9^{th} International Conference on Computational Social Science IC^{2}S^2
July 17-20, 2023, Copenhagen, Denmark

Metrics at Work: How Social Media Metrics Shape News Production on Facebook

Keywords: social media, time series, topic modeling, panel data analysis, Facebook

Extended Abstract

Social media platforms have become a major source of news and information for American citizens [4] where journalists monitor public opinion and popular trends, and newsrooms try to attract audiences and traffic [7]. These platforms provide media practitioners with a variety of engagement metrics, such as the number of comments, reactions, and shares, which may signal what the audience demands and can in turn be used by newsrooms to decide what content to publish [1, 3]. While this can give citizens more voice in determining the news agenda, it may also lead to a focus on attention-grabbing "soft news" rather than important topics of public interest [5]. Additionally, media outlets may choose to cover more divisive, toxic, or hyper-partisan topics to boost engagement metrics and attract a larger audience [2]. Understanding how social media metrics influence news content production is crucial for promoting a healthy information environment, which is considered essential for a functioning democracy.

In this study, we seek to quantify how audience engagement metrics shape news organizations' content production on Facebook, a platform that two-thirds of U.S. adults report using and one-third report regularly getting news from [4]. We leverage a dataset of over 2.2 million Facebook posts published by 29 outlets over a five-year period. This set, which includes all English language outlets from Pew's 57 $^{\text{th}}$ wave of the American Trends Panel [6] comprises a relatively comprehensive list of ideologically diverse media that are among the most prominent in the U.S. It allows us to conduct a holistic appraisal of the dynamic between engagement metrics and news production, as well as investigate between-outlet heterogeneities.

We use this large-scale longitudinal dataset to design a measure to quantify media responsiveness to audience engagement signals by combining topic modeling and longitudinal data analysis techniques. Specifically, we test two main hypotheses, and ask one research question:

H1: Engagement metrics on social media affect news content production. News topics that attract greater engagement are increasingly covered in subsequent time periods.

H2: Responsiveness of outlets to topics will be associated with their ideological slants: right-wing outlets are more likely to respond to audience metrics than left-wing outlets.

RQ: How does the responsiveness of outlets to audience engagement vary between different domains, especially for partisan politics and entertainment?

We first identify 1,005 topics over 60, one-month seasons that comprise our five-year time frame. We then run fixed-effects models with a lagged dependent variable on aggregated data that are pooled within each one-month season. We define *responsiveness* as the effect β_1 of $ES_{i,m,t-1}$ (audience engagement signal with topic i for outlet m at time $t-1$) on $Freq_{i,m,t}$ (the subsequent frequency of the topic i being published by outlet m at time t) after controlling for news value ($Freq_{i,m,t-1}$, $Freq_{i,t-1}$) and topic level (ζ_i), outlet level (ξ_m), and time level (λ_t) fixed effects as shown in Equation 1.

$$\begin{aligned}\text{Freq}_{i,m,t} =& \beta_0 + \beta_1 \log(ES_{i,m,t-1}) + \beta_2 \text{Freq}_{i,m,t-1} + \beta_3 \text{Freq}_{i,t-1} + \beta_4 \log(ES_{i,t-1}) + \\ & \zeta_i + \lambda_t + \xi_m + \varepsilon\end{aligned} \tag{1}$$

Our findings suggest that engagement metrics have a significant but modest effect (Figure 1): topics that receive higher (lower) than average engagement on any day on Facebook, are covered significantly more (less) by that outlet on the very next day, but as panel A shows, this effect dissipates and becomes non-significant beyond a response window of four days. There is also substantial heterogeneity in the extent to which the outlets respond (panel B). In fact, the responsiveness of the outlets is highly correlated with their audience-based ideological slants obtained from Pew [6]: right-wing outlets are significantly more responsive to audience engagement signals than left-leaning outlets (panel C).

Next, we code each of the 1,005 topics, classifying them as partisan politics, entertainment-oriented, or other. We then repeat our analysis for partisan and entertainment-oriented topics separately and find that the correlation between outlet-slant and outlet-responsiveness is contingent on the type of topic: only significant for partisan topics (panel D), and not for entertainment-oriented topics (panel E). In other words, right-wing outlets are significantly more likely than left-wing outlets to respond to audience engagement, but only in the domain of partisan politics.

Our results are robust to a variety of model specifications and provide valuable insights into how social media metrics empirically affect journalistic production on Facebook and its implications for our news environment. We also demonstrate the application of a novel computational approach to capture the impact that the audiences exert on content production, which in turn could be extended to other domains and platforms. While our findings can be interpreted with optimism to indicate that these metrics are democratizing information flows, enabling audiences to collectively affect the news that journalists produce, there are reasons to be skeptical and exercise caution. Journalism's transition towards a majoritarian Market Model where news supply is determined by the readers may further threaten the health of the information environment, and consequently, undermine the democratic process where all voices are heard, represented, and paid attention to.

References

[1] Angèle Christin. *Metrics at work: journalism and the contested meaning of algorithms*. Princeton University Press, 2020.

[2] Molly J Crockett. Moral outrage in the digital age. *Nature human behaviour*, 1(11):769–771, 2017.

[3] Eun-Ju Lee and Edson C Tandoc Jr. When news meets the audience: How audience feedback online affects news production and consumption. *Human communication research*, 43(4):436–449, 2017.

[4] Katerina Eva Matsa and M Walker. News consumption across social media in 2021. *Pew Research*, 2021.

[5] Lukas Otto, Isabella Glogger, and Mark Boukes. The softening of journalistic political communication: A comprehensive framework model of sensationalism, soft news, infotainment, and tabloidization. *Communication Theory*, 27(2):136–155, 2017.

[6] Pew. American Trends Panel Wave 57 [Data set]. Technical report, Pew Research Center.

[7] Edson C Tandoc Jr and Tim P Vos. The journalist is marketing the news: Social media in the gatekeeping process. *Journalism practice*, 10(8):950–966, 2016.

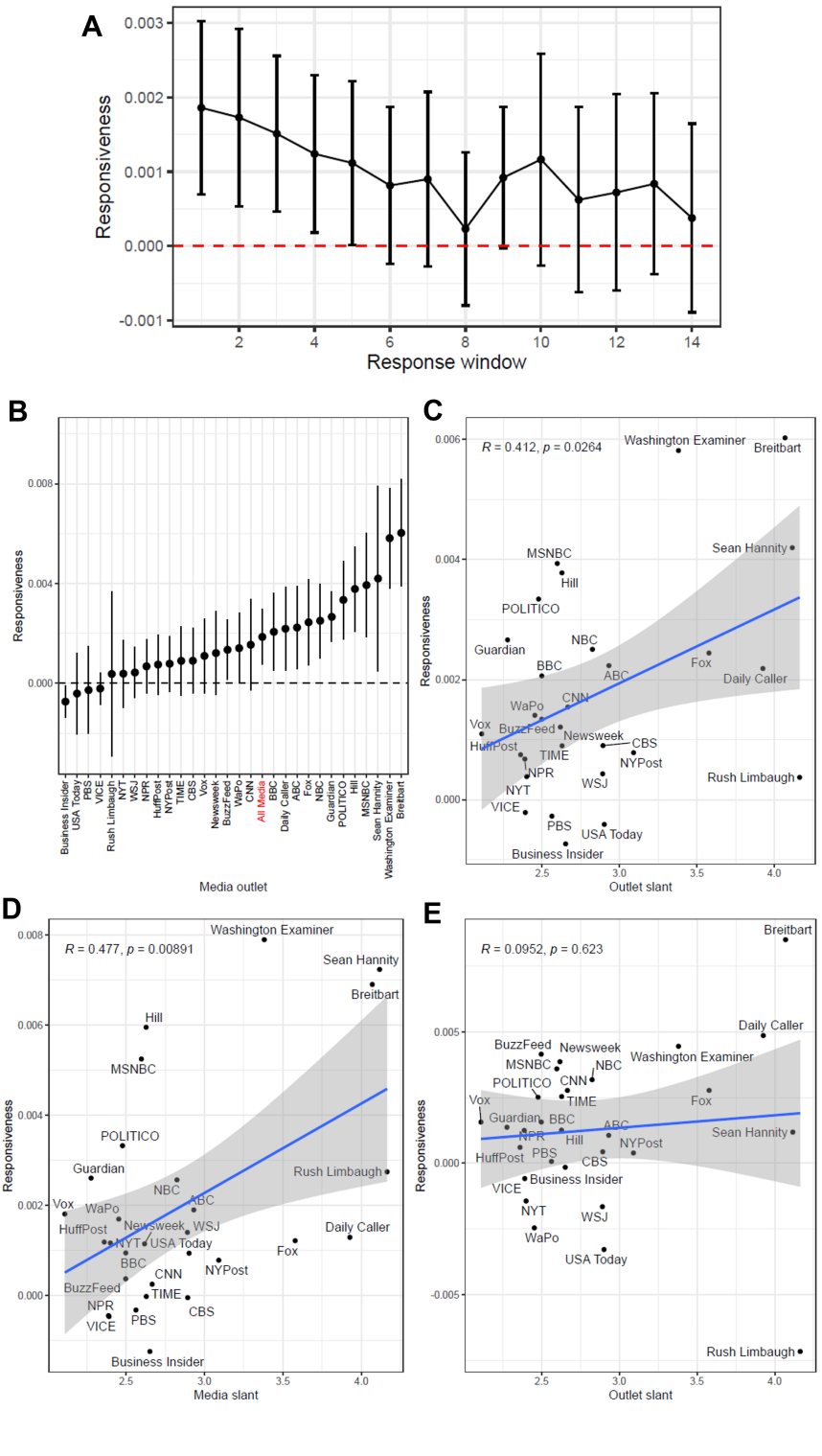

Figure 1: Panel A depicts the point estimates of responsiveness scores (and 95% CI) for response windows ranging from 1 day to 14 days. After 4 days, responsiveness scores are no longer significant (panel A). The outlets vary greatly in responsiveness scores (panel B), and their responsiveness is highly correlated with their partisan slants: right-wing outlets being significantly more responsive to audience engagement than left-wing outlets (panel C). This correlation is however, contingent on the type of topics: right-wing outlets are significantly more responsive than left-wing outlets but only to audience engagement with partisan topics (panel D) and not with entertainment-oriented topics (panel E).