative to intrinsic user demand. Media outlets responded by shifting the slant of online headlines toward more

slanted language, consistent with their estimated τ_j^V values emphasizing view-maximization. Following Facebook's policy change to downweight political content, both algorithmic amplification (ζ_t) and outlet-level slant adjustments declined measurably, confirming that algorithm design can influence not only what users see but also what media outlets produce.

8 Counterfactual Analysis

This section uses the estimated structural model to quantify how much of the observed ideological slant on social media can be attributed to algorithmic amplification. The analysis compares the original equilibrium to several counterfactual scenarios in which the recommendation algorithm's sensitivity to ideological and engagement signals is modified. All counterfactuals are simulated holding the estimated demand-side and supply-side parameters fixed, and recomputing equilibrium outcomes for post-level viewing, engagement, and editorial slant choices.

8.1 Counterfactual Design

The observed case represents the upper bound on algorithmically induced polarization. In this scenario, the algorithm fully amplifies user-level ideological preferences and engagement responses through positive values of $\hat{\alpha}_t$, $\hat{\beta}_t$, and ζ_t. This is the environment estimated from the data, corresponding to the original equilibrium during 2017–2022 when Facebook's feed emphasized engagement metrics.

To isolate the mechanisms, I consider a sequence of counterfactual experiments that sequentially remove different components of algorithmic amplification:

- Scenario 1: No Extra Effect on Slant. The algorithm no longer rewards ideological alignment beyond users' own preferences. Formally, I set $\hat{\alpha}_t = \hat{\beta}_t = 0$, keeping ζ_t fixed. This removes algorithmic reinforcement of slant but maintains engagement weighting.

- Scenario 2: Ignore Engagement. The algorithm ceases to prioritize predicted engagement in its recommendation function, i.e., $\zeta_t = 0$. It still perceives ideological distances, but places no additional value on posts with higher expected engagement. This experiment identifies the importance of engagement-based amplification separate from ideological bias.

- Scenario 3: No Algorithm Effect on Slant. A fully neutral algorithm that neither rewards ideological similarity nor engagement intensity: $\hat{\alpha}_t = \hat{\beta}_t = \zeta_t = 0$. The decision is only based on the costs and content values in this case.

For reference, I also compute a hypothetical lower bound in which all parameters related to ideological distance are shut down—both user and algorithmic components: $\alpha_t = \beta_t = \hat{\alpha}_t = \hat{\beta}_t = 0$ as the Benchmark Scenario. This represents the most extra slant that could ever be removed, since it eliminates ideological differentiation altogether. In practice, total slant cannot fall to zero even under this benchmark because the logit form of outlet's value function (υ_a^n) allow outlets to retain some dispersion in equilibrium slant purely due to randomness and unobserved heterogeneity.

The model simulates the results below by iterating the algorithmic recommendation and outlet choice functions under each parameter restriction. Outlets adjust their post slant choices through the estimated policy functions $P_j(a \mid \Theta)$, and the resulting equilibrium slant distribution is compared to the baseline.

8.2 Effects on Slant: Supply vs Exposure

For each counterfactual scenario, I compute two distinct measures of polarization. The first is the *supplied extra slant*, defined as the average difference between the slant of online posts and their corresponding print headlines. The second is the *viewed extra slant*, which measures the same difference but weighted by the equilibrium probability that each post is recommended and viewed by users.

Formally, the supplied extra slant captures how much media outlets adjust their content, while the viewed extra slant reflects the distribution of information that users are actually exposed to after algorithmic recommendation. Comparing the two allows me to separate changes in editorial behavior from changes in exposure induced by the platform.

Table 5: Counterfactual Reductions in Slant: Supplied vs Viewed

Scenario	Supplied Reduction (%)	Viewed Reduction (%)
Scenario 1	89.52%	90.24%
Scenario 2	16.82%	9.24%
Scenario 3	100.00%	99.86%

Notes: Each scenario recomputes equilibrium slant, visibility, and engagement given the parameter restrictions described above. Supplied reduction is based on the average extra slant across posts, while viewed reduction weights each post by its equilibrium probability of being recommended. All reductions are normalized relative to the benchmark scenario, which provides the maximum achievable reduction in slant.

Table 5 reports the percentage reduction in extra slant under each scenario for both measures, normalized relative to the benchmark case as described above. Several patterns emerge from the comparison between supplied and viewed slant.

First, removing only the algorithm's ideological weighting $(\hat{\alpha}_t, \hat{\beta}_t)$ in Scenario 1 yields a major reduction in extra online slant, indicating that only a small part of the bias originates from users' own preferences. Shutting down engagement weighting ($\zeta_t = 0$) in Scenario 2 produces a significant but not major decline, showing that the platform's weight on predicted user utility is a not major driver of polarized content.

Second, the reduction in viewed slant does not generally coincide with the reduction in supplied slant. While Scenario 1 delivers nearly identical reductions in both measures, Scenario 2 exhibits a substantial divergence: it achieves a modest reduction in supplied slant but a noticeably smaller reduction in viewed slant. This indicates that even when outlets adjust their content, the algorithm can continue to concentrate attention on relatively more slanted posts.

Finally, the fully neutral algorithm scenario ($\hat{\alpha}_t = \hat{\beta}_t = \zeta_t = 0$) achieves the full reduction. As newspapers only care about views but not engagement, this scenario effectively shuts down

all channels through which slant affects outlets' decision making problem.

Overall, the counterfactual analysis highlights the contribution of algorithmic amplification to the equilibrium polarization of online news.

8.3 Algorithm Effect

The previous results distinguish between extra slant at the content level and extra slant in user exposure. To quantify the role of the algorithm in bridging these two objects, I decompose viewed extra slant into a component attributable to media supply and a component attributable to algorithmic amplification.

For each scenario, let G^V denote viewed extra slant and G^S denote supplied extra slant. I define the algorithmic share as

$$\text{Algorithm Share} = \frac{G^V - G^S}{G^V}.$$

This measure captures the fraction of viewed extra slant that cannot be explained by editorial choices alone, and is therefore attributed to the recommendation algorithm's allocation of attention across posts.

Table 6: Decomposition of Viewed Extra Slant

Scenario	Algorithm Share (%)
Benchmark	12.65
Scenario 1	13.33
Scenario 2	18.57
Scenario 3	12.66
Original	19.10

Notes: Supplied extra slant measures the average difference between online and print slant across posts, while viewed extra slant weights each post by its equilibrium probability of being recommended. Algorithm share is defined as the fraction of viewed extra slant not accounted for by supplied extra slant.

Table 6 reports the shares across all scenarios, including the original equilibrium and the benchmark. Several insights emerge from this decomposition.

First, the algorithm plays a quantitatively important role in shaping exposure. In the original equilibrium, approximately 19% of viewed extra slant is attributable to the algorithm, indicating that recommendation behavior significantly amplifies polarization beyond what is generated by media supply alone. At the same time, media outlets' response to algorithm accounts for the majority of viewed extra slant.

Second, a nontrivial share of viewed extra slant persists even in scenarios where slant is irrelevant in the algorithm's objective. In both the benchmark and Scenario 3, the algorithm share remains around 12–13%. One possible explanation is that media outlets differ in their baseline prominence and editorial capacity. More established outlets are more likely to engage in strategic headline editing and are also more likely to receive attention from users. As a result, differences in exposure across outlets can generate residual viewed extra slant even in the absence of explicit algorithmic amplification.

Finally, Scenario 1 delivers an outcome close to this lower bound. Removing the algorithm's direct sensitivity to ideological slant ($\hat{\alpha}_t = \hat{\beta}_t = 0$) reduces the algorithm share to a level similar to the benchmark and Scenario 3. This again indicates that most of the algorithm's contribution to polarization operates through its amplification of slant-related features. Scenario 2 highlights the role of engagement-based amplification. Although shutting down the engagement channel ($\zeta_t = 0$) reduces overall slant, the algorithm share remains relatively high, consistent with previous results.

Overall, these results show that algorithms accounts for a sizable share of slant exposure and interacts with endogenous supply decisions. Understanding both channels is essential for evaluating how platform design affects the equilibrium distribution of news exposure.

8.4 Welfare and Diversity Trade-offs

The counterfactual analysis reveals a central trade-off between short-run user utility and the diversity of information exposure.

To quantify welfare, I compute the total expected user utility under each scenario using

the estimated demand system. The results show that removing algorithmic amplification reduces user utility. Scenarios that shut down the algorithm's sensitivity to slant or engagement lead to substantial utility losses relative to the original equilibrium. This is expected, as the algorithm amplifies components of content that users find engaging, including ideological alignment and salience.

However, these welfare losses are accompanied by important changes in the distribution of viewed content. Figure 10 shows the share of viewed slant categories (negative, neutral, positive) across users with different ideological positions under each scenario. The original equilibrium exhibits clear polarization in exposure: left-leaning users disproportionately consume negatively slanted content, while right-leaning users consume more positively slanted content. This sorting reflects both user preferences and algorithmic amplification.

Scenario 3 delivers the most diverse exposure environment, corresponding to the case in which recommendation has no consideration on slant. In this scenario, users across the ideological spectrum are exposed to the same mix of content, with less concentration in any ideological categories. Scenario 1 achieves a very similar degree of flattening in exposure.

Importantly, these scenarios differ substantially in their welfare implications. While Scenario 3 generates a large welfare loss of approximately 75%, Scenario 1 results in a much smaller loss of about 20%. This suggests that it is possible for the platform to provide a more balanced and diverse information environment without substantially reducing user utility. In contrast, Scenario 2—closely related to Facebook's policy change—produces a comparable welfare loss (around 19.5%) but leads to only limited improvements in exposure diversity.

Taken together, these results highlight a fundamental trade-off. Algorithmic amplification increases short-run user utility by directing attention toward content that users are more likely to engage with, but at the cost of greater polarization in exposure. Conversely, reducing amplification lowers engagement-based utility while improving the diversity and balance of information consumption. This trade-off suggests that platform design involves balancing engagement objectives against broader diversity outcomes. While engagement-maximizing

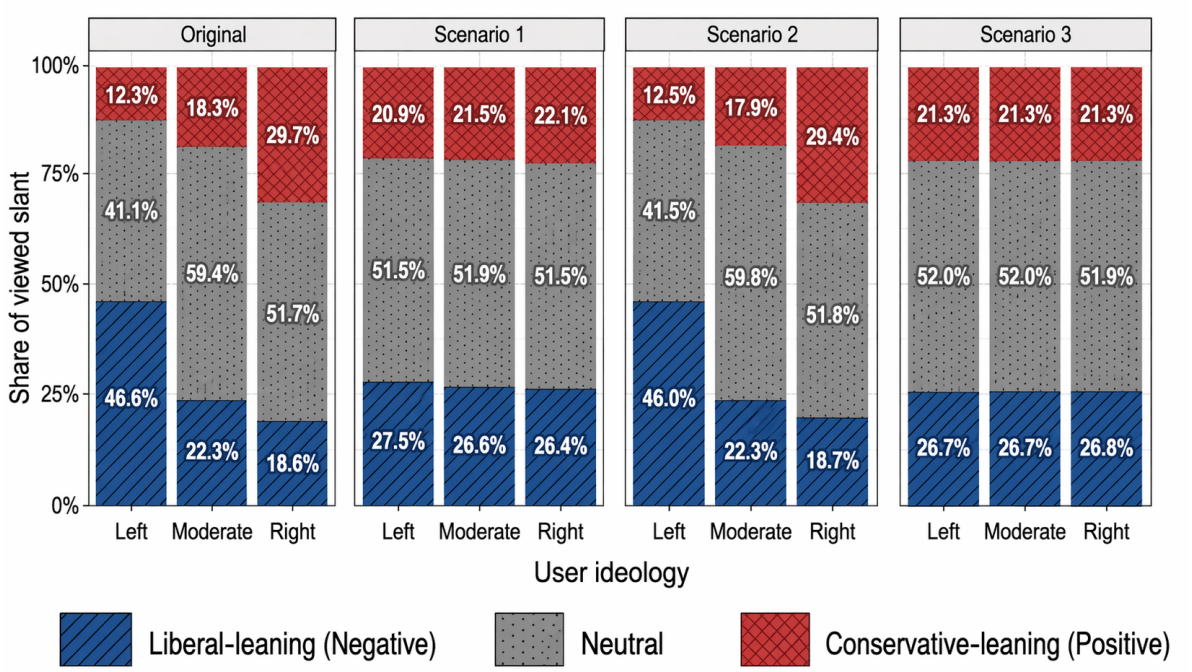


Figure 10: Share of viewed slant by user ideology across counterfactual scenarios. The figure reports the distribution of viewed slant categories (negative, neutral, positive) across user ideology groups under each scenario. For presentation, the five ideological types used in estimation are aggregated into three groups: Left combines very liberal and liberal users, Moderate represents moderate users, and Right combines conservative and very conservative users. For each group, the shares sum to one and reflect the composition of content that users are exposed to in equilibrium. The measure is based on the slant of viewed posts, rather than extra slant relative to print benchmarks. Differences across ideology groups capture polarization in exposure, while differences across scenarios reflect the effects of algorithmic and supply-side changes on the distribution of content.

algorithms increase immediate user satisfaction, they may also contribute to segregated information environments with potential long-run consequences for belief formation and media bias.

9 Policy Implications

The results of this paper have direct implications for the design and regulation of social-media platforms. A central insight is that algorithmic curation does not merely reallocate attention across existing content; it alters the incentives of media outlets and thereby reshapes the equilibrium supply of news. As a result, policies that focus solely on user exposure or short-

run behavioral responses may fail to address the underlying source of distortion. Effective intervention must account for how platform design affects both demand (what users see) and supply (what outlets choose to produce).

9.1 Supply-Side Responses and Regulatory Monitoring

An important implication is that evaluations of recommendation systems should account for endogenous responses by content producers. In the model, media outlets choose their online slant partly to compete for algorithmically allocated visibility. A change in the recommendation rule therefore affects not only which existing posts users see, but also the content that publishers subsequently choose to supply. Evaluating a recommendation system solely from the allocation of a fixed set of content can consequently miss an important equilibrium effect.

The regression evidence illustrates this distinction. Following Facebook's 2025 restoration of political content, political posts receive substantially more views, and outlets increase the extent to which their online headlines depart from the print headlines of the same stories. At the same time, engagement per view does not increase for political posts and instead declines across both political-content classifications. This pattern does not support the simple explanation that the increase in exposure and editorial adjustment is driven by a contemporaneous increase in users' propensity to engage with political content. Complementary evidence from Facebook's earlier reduction in political-content distribution points in the opposite direction: as the platform reduced the distribution of political content, the online–print slant gap declined.

These results suggest that regulatory monitoring should distinguish between changes in *exposure*, changes in *user response conditional on exposure*, and changes in *content supply*. Exposure-based audits can measure whether content with different characteristics receives systematically different amounts of algorithmically allocated attention. Engagement conditional on exposure provides complementary evidence about whether changes in exposure

are accompanied by corresponding changes in users' revealed propensity to engage. Finally, tracking publishers' content choices—for example, by comparing online and print presentations of the same underlying stories—can reveal whether changes in recommendation incentives induce supply-side editorial responses.

This distinction matters for the interpretation of platform regulation. A policy that changes recommendation probabilities can have effects beyond the posts whose exposure is directly altered: publishers may adapt their content in anticipation of the new allocation rule. Regulatory assessments that ignore these endogenous supply responses may therefore understate the effect of recommendation systems on the ideological composition of the news ultimately consumed by users. Importantly, these forms of monitoring do not require direct access to the platform's recommendation code; they can instead use observed changes in exposure, conditional engagement, and content supply to assess how platform incentives propagate through the market.

9.2 Polarization and Welfare

The counterfactual results reveal a central trade-off in platform design. Reducing algorithmic amplification lowers user engagement and short-run utility, but simultaneously leads to a more balanced distribution of information exposure.

On the one hand, the algorithm amplifies components of user utility that are positively valued—namely ideological alignment and the salience of unexpected content. As a result, removing these amplification channels mechanically reduces expected utility in the short run. This effect is reflected in the utility loss estimates reported in Section 8.4.

On the other hand, weakening algorithmic amplification substantially changes the composition of viewed content. As shown in Figure 10, users are exposed to a more diverse and ideologically balanced set of news when amplification is reduced. This suggests that the algorithm not only increases engagement but also concentrates attention on ideologically extreme content.

These findings highlight a tension between engagement maximization and exposure diversity. While engagement-based metrics favor content that is aligned with user preferences or highly salient, such content may reduce exposure to opposing viewpoints and contribute to polarization. A large literature is based on the idea that more diverse media environments can mitigate bias and improve information aggregation (e.g., Gentzkow et al., 2014, Bakshy et al., 2015, Levy, 2021).

Together, the results imply that policies aimed at reducing polarization through algorithmic design are likely to involve short-run costs in terms of engagement and user utility. However, these costs may be offset by long-run gains from a more balanced and informative media environment. This trade-off is central to evaluating the welfare consequences of platform regulation.

9.3 Broader Implications

More broadly, the findings highlight a general principle for digital markets: when intermediaries control visibility, their objective functions shape equilibrium outcomes in upstream markets. In such environments, algorithm design becomes a central policy lever, analogous to pricing or entry regulation in traditional industrial organization settings.

This insight extends beyond news to other contexts in which platforms mediate information or product discovery. By shaping the mapping from content to exposure, platforms influence not only what users consume, but also what producers choose to supply. As a result, understanding and regulating algorithmic design is essential for addressing distortions in modern information markets.

10 Conclusion

This paper documents and quantifies how social-media algorithms shape the supply of news. Using newly linked data that match print and online newspaper headlines to millions of

Facebook posts and user engagement metrics from 2017–2025, I show that the design of platform algorithms not only determines what users see, but also feeds back to influence what news outlets produce.

The regression and event-study evidence around Facebook's 2025 restoration of political content provides the main causal evidence in the paper. Following the policy change, political posts receive substantially more exposure, and outlets increase the extent to which their online headlines depart from the print headlines of the same stories. The effect is robust across alternative political-content classifiers and is not accompanied by a corresponding increase in engagement per view. These patterns are consistent with outlets adjusting their online editorial choices in response to changes in platform incentives rather than simply reflecting a contemporaneous increase in users' propensity to engage with political content.

Complementary evidence from Facebook's earlier reduction in political-content distribution points in the opposite direction: as political content was reduced, the online–print slant gap declined. Because that earlier policy was rolled out gradually, I do not interpret it as the primary quasi-experiment, but the opposite responses across the two policy regimes provide additional evidence that recommendation incentives shape the ideological content supplied online.

Building on this evidence, I develop a dynamic structural model in which users, algorithms, and media outlets interact in a digital news market. Users engage with content based on ideological alignment and salience; algorithms recommend posts that maximize predicted engagement; and media outlets dynamically choose headline slant to attract visibility. Estimating the model with post-level data reveals that prior to Facebook's policy change, the algorithm introduces substantial additional incentives for politically slanted content relative to intrinsic user preferences. Outlets responded by increasing online slant to maximize expected views, consistent with their estimated objective parameters that heavily weight visibility but place little value on engagement metrics.

Counterfactual simulations using the estimated model quantify the magnitude and chan-

nels of algorithmic amplification. Eliminating ideological weighting from the recommendation system reduces the extra slant of online news by roughly one order of magnitude and substantially flattens the distribution of exposure across users with different ideologies, while generating only modest welfare losses. In contrast, reducing engagement-based targeting—similar to recent platform interventions—yields comparable welfare losses but leads to only limited improvements in exposure diversity. A fully neutral algorithm that removes both channels would eliminate nearly all of the observed divergence between print and online headlines, but at a large welfare cost. Because newsrooms primarily optimize for visibility, the platform's design shapes not only how content is distributed but also the equilibrium level of slant in online news supply.

Taken together, the findings imply that algorithmic curation creates a general-equilibrium feedback loop between user engagement, platform design, and editorial strategy. When engagement-based algorithms reward polarization, outlets respond by producing more polarized content; when platforms demote such content, supply moderates even without a shift in user tastes. These results provide the first structural quantification of how digital recommendation systems reshape the incentives of media producers, highlighting that algorithm design has consequences not only for the demand for information, but for its very supply. The approach also demonstrates how structural tools from industrial organization and information design can be applied to media economics to study dynamic platform markets where firms, algorithms, and consumers interact in real time.

Acknowledgments. This research did not receive any external funding. I have no financial relationships or conflicts of interest to disclose.

References

Aldous, K. K., Şengün, S., Salminen, J., Farooq, A., Jung, S.-G., and Jansen, B. J. (2025). Multiple engagement by an individual on a social media post is rare: Insight from an anal-

ysis of 3.5 million instagram user accounts and 29 user interviews. *Data and Information Management*, page 100113.

Alesina, A., Miano, A., and Stantcheva, S. (2020). The polarization of reality. *AEA Papers and Proceedings*, 110:324–28.

Allcott, H. and Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–36.

Armona, L., Gentzkow, M., Kamenica, E., and Shapiro, J. M. (2024). What is newsworthy? theory and evidence. NBER Working Paper 32512, National Bureau of Economic Research. Available at SSRN: https://ssrn.com/abstract=4851786.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Bakshy, E., Messing, S., and Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132.

Beattie, G., Durante, R., Knight, B., and Sen, A. (2021). Advertising spending and media bias: Evidence from news coverage of car safety recalls. *Management Science*, 67(2):698–719.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890.

Bhargava, H. K. (2022). The creator economy: Managing ecosystem supply, revenue sharing, and platform design. *Management Science*, 68(7):5233–5251.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2022). Salience. *Annual Review of Economics*, 14(Volume 14, 2022):521–544.

Braghieri, L., Eichmeyer, S., Levy, R., Mobius, M. M., Steinhardt, J., and Zhong, R. (2024). Article-level slant and polarization of news consumption on social media. Working Paper.

Cagé, J., Hervé, N., and Mazoyer, B. (2024). Social media influence on mainstream media: Evidence from two billion tweets. *SSRN Electronic Journal*.

Cagé, J., Hervé, N., and Viaud, M.-L. (2019). The production of information in an online world. *The Review of Economic Studies*, 87(5):2126–2164.

Calzada, J. and Gil, R. (2020). What do news aggregators do? evidence from google news in spain and germany. *Marketing Science*, 39(1):134–167.

Chiang, C.-F. and Knight, B. (2011). Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies*, 78(3):795–820.

Compiani, G., Morozov, I., and Seiler, S. (2026). Demand estimation with text and image data. *The RAND Journal of Economics*.

DellaVigna, S. and Kaplan, E. (2007). The fox news effect: Media bias and voting*. *The Quarterly Journal of Economics*, 122(3):1187–1234.

Djourelova, M. (2023). Persuasion through slanted language: Evidence from the media coverage of immigration. *American Economic Review*, 113(3):800–835.

Gentzkow, M. (2006). Television and voter turnout*. *The Quarterly Journal of Economics*, 121(3):931–972.

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

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? evidence from u.s. daily newspapers. *Econometrica*, 78(1):35–71.

Gentzkow, M., Shapiro, J. M., and Sinkinson, M. (2014). Competition and ideological diversity: Historical evidence from us newspapers. *The American Economic Review*, 104(10):3073–3114.

Gentzkow, M., Shapiro, J. M., and Taddy, M. (2019). Measuring group differences in high-dimensional choices: Method and application to congressional speech. *Econometrica*, 87(4):1307–1340.

Guess, A. M. (2021). (almost) everything in moderation: New evidence on americans' online media diets. *American Journal of Political Science*, 65(4):1007–1022.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

Hotz, V. J., Miller, R. A., Sanders, S., and Smith, J. (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, 61(2):265–289.

Huang, Y., Meng, J., and Weng, X. (2026). How awareness of media slant affects reaction to news: Field and experimental evidence from microblog users. *Management Science*.

Leung, T. C. and Strumpf, K. S. (2023). All the headlines that are fit to change. Working Paper.

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

Martin, G. J. and Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, 107(9):2565–99.

Moehring, A. (2024). Personalization, engagement, and content quality on social media: An evaluation of reddit's news feed. Working Paper.

Mullainathan, S. and Shleifer, A. (2005). The market for news. *American Economic Review*, 95(4):1031–1053.

Nyhan, B., Settle, J., Thorson, E., Wojcieszak, M., Barberá, P., Chen, A. Y., Allcott, H., Brown, T., Crespo-Tenorio, A., Dimmery, D., Freelon, D., Gentzkow, M., González-Bailón, S., Guess, A. M., Kennedy, E., Kim, Y. M., Lazer, D., Malhotra, N., Moehler, D., Pan, J., Thomas, D. R., Tromble, R., Rivera, C. V., Wilkins, A., Xiong, B., de Jonge, C. K., Franco, A., Mason, W., Stroud, N. J., and Tucker, J. A. (2023). Like-minded sources on facebook are prevalent but not polarizing. *Nature*, 620(7972):137–144.

Pakes, A. and McGuire, P. (2001). Stochastic algorithms, symmetric markov perfect equilibrium, and the 'curse' of dimensionality. *Econometrica*, 69(5):1261–1281.

Piccardi, T., Saveski, M., Jia, C., Hancock, J., Tsai, J. L., and Bernstein, M. S. (2025). Reranking partisan animosity in algorithmic social media feeds alters affective polarization. *Science*, 390(6776):eadu5584.

Qian, K. and Jain, S. (2024). Digital content creation: An analysis of the impact of recommendation systems. *Management Science*, 70(12):8668–8684.

Saeedi, M. (2019). Reputation and adverse selection: theory and evidence from ebay. *The RAND Journal of Economics*, 50(4):822–853.

Saeedi, M., Shen, Y., and Shourideh, A. (2024). Getting the agent to wait. Carnegie Mellon University Working Paper.

Sen, A. and Yildirim, P. (2016). Clicks bias in editorial decisions: How does popularity shape online news coverage? Available at SSRN 2619440.

Strömberg, D. (2004). Radio's impact on public spending*. *The Quarterly Journal of Economics*, 119(1):189–221.

Sundar, S. S., Snyder, E. C., Liao, M., Yin, J., Wang, J., and Chi, G. (2025). Sharing without clicking on news in social media. *Nature Human Behaviour*, 9(1):156–168.

Weintraub, G. Y., Benkard, C. L., and Van Roy, B. (2008). Markov perfect industry dynamics with many firms. *Econometrica*, 76(6):1375–1411.

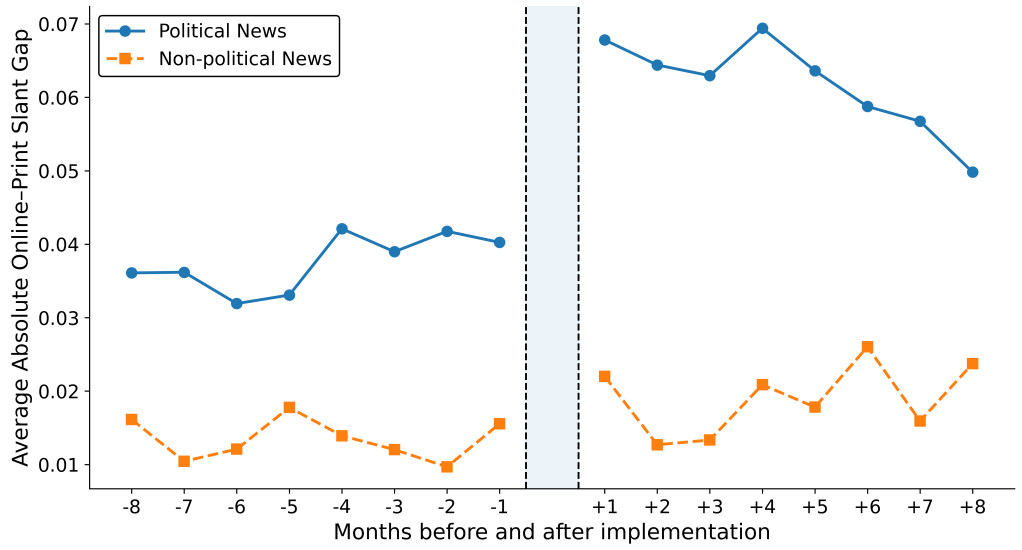
Wojcieszak, M., de Leeuw, S., Menchen-Trevino, E., Lee, S., Huang-Isherwood, K. M., and Weeks, B. (2023). No polarization from partisan news: Over-time evidence from trace data. *The International Journal of Press/Politics*, 28(3).

Zeng, Z., Zhang, Z., Chan, T., and Zhang, D. (2025). The impact of recommender systems on content consumption and production: Evidence from field experiments and structural modeling. Working Paper, SSRN 4915562.

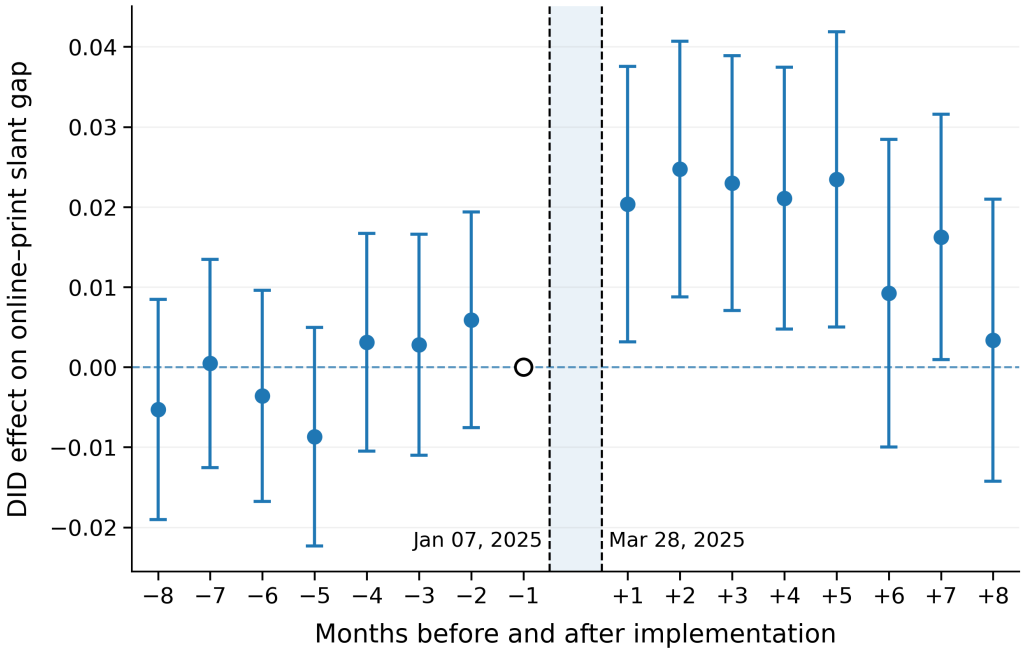
Table 7: Summary Statistics

Variable	Mean	SD
Panel A. All Facebook Posts		
Online slant	-0.0241	0.2091
Online sentiment	-0.1533	0.4127
Political content (FeedKit)	0.6751	0.4683
Political content (EXPO)	0.7231	0.4475
Views	179,861.3	520,276.7
Comments	361.4	1,294.5
Shares	258.6	1,563.4
Engagement	620.0	2,420.2
Engagement / views	0.0043	0.0106
Observations	3,117,489	
Media outlets	38	
Panel B. Posts Matched to Print Headlines		
Online slant	-0.0240	0.1906
Print slant	-0.0181	0.2177
Absolute online–print slant gap	0.0419	0.2392
Online sentiment	-0.1323	0.4032
Print sentiment	-0.1140	0.4114
Absolute online–print sentiment gap	0.1990	0.3036
Political content (FeedKit)	0.6456	0.4783
Political content (EXPO)	0.6615	0.4732
Views	133,423.6	644,057.0
Engagement / views	0.0034	0.0085
Observations	380,360	
Media outlets	11	

Notes: Panel A reports summary statistics for all Facebook posts in the sample. Panel B restricts the sample to posts matched to a corresponding print headline. Slant is measured using the paper's partisanship measure. The absolute online–print slant gap is the absolute difference between the online and print headline slant measures. Engagement is defined as comments plus shares. Political-content indicators are reported using the FeedKit and EXPO classifiers. The full sample spans April 2017 through December 2025.



(a) Average online–print slant gap



(b) Dynamic DID estimates

Figure 11: Robustness: Political Content Classification Using EXPO

Notes: The figure replicates Figure 2 using the EXPO political-content classifier. Panel (a) reports average absolute online–print slant gaps for political and non-political news. Panel (b) reports event-study DID estimates with outlet and date fixed effects, normalized to the last pre-implementation bin. Observations are grouped into 31-day intervals. The shaded region marks January 7 to March 28, 2025, which is excluded from estimation. Error bars indicate 95% confidence intervals.

Table 8: Placebo Tests for the Political-Content Difference-in-Differences

	Post × Political		
Placebo Period	FeedKit	EXPO	Observations
2024: Same Calendar Window	0.0008	-0.0026	37,978
	(0.0031)	(0.0047)	
2019: May 7–July 28	-0.0018	-0.0007	38,554
	(0.0046)	(0.0038)	
2023: July 7–September 28	-0.0010	-0.0005	38,585
	(0.0052)	(0.0050)	
Actual 2025 Policy Change	0.0205***	0.0218***	37,742
	(0.0028)	(0.0037)	

Notes: The table reports placebo estimates from the same difference-in-differences specification used for the 2025 political-content analysis. The dependent variable is the absolute online–print partisanship gap. Each entry reports the coefficient on *Post* × *Political*, where political content is classified separately using FeedKit and EXPO. The 2024 placebo uses the same calendar window as the 2025 policy change, shifted one year earlier. The other placebo windows are May 7–July 28, 2019 and July 7–September 28, 2023. All specifications include media-outlet and date fixed effects. Standard errors, reported in parentheses, are clustered by media outlet. The final row reports the updated actual 2025 policy estimate for comparison. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

A Newsroom Changes

Wall Street Journal. In August 2019, the *WSJ* built a *Newsroom Innovation* group with a dedicated SEO team and hired an SEO editor (Edward Hyatt), embedding search strategy alongside newsletters and product experimentation.[15]

Washington Post. The Post institutionalized platform work via an "Audience" team and later social-video initiatives (e.g., its well-known TikTok presence launched in 2019 under Dave Jorgenson). Public postings in the mid-to-late 2010s and early 2020s include social media editors and SEO leadership roles (e.g., "Head of AI Discovery & SEO").[16]

[15] NiemanLab (Aug. 12, 2019), "How The Wall Street Journal is building an incubator into its newsroom with new departments and plenty of hires."

[16] Vanity Fair (Feb. 2023), "Legacy Media Wants a Piece of That Gen Z TikTok Mojo," which discusses The Washington Post's TikTok team. See also Mediabistro's listing for "Head of AI Discovery & SEO," and Nina Zafar's Washington Post staff profile for an example of a social-media-editor role.

Table 9: Online–Print Slant Divergence Around the 2022 Policy Announcement

	Partisanship Gap	
	FeedKit	EXPO
Political	0.0245***	0.0235***
	(0.0038)	(0.0040)
After × *Political*	-0.0036**	-0.0033
	(0.0018)	(0.0023)
Media fixed effects	Yes	Yes
Date fixed effects	Yes	Yes
Observations	40,851	40,851
Media clusters	11	11
R^2	0.0261	0.0254

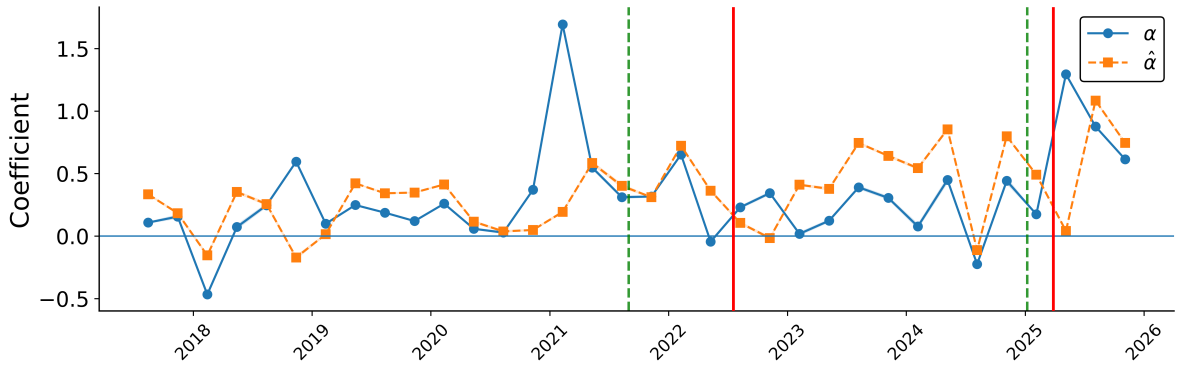
Notes: The table reports DID estimates around Facebook's July 19, 2022 announcement that changes reducing political content had been implemented globally. The dependent variable is the absolute online–print partisanship gap. Political content is classified using FeedKit or EXPO. All specifications include media-outlet and date fixed effects. Standard errors in parentheses are clustered by media outlet. Because the policy was gradually implemented before the announcement, these estimates capture differential changes around the announcement rather than the causal effect of the full rollout. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 10: Cumulative Variance Explained by Text Embedding Principal Components

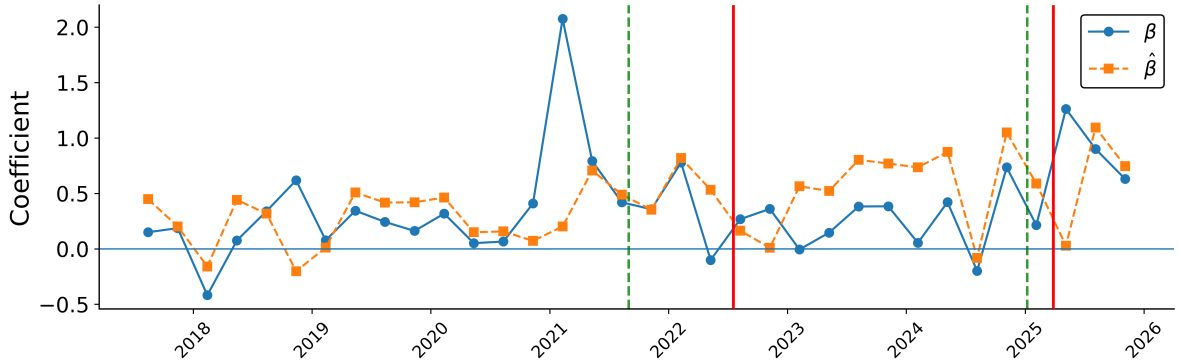
Number of PCs	Cumulative Variance Explained	Additional Variance from Previous Cutoff
10	21.5%	—
25	36.2%	14.6 pp
50	50.9%	14.7 pp
75	60.9%	9.9 pp
100	68.1%	7.2 pp

Notes: The table reports the cumulative share of variation explained by principal components of pretrained text embeddings constructed from the Facebook post `text` field. Post text is embedded using the frozen pretrained model `BAAI/bge-small-en-v1.5`, which produces a 384-dimensional embedding for each post. Principal components are estimated once using the pooled sample of Facebook posts across outlets. The first 10, 25, 50, 75, and 100 components explain 21.5%, 36.2%, 50.9%, 60.9%, and 68.1% of total embedding variation, respectively. The baseline specification uses the first 100 components. Specifications with fewer components are used to assess sensitivity to the dimensionality of the embedding controls. In the structural estimation, these components are residualized with respect to the economically interpretable content controls before entering as additional nuisance controls.

USA Today (Gannett). Gannett/USA Today Network has long maintained centralized audience/SEO and social teams across properties, reflected in frequent postings for audience,



(a) Preferences for unexpected slant: α_t and $\widehat{\alpha}_t$



(b) Preferences for ideological alignment: β_t and $\widehat{\beta}_t$

Figure 12: Estimated User and Algorithm Preferences over News Slant

Notes: The figure reports period-specific estimates of user preferences and additional algorithmic ranking preferences over news slant. Panel (a) shows the coefficients on squared deviations from an outlet's expected slant, capturing preferences for unexpected slant. Panel (b) shows the coefficients on squared ideological distance between content and users, capturing preferences for ideological alignment. Solid blue lines represent user preferences, while dashed orange lines represent additional algorithmic preferences. Green dashed vertical lines indicate August 31, 2021 and January 7, 2025; red solid lines indicate July 19, 2022 and March 28, 2025, marking major milestones in Facebook's political-content recommendation policies.

trending, and platform roles (mid-2010s onward). These roles focus on search optimization, real-time packaging, and cross-network distribution.[17]

Los Angeles Times. The *LA Times* scaled short-form and platform teams in the early 2020s (e.g., explicit TikTok/Instagram focus noted in 2023 coverage of legacy outlets' youth strategies), complementing multiplatform editor roles responsible for headline testing, meta-

[17] See Gannett's careers portal. Representative network postings emphasize analytics-driven packaging and social programming.

data, and social copy.[18]

Chicago Tribune & Boston Globe. By the mid-2010s, both organizations maintained digital desks with multiplatform editors and social roles; public job histories and postings indicate ongoing platform work (web producing, social packaging, SEO best practices).[19]

B Description of Political-Content Labels

To identify political content independently of ideological slant, I employ two pretrained, task-specific classifiers. My preferred measure uses the FeedKit RoBERTa classifier from Piccardi et al. (2025), developed to identify political content in social-media posts. As a robustness check, I use the EXPO BERT classifier from Wojcieszak et al. (2023), trained on manually annotated news headlines. The two classifiers provide independent measures of political content and allow me to assess the robustness of the results to alternative classification methods.

The two measures differ primarily in the breadth of content classified as political. FeedKit classifies 67.6% of posts as political, compared with 72.3% for EXPO. Despite the differences in prevalence, the two task-specific classifiers agree closely: EXPO and FeedKit assign the same label to 89.4% of observations.

C Contraction of the Share Inversion Mapping

I examine the convergence of the two-layer share inversion used in the demand estimation. Suppress the outlet and time subscripts, and fix the nonlinear parameters $(\beta, \widehat{\beta}, \zeta)$ and the distribution of user ideology F.

[18] See Vanity Fair (Feb. 2023), "Legacy Media Wants a Piece of That Gen Z TikTok Mojo." Los Angeles Times newsroom announcements and job postings also reference multiplatform and audience-development roles.

[19] See Boston Globe Media's careers portal for examples of audience and multiplatform roles. Public staff profiles and CVs also document digital-editor positions at Tribune newspapers.

Table 11: Alternative Measures of Political Content

Measure	Description	Reference	Political Share
FeedKit	RoBERTa-based political-content classifier developed for social-media text and used in the algorithmic-feed experiment of Piccardi et al. (2025). I use the released classifier with the recommended 0.5 threshold and apply it to the combined Facebook post text and headline information.	Piccardi et al. (2025)	0.676
EXPO	BERT-based binary classifier developed specifically to distinguish political from non-political news. The model is fine-tuned on manually coded news article titles. I apply the classifier to the same combined Facebook post text and headline information.	Wojcieszak et al. (2023)	0.723

Notes: The sample contains 3,117,489 Facebook posts from 27 news outlets and 11 influencer accounts. The first panel summarizes two alternative binary classifications of political content. FeedKit is the preferred measure in the main regression analysis, while EXPO provides an independently developed task-specific benchmark.

For post n and ideological group g, let e_{ng} denote the probability of engagement conditional on recommendation, and let r_{ng} denote the probability of recommendation:

$$e_{ng} = \frac{\exp(W_{ng})}{1 + \exp(W_{ng})}, \qquad r_{ng} = \frac{\exp(\widehat{W}_{ng})}{1 + \sum_m \exp(\widehat{W}_{mg})} \tag{33}$$

The predicted view and engagement shares are

$$\widehat{S}_n^V = \sum_g p_g r_{ng}, \qquad \widehat{S}_n^E = \sum_g p_g e_{ng} r_{ng} \tag{34}$$

Let $x = (\delta, \widehat{\delta})$ collect the post-specific mean utilities. The inversion mapping $T(x) =$

$(T^E(x), T^V(x))$ is defined by

$$T_n^E(x) = \delta_n + \log S_n^E - \log \widehat{S}_n^E(x) \tag{35}$$

$$T_n^V(x) = \widehat{\delta}_n + \log S_n^V - \log \widehat{S}_n^V(x) \tag{36}$$

A fixed point of this mapping matches both observed shares.

Unlike the standard single-layer logit inversion, the two updates are coupled. In particular, recommendation probabilities depend on user engagement utility through ζ. Differentiating the recommendation and engagement probabilities gives

$$\frac{\partial r_{ng}}{\partial \widehat{\delta}_m} = r_{ng}\left(\mathbf{1}\{n = m\} - r_{mg}\right) \tag{37}$$

$$\frac{\partial r_{ng}}{\partial \delta_m} = \zeta e_{mg} r_{ng}\left(\mathbf{1}\{n = m\} - r_{mg}\right) \tag{38}$$

$$\frac{\partial e_{ng}}{\partial \delta_m} = \mathbf{1}\{n = m\} e_{ng}(1 - e_{ng}) \tag{39}$$

Consequently, the Jacobian of the joint inversion mapping is

$$J_T(x) = I_{2N} - \begin{pmatrix} \dfrac{\partial \log \widehat{S}^E}{\partial \delta} & \dfrac{\partial \log \widehat{S}^E}{\partial \widehat{\delta}} \\ \dfrac{\partial \log \widehat{S}^V}{\partial \delta} & \dfrac{\partial \log \widehat{S}^V}{\partial \widehat{\delta}} \end{pmatrix} \tag{40}$$

The off-diagonal blocks capture the interaction between the two demand layers.

To establish a sufficient condition for contraction, consider a closed, convex region $\mathcal{D}$ that is mapped into itself by T. Suppose that throughout this region, the user-type-specific engagement probabilities and total news recommendation probabilities satisfy

$$e_{ng} \leq \bar{e} < 1, \qquad \sum_n r_{ng} \leq \bar{r} < 1 \quad \text{for all } n, g \tag{41}$$

Write the Jacobian in block form as $J_T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$. Using the derivatives above, the

maximum absolute row-sum norms of the four blocks satisfy

$$\begin{aligned}
\|A\|_\infty &\le \bar{e}[1 + |\zeta|(1 + \bar{r})] \\
\|B\|_\infty &\le 1 + \bar{r} \\
\|C\|_\infty &\le |\zeta|\bar{e}(1 + \bar{r}) \\
\|D\|_\infty &\le \bar{r}
\end{aligned} \tag{42}$$

Consider the weighted maximum norm

$$\|(\delta, \widehat{\delta})\|_\omega = \max\left\{\|\delta\|_\infty, \omega\|\widehat{\delta}\|_\infty\right\}, \qquad \omega > 0$$

The corresponding Jacobian norm is bounded by

$$q_\omega = \max\left\{\bar{e}[1 + |\zeta|(1 + \bar{r})] + \frac{1 + \bar{r}}{\omega}, \bar{r} + \omega|\zeta|\bar{e}(1 + \bar{r})\right\} \tag{43}$$

If there exists $\omega > 0$ such that $q_\omega < 1$, the mapping is a contraction on $\mathcal{D}$. By the mean-value inequality,

$$\|T(x) - T(y)\|_\omega \le q_\omega\|x - y\|_\omega \quad \text{for all } x, y \in \mathcal{D}$$

Since $\mathcal{D}$ is closed and invariant under T, the Banach fixed-point theorem guarantees a unique fixed point in $\mathcal{D}$ and convergence of the iterative updates from any initialization within this region.

The sufficient condition reflects the two-layer structure of the model. A large outside option limits interactions across recommendation shares, while sufficiently small conditional engagement probabilities and a bounded algorithmic weight ζ limit feedback between recommendation and engagement utility.