

Whose Advantage? Measuring Attention Dynamics across YouTube and Twitter on Controversial Topics

JooYoung Lee,^{*1} Siqi Wu,^{*3,1} Ali Mert Ertugrul,^{*2} Yu-Ru Lin,² Lexing Xie¹

¹ School of Computing, Australian National University

² School of Computing and Information, University of Pittsburgh

³ School of Information, University of Michigan

{jooyoung.lee,lexing.xie}@anu.edu.au, {ertugrul,yurulin}@pitt.edu, siqiwu@umich.edu

Abstract

The ideological asymmetries have been recently observed in contested online spaces, where conservative voices seem to be relatively more pronounced even though liberals are known to have the population advantage on digital platforms. Most prior research, however, focused on either one single platform or one single political topic. Whether an ideological group garners more attention across platforms and topics, and how the attention dynamics evolve over time, have not been explored. In this work, we present a quantitative study that links collective attention across two social media platforms – YouTube and Twitter, centered on online activities surrounding videos of three controversial political topics including Abortion, Gun control, and Black Lives Matter over 16 months. We propose several video-centric metrics to characterize how online attention is accumulated for different ideological groups. We find that neither side is on a winning streak: left-leaning videos are overall more viewed, more engaging, but less tweeted than right-leaning videos. The attention time series unfold quicker for left-leaning videos, but span a longer time for right-leaning videos. Network analysis on the early adopters and tweet cascades show that the information diffusion for left-leaning videos tends to involve centralized actors; while that for right-leaning videos starts earlier in the attention lifecycle. In sum, our findings go beyond the static picture of ideological asymmetries in digital spaces and provide a set of methods to quantify attention dynamics across different social platforms.

1 Introduction

Several recent studies have documented the ideological asymmetries between the left-wing and right-wing activism (Brady et al. 2019; Schradie 2019; Freelon, Marwick, and Kreiss 2020; Waller and Anderson 2021). Some highlight the dominance of conservative voices on social media (Brady et al. 2019); others portray the widespread symbolic support for progressive social movements (Jackson, Bailey, and Welles 2020). The term “conservative advantage” is coined to describe the strategic dissemination of right-wing users to spread their messages (Schradie 2019). However, most of the existing research bases on the analysis

of a single platform or a single political topic. Relatively little is known about how different ideological groups garner attention across platforms, and whether the group advantage of gaining visibility remains across topics and over time. To answer these questions, this work designs several sets of cross-platform measurements on the collective attention dynamics of two different ideological groups across three controversial political topics.

Online platforms, such as Twitter, YouTube, Reddit, and Facebook, are social-technological artifacts that segregate online attention into silos defined by the underlying software and hardware systems. Video views on YouTube are known to be driven by discussions outside the platform (Rizou et al. 2017), and to be part of users’ broader information diet (HosseiniMardi et al. 2021). What is not known, however, is how groups of related content comparatively evolve across different social platforms. Collective attention on political content have been studied on one topic, such as the Occupy Movement (Thorson et al. 2013), Gun Control/Rights (Zhang et al. 2019), and Black Lives Matters (De Choudhury et al. 2016; Stewart et al. 2017). Yet, cross-cutting studies that compare different movements are rare. With data from three long-running controversial topics, this work seeks to provide measures across YouTube and Twitter and paint a nuanced picture about the temporal patterns of attention from left to right.

We choose three topics: Abortion, Gun Control, and Black Lives Matter (BLM). We rely on video hyperlinks to connect the content from YouTube to Twitter. A motivating example is given in Figure 1. We plot the time series of daily view count for the collected BLM videos from YouTube (top panel) and daily volume of tweets mentioning these BLM videos from Twitter (bottom panel). Both time series are further disaggregated by video uploaders’ political leanings. Visually, the view count dynamics of both left- and right-leaning videos are relatively stable in year 2017, except a sharp spike caused by the “Unite the Right rally”¹ event in Charlottesville, USA. On the bottom panel, the tweet count dynamics of right-leaning videos have many spikes, which can be attributed to the upload of new videos from far-right YouTube political commentators. The measures on YouTube and Twitter present a contrasting story here: if we

^{*}These authors contributed equally.

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

¹https://en.wikipedia.org/wiki/Unite_the_Right_rally

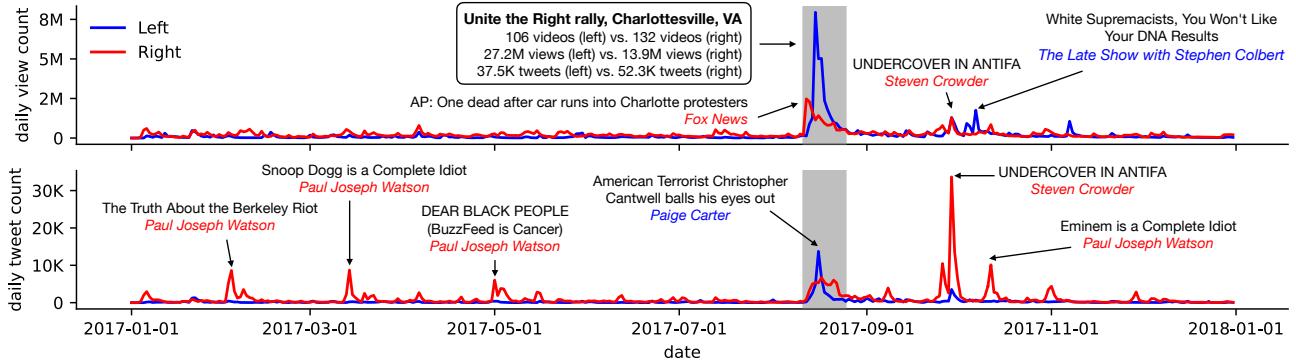


Figure 1: Attention time series (top: daily view count; bottom: daily tweet count) related to BLM throughout 2017. The view count dynamics of both left- and right-leaning videos have a handful of sharp peaks, while the tweet count dynamics of right-leaning videos peak more frequently. Notable video releases (with the most views and tweets) are labeled. YouTube channel titles are *italicized* and colored by video leanings. Best viewed in colors.

focus on the two weeks period after the rally, left-leaning videos attracted more attention on YouTube (measured by views, left: 27.2M, right: 13.9M) while right-leaning videos had higher exposure on Twitter (measured by tweets, left: 37.5K, right: 52.3K). This example demonstrates the need for cross-platform analysis – findings on one platform may not generalize to another.

We design a set of metrics from publicly available data on YouTube and Twitter, which include total views, video watch engagement, tweet reactions, the evolution of attention over time, and early adopter networks among tweets and Twitter users. On YouTube, we find that left-leaning videos accumulate more views, are more engaging, and have higher viral potential than right-leaning videos. In contrast, right-leaning videos have higher numbers of total tweets and retweets on Twitter. Statistics on the unfolding speed for views and tweets show that the attention on left-leaning videos attenuates faster, while that on right-leaning videos persists for longer. Note that these observations are not generalized unanimously across topics, e.g., for some metrics, we observe significant differences for Abortion and Gun Control, but not for BLM. These findings expand current wisdom on ideological asymmetries in two ways: the first is exposing the novel facet that left-leaning content attracts more attention in a shorter period of time; the second is the need of contrasting temporal attention statistics between platforms, such as right-leaning tweet cascades tend to start earlier and YouTube views on right-leaning content sustain longer. In sum, our observations paint a richer picture of attention patterns across the political spectrum, provide a basis for further studying political framing and group behavior, and supply fundamental metrics for understanding influences that transcend platforms.

The main contributions of this work include:

- a data curation procedure linking content on YouTube and Twitter for longitudinal topic monitoring.²

²Our datasets and analysis code are publicly available at <https://github.com/picslab/Measuring-Online-Information-Campaigns>

- several sets of cross-platform metrics that support statistical comparisons for different ideological groups, encompassing the volume and quality of attention, networks of tweets and users, as well as relative temporal evolution.
- adding the temporal and cross-platform dimensions to recent observations on ideological asymmetries. We find that polarized content engage users in distinct ways – more views, more engagement, and faster reactions for videos on the left; comparing to more tweets, more sustained attention for videos on the right.

2 Related Work

Online behavior of political groups. Measurement studies have quantified different aspects of users, contents, and their interactions under political polarization on social media. Conover et al. (2011) presented one of the first profiling studies of polarized political groups on Twitter. They studied the communication patterns between two opposing ideological groups and found that the retweet network was highly polarized but the mention network was more mixed. There have also been evidences that liberal and conservative groups attract online attention in different manners. Abisheva et al. (2014) focused on a set of influential Twitter users who promoted YouTube videos, and they found that conservatives tweeted more diverse topics than liberals and that conservatives shared new videos faster. Bakshy, Messing, and Adamic (2015) quantified the extent to which Facebook users were exposed to politically opposing contents, and they found that conservatives tended to seek out more cross-partisan content. Garimella et al. (2018) defined several consumption and production metrics and profiled key user behavior patterns. Ottoni et al. (2018) showed that conservatives used more specific language to discuss political topics and showed more negative emotions in the language. On YouTube, a recent study from Wu and Resnick (2021) found that left-leaning videos attracted more comments from conservatives than right-leaning videos from liberals. However, all of these works are conducted platform-wide, and are not specialized into particular topics or movements.

Online activism, also known as online social movement or digital political campaign, has also been an active topic in computational social science. For example, De Choudhury et al. (2016) presented one of the first studies on the BLM movement, measuring geographical differences in participation, and relationships to offline protests. Stewart et al. (2017) constructed a shared audience network of users who talked about BLM on Twitter and found the existence of superclusters among liberals and conservatives. Zhang and Counts (2015) performed policy decision prediction based on tweet texts analysis on same-sex marriage. In a follow-up work, Zhang and Counts (2016) discussed gender disparity by linking tweet texts to the state-level Abortion policy events. Ertugrul et al. (2019) examined the relation between offline protest events and their social and geographical contexts. Freelon, Marwick, and Kreiss (2020) explained different tactics of liberals and conservatives when approaching audience on social media and articulated the asymmetries of measured behavior between conservatives and liberals. However, all of these works focus on a single controversial topic. In contrast, we examine three political topics and assess consistency of findings across the topics.

Cross-platform measurement studies. One early attempt in linking Twitter and YouTube data is from Abisheva et al. (2014), in which the authors found that the features of early adopters on Twitter were predictive for the video view counts on YouTube. Rizoiu et al. (2017) proposed the Hawkes Intensity Process that linked the time series of tweets and views, which led to a metric called *viral potential* for measuring the expected number of views that a video would obtain if mentioned by an average tweet (Rizoiu and Xie 2017). Zannettou et al. (2017) measured the sharing of alternative and mainstream news articles on three different platforms – Twitter, Reddit, and 4chan. This seminal study characterized the role of fringe communities in spreading news. HosseiniMardi et al. (2021) used browsing histories to infer video watch behavior, and quantified the consumption and driver of extreme content with respect to users’ information diet. However, all of these works cover a breadth of content sources and categories, but none is focused around social movements or consistent political topics.

This presented work bridges the gap of cross-platform measurement studies on multiple social movements, aiming to provide richer understanding that balances the ideological asymmetry between the left and the right.

3 Curating Tweeted Video Datasets

We constructed three new cross-platform datasets by tracking videos on YouTube and posts on Twitter over three controversial topics: Abortion, Gun Control, and BLM. Those topics have been studied extensively by social and political scientists (Zhang and Counts 2016; De Choudhury et al. 2016; Stewart et al. 2017; Garimella et al. 2018). In this section, we first describe the data collection strategy. We then introduce our methods for estimating the political leanings of Twitter users and YouTube videos. Table 1 summarizes the overall statistics of the three topical datasets.

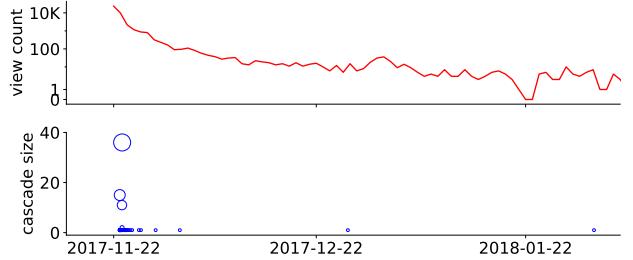


Figure 2: Time series of daily view count (top panel, in log scale) and tweet cascades (bottom panel) for an example video (YouTube video id: PTjMngQwGh8). A tweet cascade is placed at the time of its first tweet, with y-axis and size proportional to the number of tweets in the cascade (see Section 4.3). The viewing dynamic unfolds for months, but tweets only happen sporadically after the first week.

3.1 Finding YouTube Videos and Twitter Posts of Controversial Topics

We are interested in topical YouTube videos and the discussions about them on Twitter. Following the approach used in (Wu, Rizoiu, and Xie 2018), we collected public tweets that mentioned any YouTube URLs via the Twitter filtered streaming API. Our raw Twitter stream spanned 16 months (2017-01-01 to 2018-04-30), and contained more than 1.8 billion tweets. To subsample videos that attracted a reasonable amount of attention before the end of the observation period, we required that the videos must be published in 2017, receive at least 100 tweets and at least 100 views within the first 120 days after upload. We make this filtering choice because (a) analyzing videos with little attention (<1 view per day) is bound to generate noises when comparing different groups; (b) characterizing the timing and structure of a video’s tweeting cascades requires a non-trivial number of tweets. This yielded 328,557 videos, which were mentioned in 242M tweets by 29.9M users. For each video, we collected its metadata, daily time series of view count and watch time, and all tweets mentioning it.

To identify topic-relevant videos and tweets, we first curated three separate keyword lists for three controversial topics – Abortion, Gun Control, and BLM. We consider a video is potentially relevant if (a) it contains at least one keyword in the video title or description; or (b) it is mentioned in a tweet that contains at least one keyword in the tweet text. Next, we used a mix of manual and semi-automated approaches to annotate the potentially relevant videos. Section A of (Appendix 2022) details the topical keywords, their curation, and our video annotation protocol. In total, we obtained 179 Abortion, 268 Gun Control, and 777 BLM videos, which were mentioned in 970K+ tweets.

Figure 2 shows the daily view count series and tweet cascades for an example video. A notable contrast between the two sources is that YouTube attention data is only available in daily aggregates, i.e., without individual user logs, while tweets have precise timestamp, user information, and relations between tweets and users. Section 4 is dedicated to designing measures for such cross-platform multi-relational

	YouTube video				Tweet Total	User Total	20% Early Adopters (Twitter user)				(Tweet)	
	L	C	R	Total			Lib.	Neu.	Con.	Total	Total	
Abortion	58	10	111	179	106,776	76,337	2,534	1,708	8,202	12,444	15,843	
Gun Control	81	33	154	268	270,543	156,145	10,987	3,371	13,591	27,949	37,264	
BLM	297	84	396	777	593,574	262,580	10,419	5,745	33,934	50,098	78,969	

Table 1: Summary of cross-platform datasets on three controversial topics. L: Left-leaning video; C: Center video; R: Right-leaning video. Lib.: Liberal user; Neu.: Neutral user; Con.: Conservative user. See Section A of (Appendix 2022) for methods of curating tweeted video datasets across YouTube and Twitter, and Section B&C for methods of estimating the political leanings of Twitter users and YouTube videos. The ratios of early adopters are slightly below 20% due to banned and protected users.

temporal data. We bootstrapped the political information about Twitter users and YouTube videos. In particular, we gathered more information about videos’ early adopters, defined as the first 20% users who tweeted about each video. We collected the follower lists for all early adopters. Network measures of early adopters are found to indicate future popularity (Romero, Tan, and Ugander 2013). The threshold (first 20%) is chosen to balance the need for data and the burden of collecting network information within practical API limits. After filtering out banned and protected users, we extracted 132K early adopter tweets posted by tens of thousands of users across the three topics (see Table 1).

3.2 Estimating Political Leanings of Twitter Users and YouTube Videos

We classified the political leanings of early adopters on Twitter into liberal, neutral, and conservative. Meanwhile, we classified the video leanings into left, center, and right. Twitter users and YouTube videos are related in that left-leaning contents (e.g., Gun Control videos) are generally shared by liberal users while right-leaning contents (e.g., Gun Rights videos) are shared by conservative users.

We estimated Twitter users’ political leanings by first identifying a group of seed users who included political hashtags in their profile descriptions, and then by using a label propagation algorithm (Zhou et al. 2004) to propagate the labels of seed users to other users based on the shared follower network. This is a common approach for classifying user leaning on Twitter (Stewart et al. 2017) and the *follow* relation is found to be the most important in predicting user ideology (Xiao et al. 2020). We performed 10-fold cross-validation to evaluate the classification performance. We observed very high scores in precision, recall, and F-score across all three topics (> 95% in all metrics). Section B of (Appendix 2022) describes our classification and evaluation methods for Twitter users in more detail.

We estimated YouTube videos’ political leanings by first averaging the leaning scores of videos’ early adopters on Twitter. We then used an external YouTube media bias dataset (Ledwich and Zaitsev 2020) to label the video leanings and identified optimal classification thresholds. We were able to find 58/10/111 left, center, and right-leaning videos for Abortion (analogously, 81/33/154 for Gun Control and 297/84/396 for BLM). To validate our estimation, we performed one round of manual annotation for videos in Gun Control. We used stratified sampling to

sample 50 videos based on the video leaning scores. These videos were annotated independently by three authors. The Fleiss’ Kappa was 0.69, suggesting a moderate inter-rater agreement. Section C of (Appendix 2022) details our classification and evaluation methods for YouTube videos.

4 Measures for Cross-Platform Attention

This section designs several sets of metrics for the cross-platform data, in order to compare content across different political ideologies, and examine whether the differences are consistent across topics, across platforms, and over time.

4.1 Aggregate Attention on YouTube and Twitter

We present four metrics for the total video attention on YouTube.

Total view count sums up a video’s view count time series until day 120.

Relative engagement is a metric proposed in (Wu, Rizoiu, and Xie 2018) for quantifying the average video watching behavior. Specifically, for each video, we first compute *average watch percentage*, defined as the total watch time divided by total number of views (both at 120 days) and then normalized by the video length (in seconds). The relative engagement score is the percentile ranking of average watch percentage among videos of similar lengths. It is a normalized score between 0 and 1. A higher score means more engaging, e.g., a score of 0.8 suggests that this video is on average watched for longer time than 80% videos of similar length. Note that relative engagement is shown to be stable over time, hence there is no need to examine the temporal variations of watch time, as it would strongly correlate with view counts. In this work, relative engagement is computed based on a publicly available collection of 5.3M YouTube videos (Wu, Rizoiu, and Xie 2018), with details described in Section D of (Appendix 2022).

Fraction of likes measures the video reaction – provided by YouTube as the total counts of likes and dislikes and collected via the thumb-up and thumb-down icon on the video page. A relatively lower fraction of likes indicates a more diverse audience reaction to the video content. Note that the majority of videos receive a lot more likes than dislikes.³

Viral potential is a positive number, representing the *expected* number of views that a YouTube video will obtain

³YouTube announced that the dislike count will no longer be available to the public on Nov 10, 2021. <https://blog.youtube/news-and-events/update-to-youtube/>

if mentioned by an *average* tweet on Twitter (Rizou and Xie 2017). More specifically, it is the area under the impulse response function of an integral equation known as Hawkes Intensity Process (HIP) (Rizou et al. 2017), which is learned for each video by using the first 120 days of tweeting and viewing history. We choose this quantity rather than simply dividing the number of views by the number of tweets, because the model takes into account views that are yet to unfold due to its sustained circulation via sharing and tweeting. A self-contained summary about HIP and viral potential computation is given in Section E of (Appendix 2022).

On Twitter, tweets can be categorized into four types: original tweets, retweets, quotes, and replies. This leads to five counting metrics: the total number of *tweets*, *original tweets*, *retweets*, *quoted tweets*, and *replies*.

4.2 Views and Tweets over Time

Viewing half-life is computed as the number of days to achieve half of its total views at day 120.

Tweeting half-life is computed as the number of days to achieve half of its total tweets at day 120.

Tweeting lifetime is time gap between the first and the last tweets. We do not measure lifetime on viewing because the view count of a video rarely becomes zero even towards the end of the measurement period, but tweets tend to exhaust much sooner.

Tweeting inter-arrival time is the average time difference between every two consecutive tweets about each video.

Accumulation of views and tweets. In addition to the summary metrics above, we also compare the attention accumulation on the left- and right-leaning content on a daily basis. On each day t , we compute the fraction of the total views that each video has achieved. This leads to two sets of samples $\{v_i^{(L)}\}_{i=1}^n$ and $\{v_j^{(R)}\}_{j=1}^m$, where n is the number of left-leaning videos and m is the number of right-leaning videos. We then compute the normalized Mann-Whitney U (MWU) statistic (Mann and Whitney 1947),

$$\bar{U}_t = \frac{1}{nm} \sum_i \sum_j \{I[v_i^{(L)} > v_j^{(R)}] + 0.5 I[v_i^{(L)} = v_j^{(R)}]\}$$

Here $I[\cdot]$ is the indicator function that takes value 1 when the argument is true, 0 otherwise. The U statistic intuitively corresponds to the fraction of sample pairs $(v_i^{(L)}, v_j^{(R)})$ where the sample from left-leaning distribution is larger, accounting for ties. If the distributions of $v^{(L)}$ and $v^{(R)}$ are indistinguishable, then \bar{U} would be around 0.5. We compute the statistic \bar{U}_t on tweets in the same fashion, and both statistics are computed for each day. These two series of statistics allows us to quantify the differences between left- and right-leaning content, and compare the trends on the accumulation of views and tweets over time.

4.3 Videos' Tweet Cascades

Cascade size. We define that a *cascade* consists of a root tweet and all of its retweets, replies, and quotes. It is well-known that the vast majority of cascades in online diffusion networks are very small and only a very small fraction of

cascades would become very big (Goel, Watts, and Goldstein 2012). Based on the number of tweets in a cascade, we divide the cascades into isolated (only root tweet), small (2-4 tweets), and large (≥ 5 tweets) groups. For videos of each leaning on each topic, we compute the fractions of isolated/small/large cascades and the fraction of tweets in each cascade group. These metrics quantify the structure of online diffusion and allow us to compare behavior on controversial political topics with what was known about tweeted videos in general.

Cascade start time is the percentage of accumulated views of the video when the root tweet of the cascade is posted. It measures how much view attention is accumulated on YouTube before the infusion on Twitter starts. We choose to describe cascade timing relative to the accumulation of view, rather than in absolute number of days since upload, because (a) such relative time more directly correlates the amount of cascades with respect to the views they can potentially drive (rather than through another variable, days); and (b) the percentage of views provides more granularity, since many videos have all views and tweets unfold within a few days after upload.

4.4 Networks among Early Adopters on Twitter

For each video, we obtain its follower network among the early adopters. If there exists a following relationship between a pair of users, a directed edge is established. This results in one network for each shared videos. We compute a set of metrics per video, and then compare their distributions on each topic for left- and right-leaning videos. We describe two key metrics here, and discuss four additional metrics in Section H of (Appendix 2022).

Gini coefficient of indegree centrality. We calculate the indegree centrality for each node in the network. To have a video-level metric, we use the Gini coefficient, which ranges from 0 to 1 and measures the distribution inequality. Specifically, the Gini coefficient of indegree centrality quantifies the degree of inequality of the indegree distribution. A higher value indicates that a few early adopters are followed more by other early adopters, and a lower value indicates that the indegree distribution is more equal.

Gini coefficient of closeness centrality captures the dispersion in inverse of average shortest path length from one early adopter to all other early adopters of a given video. Higher coefficient implies that a few early adopters can reach the rest of the early adopters within a few hops.

5 Observations on Cross-Platform Attention

We report the results on all metrics described in Section 4, in mirroring subsections to aid navigation. Many results in this section are presented as violin plots. The outlines are kernel density estimates for the left-leaning (blue) and right-leaning (red) videos, respectively. The center dashed line is the median, whereas the two outer lines denote the interquartile range. To compare the distributions of each metric for the left- and right-leaning videos, we adopt the one-sided Mann-Whitney U test. We summarize our results in Table 2 at the end of this paper.

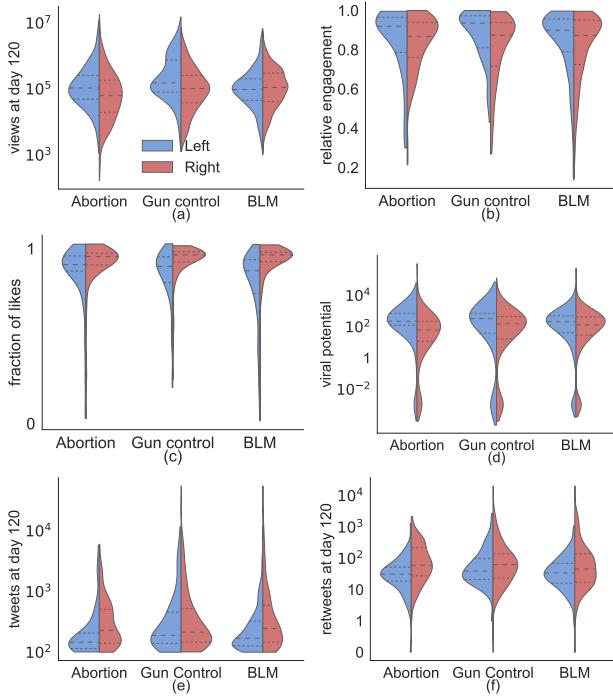


Figure 3: Violin plots comparing the left-leaning (blue) and right-leaning (red) videos in (a) total number of views at 120 days after upload, (b) relative engagement, (c) fraction of likes, (d) viral potential, (e) total number of tweets, and (f) total number of retweets at 120 days after the video upload. Left-leaning videos are more viewed, more engaging, having more diverse reactions, and more viral than right-leaning videos across all three topics (except views in BLM). In Abortion and BLM, right-leaning videos are significantly more tweeted, especially with more retweets.

5.1 Aggregate Attention on YouTube and Twitter

Total view count. Figure 3(a) shows the distribution of video views at day 120 after upload. Using the view count at the same day removes the effects of video age, so that the videos published for longer time are not taking an unfair advantage. In Abortion and Gun Control, the median, as well as 25th and 75th percentile of views of left-leaning videos are higher than that of right-leaning videos. The median views for left-leaning videos are 107,346 for Abortion and 153,482 for Gun Control, versus 62,780 and 103,373 for right-leaning ones. The differences in view distribution are statistically significant ($p < 0.01$, Table 2 row 1). For BLM, right-leaning videos have higher median and 75th of views, but the effect is not significant.

Relative engagement. From Figure 3(b), we can see that videos in all three topics are highly engaging, with mean relative engagement at 0.834 for Abortion, 0.824 for Gun Control, and 0.831 for BLM. This is because our data processing procedure requires videos to have at least 100 tweets and 100 views, which tends to select videos with significant amount of interests. Left-leaning videos are significantly more engaging than their right-leaning counterparts

across all three topics ($p < 0.05$, Table 2 row 2).

Fraction of likes. Figure 3(c) presents the proportion of likes in videos' reactions. Left-leaning videos across all topics have significantly smaller fraction of likes than right-leaning videos ($p < 0.001$, Table 2 row 3). This may be explained by the observation that there are far more cross-partisan talks on left-leaning videos (Wu and Resnick 2021).

Viral potential. Figure 3(d) shows the distributions of viral potential. We find that the left-leaning videos have significantly higher viral scores than the right-leaning videos across all three topics ($p < 0.05$, Table 2 row 4), meaning that given the same amount of tweets exposing the video on Twitter, an average left-leaning video can effectively attracts more views than an average right-leaning video. The difference is most notable in Abortion: a typical left-leaning video receives 224 views from an average tweet, whereas a typical right-leaning video receives only 63 views.

Tweet counts. Figure 3(e) and (f) show the distributions of total tweets and retweets. Contrasting to the observation that left-leaning videos are more viewed, here we find that right-leaning videos are significantly more tweeted, especially with more retweets and more replies ($p < 0.001$, Table 2 row 5-7) in Abortion and BLM. On the other hand, we do not observe a significant difference in original tweets and quotes, except for BLM where right-leaning videos have prevailing volume across all tweet types.

To examine the robustness of presented results in this section, we bootstrapped videos for each topic and for each ideological group. Specifically, for each group, we created 1,000 bootstrapped sets of videos that are of the same size as the original group (shown in Table 1). Next, we computed the mean of proposed metrics (shown in Figure 3) for each bootstrapped set. Lastly, we used the *independent t-test* to check the statistical significance between left- and right-leaning groups. The results of the *t-tests* support all reported relations in Table 2 row 1-6 with $p < 0.001$.

5.2 Views and Tweets over Time

We measure how quickly left- and right-leaning videos attract views and tweets. We find that left-leaning videos are reacted on YouTube and Twitter quicker across all topics.

Viewing half-life and Tweeting half-life. We notice that there are significant differences in the attention consumption patterns: right-leaning videos have more prolonged attention spans on YouTube across all topics ($p < 0.01$, Table 2 row 10). Right-leaning videos also have longer attention spans on Twitter for Abortion and Gun Control ($p < 0.05$, Table 2 row 11). For example, Figure 4(a) shows that right-leaning videos for Abortion have the longest attention span – taking 9 days for 75% videos to achieve viewing half-life, while left-leaning videos only take 3 days. Comparing Figure 4(a) to Figure 4(b), we find that attention spans on Twitter are shorter than that on YouTube. In Abortion, left-leaning videos take 2 days for 75% of videos to reach tweeting half-life (vs. 3 days for views) and right-leaning videos take 5 days for 75% of videos to reach tweeting half-life (vs. 9 days for views).

Tweeting lifetime and Tweeting inter-arrival time. Figure 4 (c) and (d) shows the distributions of tweeting lifetime

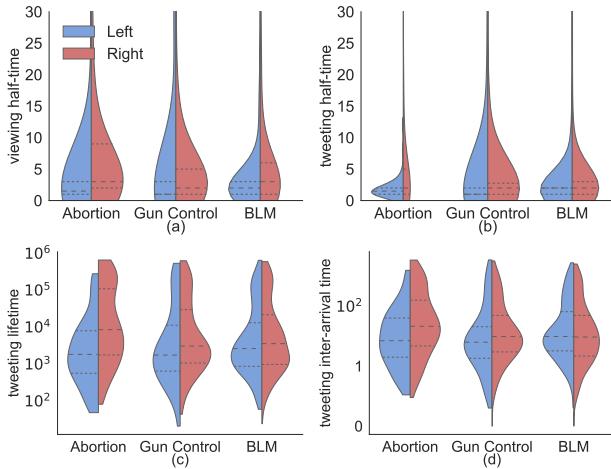


Figure 4: The number of days to receive 50% (a) views and (b) tweets. The time gaps (in minutes) between tweets measured as (c) tweeting lifetime and (d) tweeting inter-arrival time. Left-leaning videos are mentioned more quickly on Twitter and have shorter circulation duration for all metrics in Abortion and Gun Control.

and inter-arrival time. The results are mixed across topics. For Abortion and Gun Control, the MWU test results show that left-leaning videos have significantly less scores in both metrics than right-leaning videos ($p < 0.05$, Table 2 row 12-13). But for BLM, both metrics show similar distributions between left- and right-leaning videos.

These results suggest that left-leaning videos have shorter circulation duration and are mentioned more quickly (except BLM). The most notable difference is in Abortion where the median of tweeting lifetime and inter-arrival time for right-leaning videos are more than three times of those for left-leaning videos. For instance, the median tweeting lifetime is 1811.6 minutes for left-leaning videos, while the median is 8453.8 minutes for right-leaning videos.

Accumulation of views and tweets. We examine how much views and tweets are accumulated each day. Figure 5(a) and (b) show the Complementary Cumulative Distribution Function (CCDF) of views and tweets percentages accumulated for the first day (video published date) and 30th day for Abortion videos. We observe that left-leaning videos tend to achieve views/tweets faster than right-leaning videos, which is consistent with Figure 4. For example, after day 1, 56.9% left-leaning videos have achieved viewing half-life, but only 26.1% right-leaning videos achieved the same. For tweet accumulation, after 1 day, the gap of tweeting half-life between left- and right-leaning videos is 11.4% (46.6% and 35.1%, respectively). By day 30, only 3 left-leaning videos have not accumulated 80% of views.

Figure 5(c) compares the normalized MWU statistic values of left- and right-leaning videos in views \bar{U}_v and tweets \bar{U}_t on each of the 120 days since upload. It shows that the difference between left- and right-leaning videos is larger in the beginning and decreases towards 0.5 over time. At the 120th day (as the observation period ends), both will

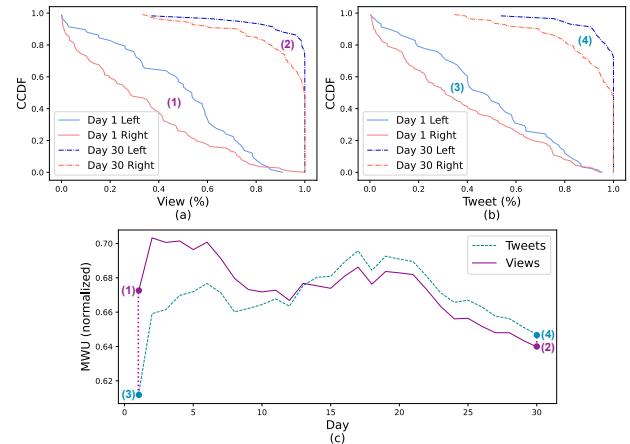


Figure 5: The distribution of accumulated views and tweets in Abortion videos on the first day of video upload and at 30th day: (a) percentage of views achieved, (b) percentage of tweets achieved, (c) MWU statistic values for each day from the first day to day 30. The area marked with (1,2,3,4) in (a) and (b) are shown in (c) as a normalized MWU statistic. For example, (1) shows difference between left- and right-leaning videos in terms of view accumulation. This difference is measured using MWU test and shown in (c) where (1) is marked. The difference between left- and right-leaning videos, in both views and tweets accumulation, tend to become smaller in later days. Section F of (Appendix 2022) includes the same set of plots for Gun Control and BLM.

be 0.5 by definition, we thus truncate the plot at day 30. We also observe that the differences in views is larger than that of tweets across left- and right-leaning videos initially, whereas the discrepancy between views and tweets narrows as videos get older. This is because more videos have already fully achieved all the views and tweets.

5.3 Videos' Tweet Cascades

Cascade size. Goel, Watts, and Goldstein (2012) found that one-node-cascades (isolated cascades) account for 96% of all cascades in their *Twitter Videos* dataset. Figure 6(a) shows that the proportions of isolated cascades in our datasets are lower, measured at 91.5%, 91.2%, and 91.9% respectively for Abortion, Gun Control, and BLM. Notwithstanding a confounder that the dynamics on Twitter has changed significantly since (Goel, Watts, and Goldstein 2012), this may still suggest that tweets on controversial topics are less isolated than tweets of videos about any topics. Figure 6(b) shows the volume of tweets involved in each cascade group. It is interesting to find that most tweets belong to either isolated cascades or large cascades. Aggregated over left and right-leaning videos in Abortion, Gun Control, and BLM, 47%, 43.4%, and 47.3% of tweets are isolated, whereas 44.7%, 48.4%, and 44.3% are in large cascades of size 5 and above, dominated by a handful of cascades over 1,000 tweets.

Cascade start time. We compare when tweet cascades start in the process of view accumulation, grouped by differ-

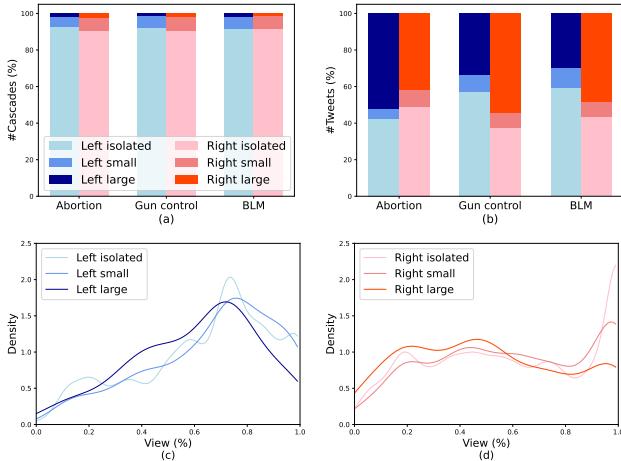


Figure 6: Comparing the left- and right-leaning videos in (a) proportions of cascade sizes – isolated (1 tweet), small (2-4 tweets) and large (5+ tweets); (b) volume of tweets in each cascade type. (c) and (d) show the density of cascade start time in relation to accumulated view percentage for Abortion videos. Because the left-leaning videos accumulate views quicker (90% of left-leaning videos reach viewing half-life within 3 days after upload while 59.4% of right-leaning videos do), more left-leaning tweet cascades are shown to start after viewing half-life. Section G of (Appendix 2022) includes the same set of plots for Gun Control and BLM.

ent cascade sizes. Figure 6(c) and (d) show the distribution of cascade start times of left- and right-leaning videos in Abortion relative to percentages of views. For right-leaning videos, there is a peak of isolated cascades started at the end of the videos' viewing lifetime. We also observe that for left-leaning videos, the peaks for isolated, small and large cascades are concentrated near 70% of view accumulation while peaks for right-leaning videos are more distributed over different stages of view accumulation. Moreover, in all topics, more right-leaning tweet cascades start before viewing half-life regardless of cascades size. For example, 41% of right-leaning isolated cascades started before viewing half-life while 25% of left-leaning isolated cascades started before viewing half-life in Abortion. This is consistent with the observation that views of right-leaning videos unfold much slower (See Figure 4(a) and Figure 5(a)), allowing tweet cascades to start at earlier stages of view accumulation process. The difference in cascade start time is significant between left- and right-leaning videos in all topics for isolated and small cascades ($p < 0.001$, Table 2 row 14-15) and is significant for Abortion and Gun Control for large cascades ($p < 0.001$, Table 2 row 16).

5.4 Networks among Early Adopters on Twitter

Figure 7(c) and (d) show the distributions of the Gini coefficient of indegree centrality and the Gini coefficient of closeness centrality. Gini coef. of indegree centrality of left-leaning videos' networks is significantly larger in Gun

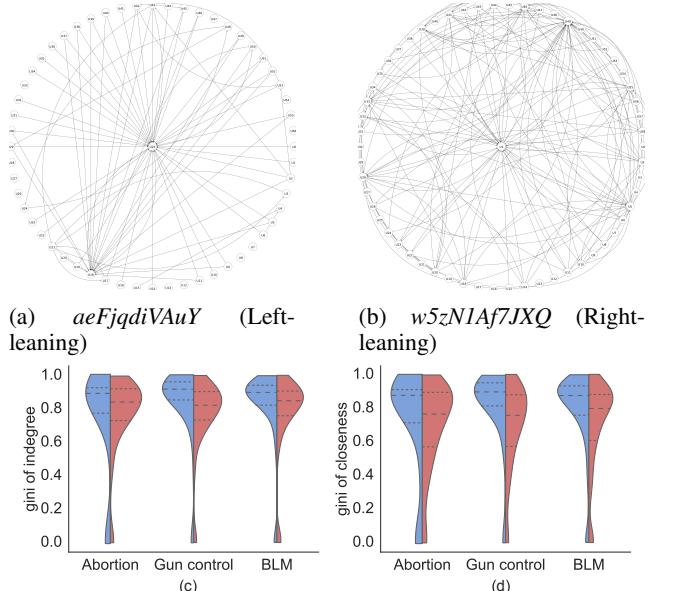


Figure 7: Follower networks of early adopters for (a) a left-leaning and (b) a right-leaning video with similar network size in Abortion. We compare the left- and right-leaning videos in (c) Gini coefficient of indegree centrality, and (d) Gini coefficient of closeness centrality.

Control and BLM ($p < 0.001$, Table 2 row 17). For Gini coefficient of closeness centrality, the MWU test results indicate that left-leaning videos' networks have significantly greater Gini index than those of right-leaning videos' networks across all topics ($p < 0.05$, Table 2 row 18).

This suggests that the networks of early adopters for left-leaning videos have more users serving as hubs, i.e., who are followed by more early adopters and have shorter path to other early adopters. This also suggests that in the networks of early adopters for right-leaning videos, users are more equally facilitating dissemination of political information which is consistent with the findings shown in (Conover et al. 2012). As an example of this, we present the follower networks of early adopters of one left-leaning video and one right-leaning video in Abortion in Figure 7(a,b). To have a fair comparison we sample two videos having similar network size (57 and 59 for left and right-leaning videos, respectively). The left-leaning video has Gini coef. of indegree centrality: 0.918, Gini coef. of closeness centrality: 0.90. The right-leaning video has Gini coef. of indegree centrality: 0.748, Gini coef. of closeness centrality: 0.536. It can be observed that the sharing of this left-leaning video relies more on central users who are followed by more early adopters and have shorter path to others. On the other hand, in the follower network of the early adopters of this right-leaning video, indegree and closeness centrality distributions are more equal.

Apart from the reported metrics, we have also performed preliminary examination on the correlation and trends between two and more metrics. An example on linking relative engagement to the view and tweet counts is presented in Sec-

tion I of (Appendix 2022). We have not seen consistent and salient patterns that are not already captured by individual measures.

6 Conclusion and Discussion

This work presents a quantitative study that links collective attention towards online videos across YouTube and Twitter over three political topics: Abortion, Gun Control, and BLM. For each topic, we curated a cross-platform dataset that contained hundreds of videos and hundreds of thousands of tweets spanning 16 months. The extracted videos all have a non-trivial amount of views and tweets. The key contributions include several sets of metrics for comparing attention consumption patterns between left-leaning and right-leaning videos across two platforms. We find that left-leaning videos are more viewed and more engaging, while right-leaning videos are more tweeted and have longer attention spans. We also found that the follower networks of early adopters on left-leaning videos are of higher centrality, whereas tweet cascades for right-leaning videos start earlier in the attention lifecycle. This study enriches the current understanding of ideological asymmetries by adding a set of temporal and cross-platform analyses.

Limitations. Extensive discussions about social data biases are presented in (Olteanu et al. 2019). The biases can be introduced due to the choice of social platforms, data (un)availability, sampling methods, etc. Here we discuss three limitations in our data collection process.

A recent study found that Twitter filtered streaming API might subsample high-volume data streams that consist of more than 1% of all tweets (Wu, Rizou, and Xie 2020). The authors proposed a method of using the Twitter rate limit messages to quantify the data loss. By using the suggested method, we find our 16-month Twitter stream has a sampling rate of 79.4 % – we collected 1,802,230,572 out of 2,270,223,254 estimated total tweets. Assuming the Twitter sampling is a Bernoulli process, for a video that has been tweeted twice, we have 95.8% chance of collecting this video in our tweet stream. And most missing videos are tweeted sporadically. Therefore, the sampling effect of Twitter APIs for decomposing the video list is small, whereas the measures on tweeting activities, including attention volumes, timing, cascade sizes, are affected by such sampling. Confidence intervals for simple measures such as volume can be derived from the Bernoulli model (Wu, Rizou, and Xie 2020). Estimates for the unsampled version of more complex measures can be obtained by simulation, and left as future work.

In this paper, we present various measurements focused on YouTube videos, which are the main entities that link the two platforms. YouTube viewers are unknowable (via publicly available data) and Twitter users are hard to track consistently over time. Therefore, we track videos which attract views and tweets. All the presented metrics are video-centric and we do not assume that the viewers or tweeters of the videos represent specific groups of users. We believe that each set of videos (Abortion, Gun Control, and BLM) represent YouTube videos that are relevant to the topic, curated by keyword queries and semi-manual coding. The

number of videos belonging to each topic is not large but we attempted to include all published videos shared on Twitter which are relevant and have non-trivial activities. Thus our results intend to explain attention gathering behavior of topic-relevant videos.

Another data integrity limitation is the time gap between tweet collection in 2017–2018 and early adopters’ follower networks collected in early 2020. Our data collection is limited by the Twitter search API quota. It is difficult to collect tweets and Twitter user followers simultaneously due to the tight quota. The tweeted videos stream has on average 3.7M tweets in each day. Collecting the follower network for all these tweets far exceeds the capacity of Twitter API, focusing on the early-adopter network is a practical trade-off between still having informative results and making data collection feasible. An associated issue comes from the unavailable YouTube videos and Twitter users since content publicly available in 2017 may become deleted, banned, or protected in 2020. In our experiments, we find that between 17% to 19% candidate videos become unavailable.

Practical implications and future work. We believe this work adds a new dimension to the understanding of online political behavior and discourse – cross platform links. Further examination in this direction could bear theoretical and empirical fruits. The measurements presented in this work are mostly quantitative. One direction of future work is to complement with qualitative analysis. For example, to strengthen our observations that left-leaning YouTube videos reply more on a group of elite early adopters, one can visualize two typical diffusion networks from both left- and right-leaning groups and the diffusion process of the video spreading. One could also examine the framing of left- and right-leaning content in both video descriptions and tweets about them. The second direction is to expand and enrich the quantitative analyses, such as investigating the linguistic features of YouTube descriptions and tweet cascades, and verify whether cascade structure is predictive of the amount and timing of views. Finally, understanding the collective attention across multiple social platforms is important for content producers, who could devise better strategies in promoting their content in another domain.

Ethical Statement

All data that we obtained was publicly available at the time of data collection. We discarded deleted, protected, and private content at the time of analysis. In our released dataset, we will anonymize user identities. Therefore, the analyses reported in this work do not compromise any user privacy.

Acknowledgments

We thank Xian Teng and Muheng Yan for their help in data annotation. We would like to acknowledge the support from NSF #2027713, ARC DP180101985, and AFOSR projects 19IOA078, 19RT0797, 20IOA064, and 20IOA065. Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect the views of the funding sources.

Table 2: Summary of all metrics and comparison across political leanings. **L > R*** means that the metric of a randomly selected left-leaning video is significantly larger than that of a randomly selected right-leaning video. Significance is measured by one-sided Mann–Whitney U test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The significantly larger leaning is boldfaced. \times sign indicates non-significant relation. Sample size: Abortion (L: 58; R: 111), Gun Control (L: 81; R: 154), BLM (L: 297; R: 396). See the corresponding cross-referenced figures and discussions in Section 5 for more details, – sign means the metric is not presented in a figure.

row	crossref	metric	definition	significance		
				Abortion	Gun Control	BLM
YouTube and Twitter attention metrics (Section 5.1)						
1	fig. 3	view_x	Total number of views at day x	L > R**	L > R***	\times
2	fig. 3	relative_engagement	Rank percentile of watch percentage among videos of similar lengths	L > R*	L > R***	L > R**
3	fig. 3	fraction_of_likes	Number of likes divided by total number of reactions	L < R***	L < R***	L < R***
4	fig. 3	viral_potential	Number of views potentially excited by one tweet	L > R***	L > R**	L > R*
5	fig. 3	tweet_x	Total number of tweets at day x	L < R***	\times	L < R***
6	fig. 3	retweet_x	Total number of retweets at day x	L < R***	\times	L < R***
7	–	reply_x	Total number of replies at day x	L < R***	\times	L < R***
8	–	original_tweet_x	Total number of original tweets at day x	\times	\times	L < R***
9	–	quote_x	Total number of quotes at day x	\times	\times	L < R***
Temporal metrics of views and tweets (Section 5.2)						
10	fig. 4	viewing_half-life	Number of days to reach 50% views	L < R***	L < R**	L < R***
11	fig. 4	tweeting_half-life	Number of days to reach 50% tweets	L < R**	L < R*	\times
12	fig. 4	tweeting_lifetime	Time gap between the first and the last tweets in minutes	L < R***	L < R*	\times
13	fig. 4	tweeting_inter-arrival_time	Average time gap between every two consecutive tweets in minutes	L < R**	L < R*	\times
Tweet cascades measures (Section 5.3)						
		<u>cascade start time</u>	Percentage of accumulated views of the video when the root of the cascade is tweeted			
14	fig. 6	(isolated cascades)		L > R***	L > R***	L > R***
15	fig. 6	(small cascades)		L > R***	L > R***	L > R***
16	fig. 6	(large cascades)		L > R***	L > R***	\times
Network metrics of early adopters (Section 5.4)						
17	fig. 7	Gini_indegree	Gini coefficient of indegree centrality	\times	L > R***	L > R***
18	fig. 7	Gini_closeness	Gini coefficient of closeness centrality	L > R*	L > R***	L > R***
19	–	Gini_betweenness	Gini coefficient of betweenness centrality	\times	L > R*	\times
20	–	network_density	Density of early adopters' follower network	\times	\times	L > R***
21	–	max_indegree	Max indegree value in early adopters' follower network	L < R**	\times	\times
22	–	global_efficiency	Average efficiency over all pairs of distinct early adopters	L < R*	L < R**	\times

References

- Abisheva, A.; Garimella, V. R. K.; Garcia, D.; and Weber, I. 2014. Who watches (and shares) what on YouTube? And when? Using Twitter to understand YouTube viewership. In *WSDM*.
- Appendix. 2022. Online appendix: Whose advantage? Measuring attention dynamics across YouTube and Twitter on controversial topics. URLtobeupdated.
- Bakshy, E.; Messing, S.; and Adamic, L. A. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science*.
- Brady, W. J.; Wills, J. A.; Burkart, D.; Jost, J. T.; and Van Bavel, J. J. 2019. An ideological asymmetry in the diffusion of moralized content on social media among political leaders. *Journal of Experimental Psychology: General*.
- Conover, M. D.; Gonçalves, B.; Flammini, A.; and Menczer, F. 2012. Partisan asymmetries in online political activity. *EPJ Data Science*.
- Conover, M. D.; Ratkiewicz, J.; Francisco, M. R.; Gonçalves, B.; Menczer, F.; and Flammini, A. 2011. Political polarization on Twitter. In *ICWSM*.
- De Choudhury, M.; Jhaver, S.; Sugar, B.; and Weber, I. 2016. Social media participation in an activist movement for racial equality. In *ICWSM*.
- Ertugrul, A. M.; Lin, Y.-R.; Chung, W.-T.; Yan, M.; and Li, A. 2019. Activism via attention: Interpretable spatiotemporal learning to forecast protest activities. *EPJ Data Science*.
- Freelon, D.; Marwick, A.; and Kreiss, D. 2020. False equivalencies: Online activism from left to right. *Science*.
- Garimella, K.; De Francisci Morales, G.; Gionis, A.; and Mathioudakis, M. 2018. Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *TheWebConf*.
- Goel, S.; Watts, D. J.; and Goldstein, D. G. 2012. The structure of online diffusion networks. In *EC*.
- HosseiniMardi, H.; Ghasemian, A.; Clauset, A.; Mobius, M.; Rothschild, D. M.; and Watts, D. J. 2021. Examining the consumption of radical content on YouTube. *PNAS*.
- Jackson, S. J.; Bailey, M.; and Welles, B. F. 2020. *#HashtagActivism: Networks of race and gender justice*. MIT Press.
- Ledwich, M.; and Zaitsev, A. 2020. Algorithmic extremism: Examining YouTube’s rabbit hole of radicalization. *First Monday*.
- Mann, H. B.; and Whitney, D. R. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*.
- Olteanu, A.; Castillo, C.; Diaz, F.; and Kiciman, E. 2019. Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data*.
- Ottoni, R.; Cunha, E.; Magno, G.; Bernardino, P.; Meira Jr, W.; and Almeida, V. 2018. Analyzing right-wing YouTube channels: Hate, violence and discrimination. In *WebSci*.
- Rizoiu, M.-A.; and Xie, L. 2017. Online popularity under promotion: Viral potential, forecasting, and the economics of time. In *ICWSM*.
- Rizoiu, M.-A.; Xie, L.; Sanner, S.; Cebrian, M.; Yu, H.; and Van Hentenryck, P. 2017. Expecting to be HIP: Hawkes intensity processes for social media popularity. In *TheWebConf*.
- Romero, D. M.; Tan, C.; and Ugander, J. 2013. On the interplay between social and topical structure. In *ICWSM*.
- Schradie, J. 2019. *The revolution that wasn’t: How digital activism favors conservatives*. Harvard University Press.
- Stewart, L. G.; Arif, A.; Nied, A. C.; Spiro, E. S.; and Starbird, K. 2017. Drawing the lines of contention: Networked frame contests within #BlackLivesMatter discourse. In *CSCW*.
- Thorson, K.; Driscoll, K.; Ekdale, B.; Edgerly, S.; Thompson, L. G.; Schrock, A.; Swartz, L.; Vraga, E. K.; and Wells, C. 2013. YouTube, Twitter and the Occupy movement: Connecting content and circulation practices. *Information, Communication & Society*.
- Waller, I.; and Anderson, A. 2021. Quantifying social organization and political polarization in online platforms. *Nature*.
- Wu, S.; and Resnick, P. 2021. Cross-partisan discussions on YouTube: Conservatives talk to liberals but liberals don’t talk to conservatives. In *ICWSM*.
- Wu, S.; Rizoiu, M.-A.; and Xie, L. 2018. Beyond views: Measuring and predicting engagement in online videos. In *ICWSM*.
- Wu, S.; Rizoiu, M.-A.; and Xie, L. 2020. Variation across scales: Measurement fidelity under Twitter data sampling. In *ICWSM*.
- Xiao, Z.; Song, W.; Xu, H.; Ren, Z.; and Sun, Y. 2020. TIMME: Twitter ideology-detection via multi-task multi-relational embedding. In *KDD*.
- Zannettou, S.; Caulfield, T.; De Cristofaro, E.; Kourtelris, N.; Leontiadis, I.; Sirivianos, M.; Stringhini, G.; and Blackburn, J. 2017. The web centipede: Understanding how web communities influence each other through the lens of mainstream and alternative news sources. In *IMC*.
- Zhang, A. X.; and Counts, S. 2015. Modeling ideology and predicting policy change with social media: Case of same-sex marriage. In *CHI*.
- Zhang, A. X.; and Counts, S. 2016. Gender and ideology in the spread of anti-abortion policy. In *CHI*.
- Zhang, Y.; Shah, D.; Foley, J.; Abhishek, A.; Lukito, J.; Suk, J.; Kim, S. J.; Sun, Z.; Pevehouse, J.; and Garlough, C. 2019. Whose lives matter? Mass shootings and social media discourses of sympathy and policy, 2012–2014. *Journal of Computer-Mediated Communication*.
- Zhou, D.; Bousquet, O.; Lal, T. N.; Weston, J.; and Schölkopf, B. 2004. Learning with local and global consistency. In *NeurIPS*.

A Identifying Topical Videos and Tweets

We curated a keyword list for each of the three controversial topics – Abortion, Gun Control, and BLM. Keywords for Abortion and Gun Control came from another recent work (Guo et al. 2020) and were expanded using contextual expansion with sample tweets. Keywords for BLM were manually curated, separating into pro- and anti-BLM keywords respectively. We expanded the latter with white-supremacy related keywords. Table 3 lists all keywords.

We consider a video is potentially relevant if (a) it contains at least one keyword in the video title or description; or (b) it is mentioned in a tweet that contains the keywords in the tweet text. We obtained 815 candidate videos for Abortion, 974 candidate videos for Gun Control, and 8,571 candidate videos for BLM.

For Abortion and Gun Control, we manually annotated all candidate videos. All videos were randomly divided into five buckets (with 163 or 195 videos in each) and each bucket was labeled by an annotator. The five authors of this paper, along with two postgraduate students who had extensive research experience in computational social science, consisted of the annotator team. Annotators were instructed to watch a video for at least five minutes before making a decision of whether or not this video is relevant to the target topic. To measure the inter-rater reliability, we randomly selected a bucket, and had two more annotators labeling the same bucket. We computed the Fleiss' Kappa. Abortion had Fleiss' Kappa of 0.830 and Gun Control had that of 0.826, suggesting a high agreement among annotators.

For BLM, we used a semi-automatic approach because the number of candidate videos (8,571) is much larger. Based on our observations on Abortion and Gun Control, we designed the following protocol:

- For videos that contained any topic-relevant keywords in the video titles or descriptions,
 - if the videos were also mentioned by tweets with topic-relevant keywords, 89.0%/86.7% of the videos were labeled *relevant* (in Abortion and Gun Control respectively, same below). These videos consisted of 16.1%/12.5% of all candidate videos. We labeled such videos as relevant, there were 418 in BLM.
 - if the videos were not mentioned by tweets with topic-relevant keywords, 77.8%/89.4% of the videos were labeled *irrelevant*. These videos consisted of 6.7%/11.9% of all candidate videos. We labeled such videos as irrelevant.
- For videos that did not contain any topic-relevant keywords in the titles nor descriptions, but had been mentioned by some topic-relevant tweets,
 - if the videos were tagged as “Music”, or did not have English captions, or did not contain any topic-relevant keywords in the captions, 95.2%/80.6% of the videos were labeled *irrelevant*. We labeled such videos as irrelevant.
 - we used the manual annotation steps for Abortion and Gun Control to label the remaining 749 videos, out of which 359 videos were labeled relevant. The Fleiss' Kappa was 0.711.

In total, we identified 179 Abortion videos, 268 Gun Control videos, and 777 BLM videos. We extracted all tweets and Twitter users that mentioned any of these topic-relevant videos from our 16 months data stream.

Topic	Keywords
Abortion	abortion, pro choice, pro life, ru486, planned parenthood, unborn baby, unborn babies, unborn child, roe v wade, roe versus wade, reproductive health care, birth control
Gun Control	gun violence, gun reform, gun safety, gun regulation, gun death, gun law, gun right, gun grab, gun owner, gun control, gun free, gun sense, guns kill people, pro gun, firearm, no bill no break, disarm hate, second amendment, moms demand, wear orange, molon labe
BLM	blm, black lives matter, all lives matter, blue lives matter, white lives matter, michael brown, mike brown, eric garner, freddie gray, walter scott, tamir rice, john crafford, tony robinson, eric harris, ezell ford, akai gurley, kajimeme powell, tanisha anderson, victor white, jordan baker, jerame reid, yvette smith, philip white, dante parker, mckenzie cochrane, tyree woodson, jacques clemmons, fire Bannon, charlottesville, kkk, ku klux klan, klansmen, klansmen, domestic terrorism, white supremacist

Table 3: Keyword list for the three controversial topics.

B Estimating and Validating the Leanings of Twitter Users

We estimated the political leanings of early adopters on Twitter in four steps with increasing coverage.

1) Curating seed hashtags. Twitter users are known to use political hashtags in their user profiles as a marker of identity and a tool of network gatekeeping, such as marking participation in an ongoing activism, facilitating the establishment of social network ties around specific topics, or expressing community identities (Kulshrestha et al. 2017). Based on this insight, we

started by identifying users who clearly signaled political ideologies in their Twitter profile descriptions. We curated two sets of ideologically descriptive hashtags that were typically used by liberal and conservative users. We described a hashtag as ideologically descriptive if it either supported or criticized political leaders, their statements, and their campaigns (e.g., `#trump`, `#america1st`, `#maga`). We called these hashtags as political seed hashtags. The whole list can be found in Table 4. There are 25 left-leaning hashtags used by 260, 552, 817 users, and 35 right-leaning hashtags used by 1023, 2065, 3798 users in Abortion, Gun Control, and BLM, respectively.

Leaning	Hashtags
Left-leaning	theresistance, resist, resistance, fbr, impeachtrump, impeachtrumpnow, stillwithher, bluewave, uniteblue, bluewave2018, resisttrump, nevertrump, fucktrump, imwithher, demforce, voteblue, followbackresistance, bernie2020, resistanceunited, impeach45, imstillwithher, democrat, democrats, alwayswithher, flipitblue.
Right-leaning	maga, trump, americafirst, draintheswamp, trumptrain, trump2020, kag, makeamerica-greatagain, lockherup, presidenttrump, cruzcrew, trumppence2020, trumparmy, womenfortrump, istandwithtrump, keepamerica-great, neverhillary, trump2016, trumpworld, republican, republicans, hillaryforprison, kag2020, gop, trumpsupporter, trumpville, trumppence16, alwaystrump, donaldtrump, america1st, trumpismypresident, trump45, women4trump, teamtrump.

Table 4: Political seed hashtags.

2) Expanding seed hashtags. To obtain new political hashtags, we followed an expansion protocol based on co-occurrence with the political seed hashtags in the user profile descriptions. Firstly, we required that a candidate hashtag must co-occur with the seed hashtags in at least 0.1% of all users' profile descriptions. Secondly, the candidate hashtag should be relevant to only one ideology (either left or right). To achieve this, we computed the Shannon entropy based on the co-occurrence of the candidate hashtag with the seed left- and right-leaning hashtags. We selected the hashtags with entropy less than or equal to 0.1. Table 5 presents the expanded political hashtags. The number of users using left- and right-leaning hashtags are 264 and 1250 for Abortion, 677 and 2375 for Gun Control, 970 and 4412 for BLM, respectively. We noticed that the number of conservative users are about four times that of liberal users across the three topics.

	Left-leaning hashtags	Right-leaning hashtags
Abortion	notmypresident	2a, nra, conservative, 1a, patriot, prolife, buildthewall, tcot, ccot, military, bluelivesmatter, usa, qanon, deplorable
Gun Control	blacklivesmatter, blm, notmypresident	2a, nra, 1a, conservative, patriot, prolife, christian, buildthewall, qanon, deplorable, backtheblue, vets, bluelivesmatter, usa
BLM	blacklivesmatter, blm, notmypresident progressive, metoo	2a, nra, 1a, conservative, prolife, buildthewall, christian, bluelivesmatter, tcot, ccot, qanon, deplorable, israel, constitution, usa, backtheblue, vets, covfefe, banislam, codeofvets, bansharia

Table 5: Expanded political hashtags from political seed hashtags.

3) Identifying seed users. After curating and expanding the left- and right-leaning hashtags (jointly called *political hashtags*), we assigned the political ideologies to the Twitter users. For users whose profile descriptions included at least one political hashtag, we counted the numbers of left- and right-leaning hashtags in their profile descriptions. The average numbers of political hashtags per such user were 1.96, 1.97, and 2.04 for Abortion, Gun Control and BLM, respectively. If the ratio of one leaning was greater than or equal to 0.9, we assigned the corresponding ideology to the user. We called the users *seed users*, since their labels were later used to estimate the ideologies of other users. The numbers of liberal and conservative seed users are given in Table 6.

4) Propagating labels to other users. We aimed to infer the ideologies of users who had no explicit political hashtags in their profile descriptions. We first created a user-to-user network where the edge weight between two users (u, v) represented their *shared audience*, calculated using the Jaccard similarity of their follower lists:

$$J(u, v) = \frac{\text{followers}(u) \cap \text{followers}(v)}{\text{followers}(u) \cup \text{followers}(v)} \quad (1)$$

The shared audience network is shown to be an effective method for splitting users with different political behaviors (Stewart et al. 2017). When estimating the user ideologies, our assumption is that users with the same ideology share similar audience.

	#Seed users			Ideology propagation (10-fold CV performance)		
	Liberal	Conservative	Total	Precision (Lib-Con)	Recall (Lib-Con)	F-score (Lib-Con)
Abortion	258 (2.07%)	1244 (9.99%)	1502 (12.06%)	0.99 - 0.99	0.96 - 1.00	0.98 - 1.00
Gun Control	664 (2.38%)	2362 (8.45%)	3026 (10.83%)	0.98 - 0.99	0.97 - 0.99	0.98 - 0.99
BLM	950 (1.90%)	4392 (8.76%)	5342 (10.66%)	0.98 - 0.99	0.96 - 1.00	0.97 - 0.99

Table 6: Political seed users and ideology propagation results. See Appendix B for methods.

The shared audience network was a dense network since we created edges between any two users regardless of whether they had a direct follower-following relation. The initial networks consisted of approximately $28M$, $119M$, and $394M$ edges for Abortion, Gun Control, and BLM, respectively. The corresponding network densities were 18.1%, 15.2%, and 15.7%. To reduce the network complexity, we used a disparity filtering algorithm (Serrano, Boguná, and Vespignani 2009) to extract the network backbone. The disparity filtering algorithm can identify important edges in a network. It has a hyperparameter α to control the degree of edge importance. We used $\alpha = 0.05$ as suggested in the initial work (Serrano, Boguná, and Vespignani 2009). After disparity filtering, the resulting networks contained approximately $1.9M$, $7.8M$, and $25.5M$ edges for Abortion, Gun Control, and BLM, respectively. The densities of these networks ranged between 1% and 1.2%.

Next, we applied a label propagation algorithm (Zhou et al. 2004) on the filtered network to infer the political leanings of the unknown users from those of the seed users. Label propagation is a semi-supervised algorithm that builds on two intuitions: (a) the seed labels should not change too much after propagation; (b) the inferred labels should not be too different with their neighbours. A teleport factor β controls the trade-off between these two intuitions. We used $\beta = 0.85$ as suggested in the prior work (Yan et al. 2020). The label propagation algorithm returned a liberal score and a conservative score for each user in the network, which we then re-normalized into a score between 0 and 1, where 0 indicated strong liberal and 1 indicated strong conservative. Re-normalization was done by computing the ratio of conservative score to the summation of liberal and conservative scores.

We stratified one-tenth of seed users in each fold and hid their labels. Next, we compared the inferred political leanings to actual leanings of these users. Note that we employed whole network during the label propagation validation. Table 6 shows the evaluation results for the label propagation.

C Estimating and Validating the Leanings of YouTube Videos

Once we obtained the ideology probabilities of early adopters on Twitter, we computed the leaning score for each YouTube video by averaging its promoters’ ideology scores. To assign discrete leaning labels (i.e. *Left*, *Right* and *Center*) to the videos, we leveraged the Recfluence dataset (Ledwich and Zaitsev 2020) as an external source to annotate YouTube videos’ leanings. This dataset includes 816 YouTube channels where each channel has 10k+ subscribers and more than 30% of its content is relevant to US politics, or cultural news, or cultural commentary. It consists of channels of YouTubers in addition to those of traditional news outlets. In this dataset, each channel is assigned to *Left* (*L*), *Center* (*C*) or *Right* (*R*) leaning. The annotations were conducted based on two well known media bias resources⁴ as well as manual labeling.

The videos whose channels’ leanings are available in the Recfluence dataset are assigned to the same leaning. We were able to label 109 videos (*L*: 38, *R*: 66, *C*: 5) from 38 channels for Abortion, 153 videos (*L*: 52, *R*: 88, *C*: 13) from 62 channels for Gun Control, and 521 videos (*L*: 194, *R*: 285, *C*: 42) from 120 channels for BLM. To label the videos whose channels are not available in the Recfluence dataset, we identified two thresholds for each topic – $thr_{(L,C)}$ and $thr_{(C,R)}$ where $thr_{(L,C)} < thr_{(C,R)}$. For a given video ν , we assigned a discrete leaning label, $leaning_label(\nu)$, based on its leaning score, $leaning_score(\nu)$, as follows:

$$leaning_label(\nu) = \begin{cases} L & \text{if } leaning_score(\nu) < thr_{(L,C)} \\ R & \text{if } leaning_score(\nu) > thr_{(C,R)} \\ C & \text{otherwise.} \end{cases}$$

We used videos with labels from Recfluence to find optimum thresholds $thr_{(L,C)}$ and $thr_{(C,R)}$. We obtained a leaning score distribution for each group (i.e., *L*, *R* and *C*). We identified outliers for each group using inter-quartile rule and filtered them out. The leaning score distributions by groups for each topic are given in Figure 8. Next, we found the thresholds $thr_{(L,C)}$ and $thr_{(C,R)}$ using posterior probabilities as follows:

$$P(L|leaning_score(\nu)) = P(C|leaning_score(\nu)), \quad \text{where } leaning_score(\nu) = thr_{(L,C)} \quad (2)$$

$$P(C|leaning_score(\nu)) = P(R|leaning_score(\nu)), \quad \text{where } leaning_score(\nu) = thr_{(C,R)} \quad (3)$$

⁴Media Bias/Fact Check: <https://mediabiasfactcheck.com/>; Ad Fontes Media: <https://www.adfontesmedia.com/>

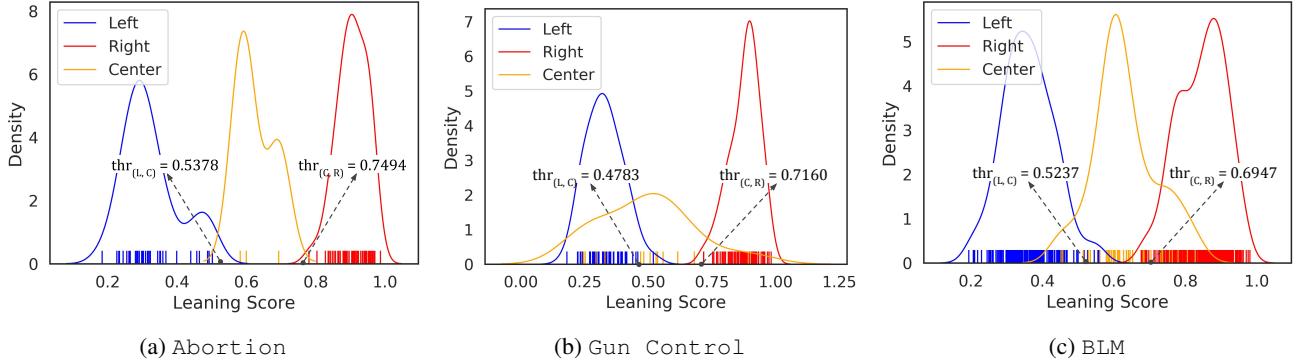


Figure 8: Distribution of leaning scores of videos (generated with kernel density estimation after removal of outliers) whose channels are available in Recfluence dataset (Ledwich and Zaitsev 2020). Each rod in the figures corresponds to a video, and its color represents the leaning label of the video obtained from Recfluence dataset.

The resulting threshold pairs ($thr_{(L,C)}$, $thr_{(C,R)}$) are (0.538, 0.749) for Abortion, (0.478, 0.716) for Gun Control and (0.524, 0.695) for BLM, respectively. Based on these thresholds, we inferred the leaning labels of 70 videos (L: 20, R: 45, C: 5) from 58 channels for Abortion, 115 videos (L: 29, R: 66, C: 20) from 98 channels for Gun Control, 256 videos (L: 103, R: 111, C: 42) from 197 channels for BLM.

To validate the video leaning estimation, we conducted a manual leaning annotation for videos in Gun Control. We used stratified sampling where we sampled 5 videos from each 10 percentile bin of their leaning scores. The leaning of these 50 videos were annotated independently by three authors considering three categories (*left-leaning*, *right-leaning*, *unknown*). The Fleiss' Kappa is 0.691, suggesting a moderate agreement among annotators. We used the majority vote to assign a video label. Table 7 summarizes the annotation results.

We observed that the only decile bin with both manually annotated left- and right-leaning video is the 30th-40th percentile of leaning scores (0.465 – 0.665). Suggesting that the annotator had high agreement with the estimated leaning scores, and that the threshold should be in this interval – which is indeed the case for Gun Control with $thr_{(L,C)}$ lies within its left boundary, and $thr_{(C,R)}$ only 0.05 away from its right boundary. In addition, most of the videos which were assigned *unknown* by the raters are the footage of shooting of Philando Castile and mass shooting events including Las Vegas shooting and TX massacre. These videos do not include any stance on their contents. However, we observed that the leaning scores of these videos align well with the leaning of channels. One example is “Comments on TX church shooting” where the video leaning score belongs to the second bin (80th – 90th) percentile. The channel owner is a far-right YouTuber. The video leaning score matches with the channel owner’s leaning, our approach is thus successful for solving such ambiguities.

Table 7 shows manual leaning annotation results of 50 Gun Control videos. Compared to the leaning score range in the second column (which is estimated), the first three rows of 14 videos (which are labeled as left-leaning) are consistently labeled (the result of manual labeling and estimated leanings agree) as well as the last five rows of 19 videos (which are labeled as right-leaning). The computed thresholds for Gun Control videos, as presented in Figure 8, are 0.478 and 0.716.

Percentiles	Leaning score range	#left-leaning videos	#right-leaning videos	#unknown videos	Videos annotated as unknown
0th - 10th	(0.138 - 0.289)	4	0	1	(1) Shooting footage of Philando Castile (by CBC News (left-center))
10th - 20th	(0.289 - 0.364)	5	0	0	-
20th - 30th	(0.364 - 0.465)	5	0	0	-
30th - 40th	(0.465 - 0.665)	2	1	2	(1) Shooting footage of Philando Castile (2) Las Vegas shooting footage
40th - 50th	(0.665 - 0.789)	0	4	1	(1) Discussion on gun violence in Chicago (by Ron Paul)
50th - 60th	(0.789 - 0.837)	0	3	2	(1) Las Vegas shooting footage. (2) Las Vegas shooting footage (multiple shooters claim).
60th - 70th	(0.837 - 0.875)	0	4	1	(1) Comments on Las Vegas shooting (by a Trump supporter.)
70th - 80th	(0.875 - 0.895)	0	4	1	(1) Las Vegas shooting footage.
80th - 90th	(0.895 - 0.924)	0	3	2	(1) Comments on TX church shooting (by a far-right YouTuber). (2) Comments on Las Vegas shooting (by Milo).
90th - 100th	(0.924 - 0.982)	0	5	0	-
Total	-	16	24	10	-

Table 7: Manual leaning annotation of Gun Control videos.

D Computing Relative Engagement of a Video

From the collected videos we extracted topic (Abortion, Gun Control and BLM) relevant videos for our topical analysis, but all collected videos were used to compute relative engagement scores. We filter out videos that receive less than 100 views within their first 30 days after upload, which is the same filter used by Brodersen, Scellato, and Wattenhofer (2012). We summarize how relative engagement score is computed. For a given video, first we compute two aggregate metrics:

- average watch time ($\bar{\omega}_t$): the total watch time $x_w[1 : t]$ divided by the total view count $x_v[1 : t]$ up to day t

$$\bar{\omega}_t = \frac{\sum_{i=1}^t x_w[i]}{\sum_{i=1}^t x_v[i]}$$

- average watch percentage ($\bar{\mu}_t$): the average watch time $\bar{\omega}_t$ normalized by video duration D

$$\bar{\mu}_t = \frac{\bar{\omega}_t}{D}$$

Then *Engagement map* of tweeted videos is constructed as follows. Two maps are created; the first where x-axis shows video duration D, and the y-axis shows average watch time over the first 120 days ($\bar{\omega}_{120}$) (Figure 9(a)) and the second with the same axis with average watch percentage over the first 120 days ($\bar{\mu}_{120}$) (Figure 9(b)) as y-axis. All videos in the tweeted videos dataset are projected onto both maps. The x-axis is split into 1,000 equally wide bins in log scale. These 2 maps are logically identical but Figure 9(b) is easier to read as y-axis is bounded between [0,1]. This second map is denoted as *Engagement map*.

Based on this *Engagement map*, relative engagement $\bar{\eta}_t \in [0, 1]$ defined as the rank percentile of video in its duration bin. This is an average engagement measure in the first t days. Figure 9(b) illustrates the relation between video duration D, watch percentage $\bar{\mu}_{120}$ and relative engagement $\bar{\eta}_{120}$ for three example videos. V_1 is a left-leaning video in Abortion topic with duration of 136 seconds. V_2 is a right-leaning video in Abortion topic with duration of 141 seconds. V_1 and V_2 have similar lengths but different watch percentages, $\bar{\mu}_{120}(V_1) = 0.83$ and $\bar{\mu}_{120}(V_2) = 0.59$. The difference becomes more apparent with relative engagement scores, $\bar{\eta}_{120}(V_1) = 0.94$ and $\bar{\eta}_{120}(V_2) = 0.27$ as shown in Figure 9(b). V_3 is also a right-leaning video in Abortion which is significantly long than the others. Although V_3 is not as much watched as the previous two, $\bar{\mu}_{120}(V_3) = 0.40$, relative engagement of V_3 is much higher than V_2 , $\bar{\eta}_{120}(V_3) = 0.86$ as longer videos tend to have lower average watch percentages.

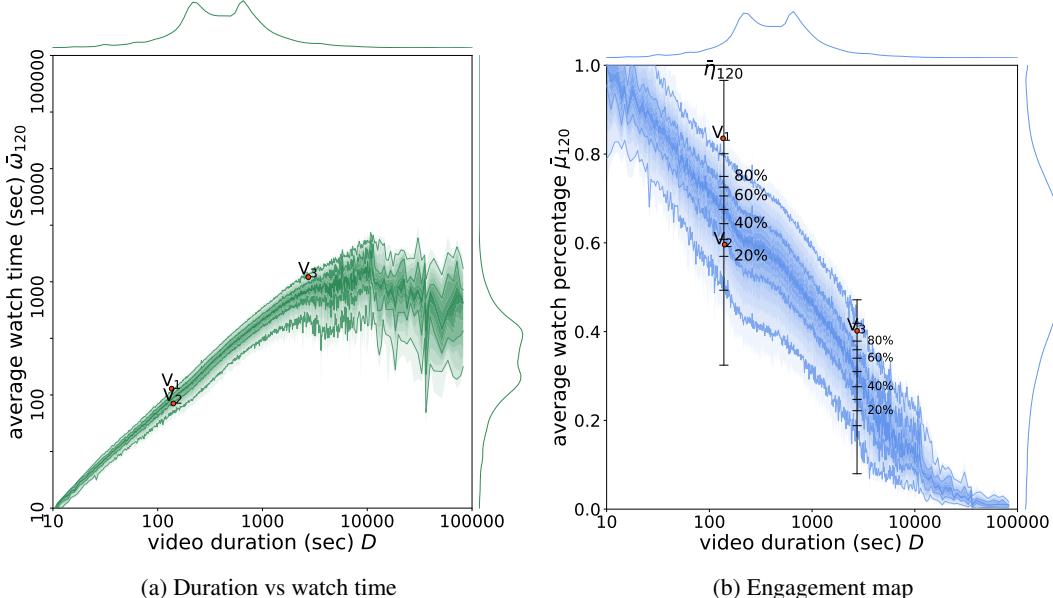


Figure 9: Video engagement in the tweeted videos dataset at the age of 120 days. (a) video duration D vs average watch time $\bar{\omega}_{120}$ (b) the engagement map: video duration D vs average watch percentage $\bar{\mu}_{120}$. This figure originally appeared in (Wu, Rizoiu, and Xie 2018).

E Computing Viral Potential of a Video via the Hawkes Intensity Process

Viral potential estimates the expected number of views a video will obtain if mentioned by an *average* tweet (Rizoiu and Xie 2017). More specifically, it is the area under the impulse response function of an integral equation known as Hawkes Intensity Process (HIP) (Rizoiu et al. 2017), which are learned for each video by using the first 120 days of tweeting and viewing history. The HIP model is designed to describe the phenomena of tweets driving or attracting video views including factors such as network effect (such as views beget views), system memory (such as interests in news videos wane within a week, but interests on music videos sustain for several years), and video quality (including factors such as sensitivity to tweet mentions and inherent attractiveness). We estimate the number of views per tweet via the HIP model rather than simply dividing the number of views by the number of tweets, because the model takes into account future views that are expected to unfold after the 120-day cut-off. The same reason motivates the name viral *potential* rather than viral score. One limitation of HIP and viral potential is the notion of an “average” tweet being implemented by marginalizing over power-law distributed number of followers. This means network size variations across different topics are not taken into account. A self-contained summary of HIP and viral potential computation is included below. The viral potential is a positive number, since views cannot be negative (i.e., no video loses views). However, it can take values less than one, corresponding to the effect of tweets being damped (one view per several tweets) rather than amplified, as often seen in videos linked by spam tweets.

Hawkes Intensity Process (HIP) (Rizoiu et al. 2017) extends the well-known Hawkes (self-exciting) process (Hawkes 1971) to describe the volume of activities within a fixed time interval (e.g. daily). This is done by taking expectations over the stochastic event history. Specifically, this model describes a self-exciting phenomenon that is commonly observed in online social network (Rizoiu and Xie 2017). It models the target quantity $x[t]$ as a self-consistent equation into three parts: the unobserved external influence, the effects of external promotions, and the influence from historical events. Formally, it can be written as

$$x[t] = \gamma \mathbb{1}[t = 0] + \eta \mathbb{1}[t > 0] + \alpha s[t] + C \sum_{\tau=1}^t x[\tau](\tau + c)^{-(1+\theta)} \quad (4)$$

The first two terms represent unobserved external influences. γ and η model the strengths of an initial impulse and a constant background rate, respectively. In the middle component, α is the sensitivity to external promotion, $s[t]$ is the volume of promotion, and $\alpha s(t)$ is the instantaneous response to promotion. In the last component, θ is the exponent of a power-law memory kernel $(\tau + c)^{-(1+\theta)}$. c is a nuisance parameter for keeping the kernel bounded, and C accounts for the latent content quality. Overall, this last component models the impact over its own event history $x[\tau]$ for $\tau=1 : t-1$.

In our case, $x[t]$ is the time series of daily views $x_v[t]$ on YouTube. $s[t]$ is the daily number of tweets on Twitter. The parameter set $\{\gamma, \eta, \alpha, C, c, \theta\}$ is estimated from the first 90-day interval of each video using the constrained L-BFGS algorithm in `SciPy` Python package.⁵

⁵<https://github.com/andrei-rizoiu/hip-popularity>

F View and Tweets over time: additional plots

Figure 10 and Figure 11 visualize the attention accumulation for Gun Control and BLM topic.

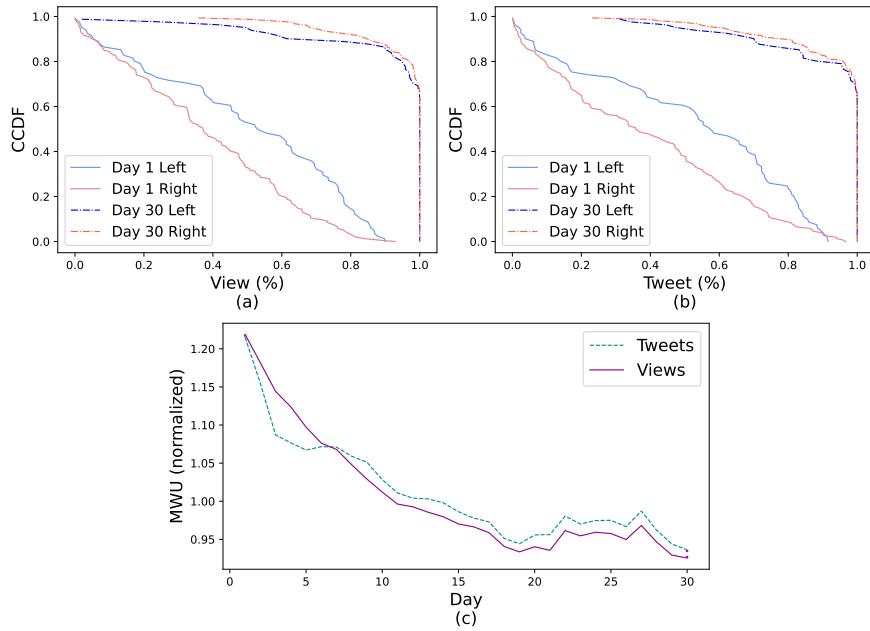


Figure 10: Comparing view and tweet accumulations of left- and right-leaning videos in Gun Control. (a) and (b) show CCDF of views and tweets accumulated for Day 1 and Day 30, respectively. As seen in (c), the differences between left- and right-leaning videos are larger than the differences of Abortion videos, but the differences between accumulation of views and tweets are smaller than Abortion videos.

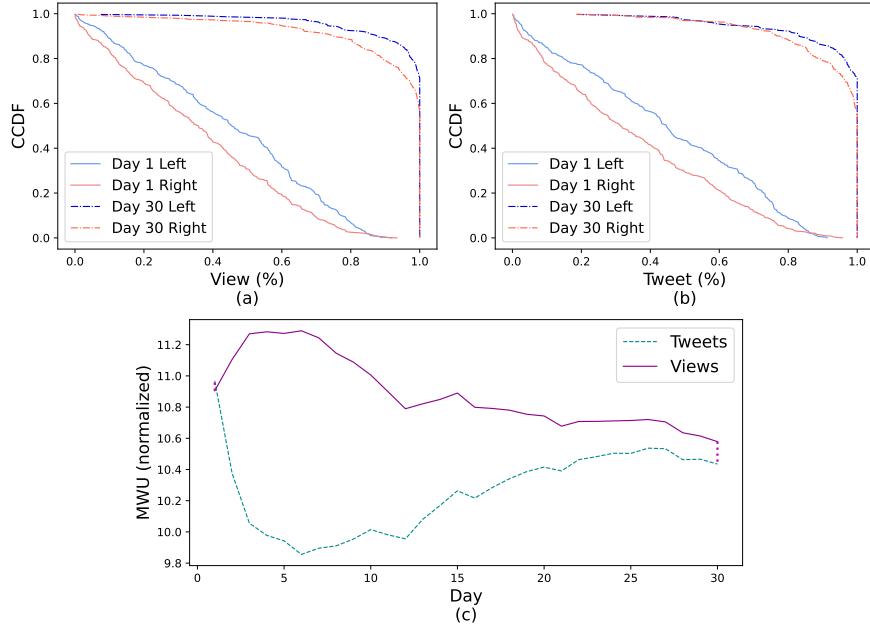
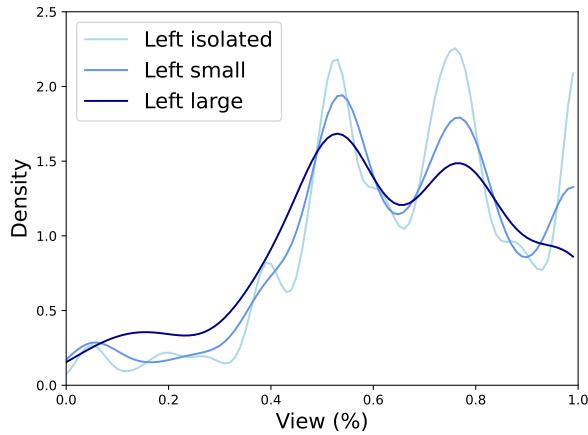


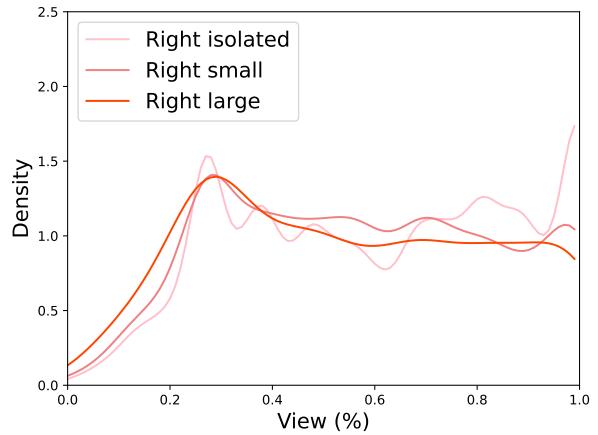
Figure 11: Comparing view and tweet accumulations of left- and right-leaning videos in BLM. (a) and (b) show CCDF of views and tweets accumulated for Day 1 and Day 30, respectively. (c) shows that the differences between left- and right-leaning videos are the largest among all topics.

G Tweet cascades: additional plots

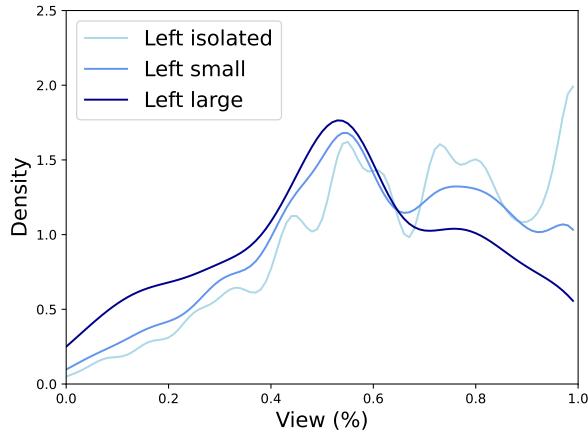
Figure 12 visualizes attention distribution of the isolated, small, and large cascades for Gun Control and BLM topic.



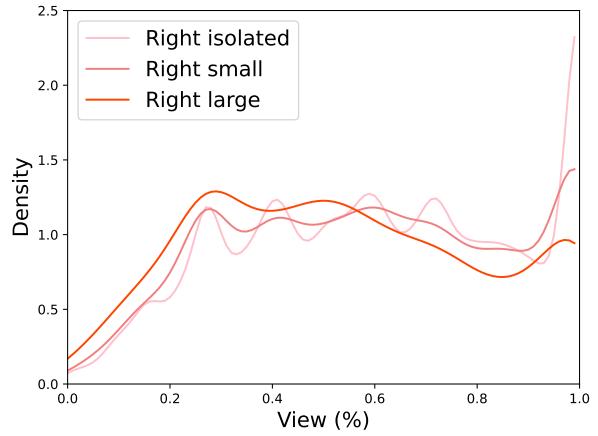
(a) Tweet cascades in Gun Control left-leaning videos.



(b) Tweet cascades in Gun Control right-leaning videos.



(c) Tweet cascades in BLM left-leaning videos.



(d) Tweet cascades in BLM right-leaning videos.

Figure 12: Cascade start times measured at view accumulation process. Table 2 reports that retweet cascades for left-leaning videos start significantly later in view accumulation process for all topics.

H Additional metrics on early adopter networks

We measured additional metrics on early-adopter networks, the comparative findings are summarized in Table 2.

Network density is the fraction of existing edges among all possible pair-wise edges. It quantifies the connectivity of the follower network. A higher value indicates the early adopters are more cohesively followed among themselves.

Maximum indegree is the maximum indegree value in the network. It provides information about the most influential user in the network.

Global efficiency is computed as the number of closed triplets over the total number of triplets. It measures the average efficiency over all pairs of distinct users. Higher global efficiency reflects higher capacity to diffuse information over the network.

Gini coefficient of betweenness centrality measures the dispersion in fraction of all shortest paths in the network that pass through a given user. Higher coefficient implies that a few early adopters are on the shortest path for most node-pairs, but the rest of the early adopters are not.

I Intersecting relative engagement with view and tweet count

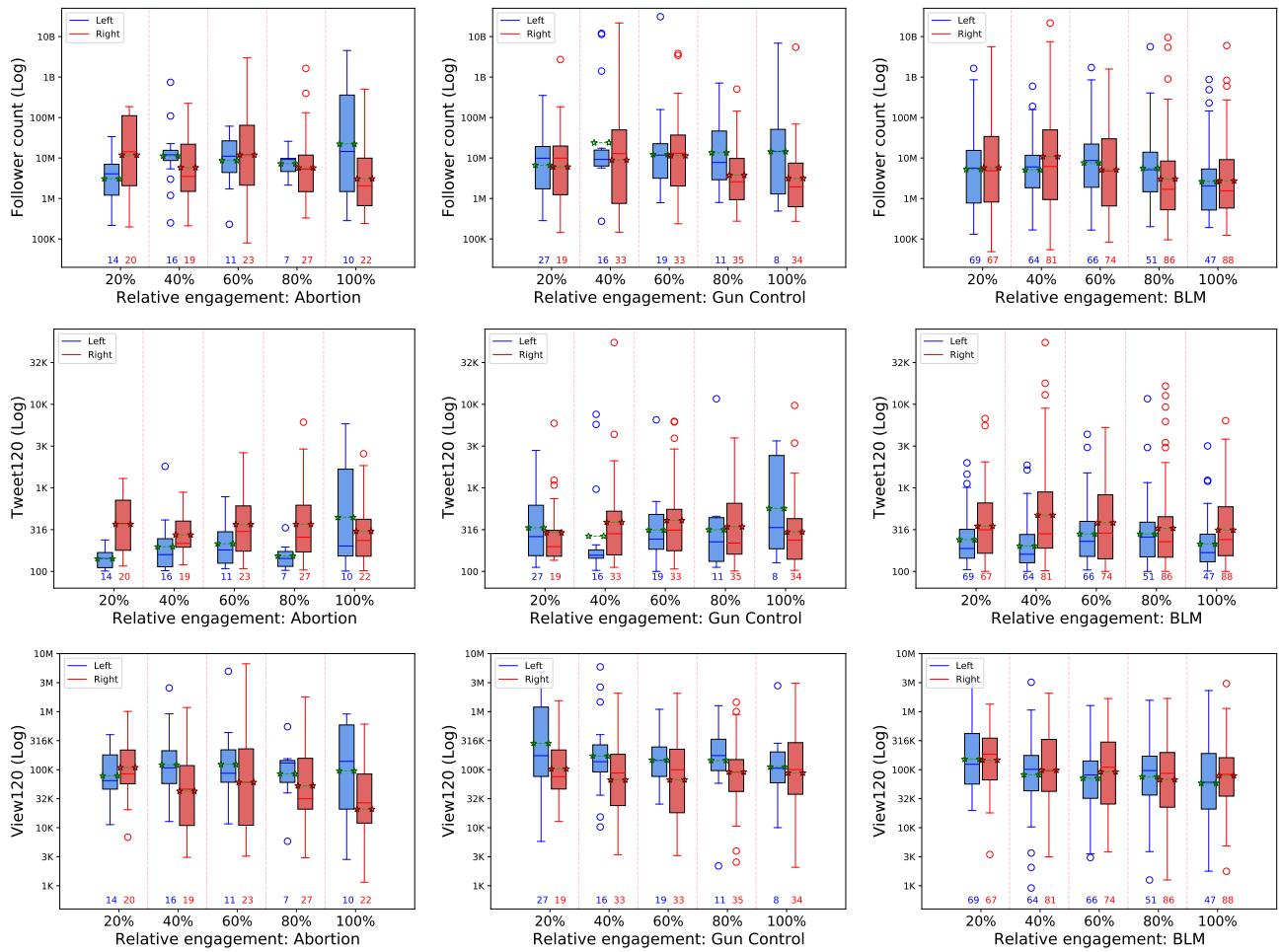


Figure 13: (**Rows**) Boxplots of three attention metrics – number of Twitter followers, the number of tweets, and the number of video views at 120 days – broken down by the relative engagement of each video in five groups of 20-percent each, (**Columns**) for each of the three topics - Abortion, Gun Control and BLM. Population means are denoted by a dashed green line ending with stars. Left-leaning videos are promoted by less tweets overall (except the most engaging group of Gun Control, and the least engaging groups for Abortion and Gun Control), but on average attracts more views (except the 40, 60 and 100 percentile group for BLM).

We link the measure of relative engagement of YouTube videos to the number of tweets, the number of followers, and the number of views. The number of tweets and views are accumulated for the first 120 days after upload. Rather than looking at the overall correlation, which is dominated by large variations, we choose to group videos according to relative engagement.

This allows us to interpret the change in Twitter user base, tweeting behavior, and YouTube attention across videos of different engagement levels.

We compute relative engagement score for each video and rank the videos in descending order of relative engagement score. Then, the videos are divided into 5 groups, each including the top 20% videos, top 20%~40% videos, top 40%~60% videos, top 60%~80% videos and top 80%~100% videos. For each group, left and right-leaning videos are shown as boxplots of *Tweet120*, *Follower count*, and *View120* (Figure 13).

As seen earlier, left-leaning videos have less number of tweets on average which can be seen with *Tweet120*. Especially, in the case of *Abortion*, the middle buckets (20%, 40% and 60%) show that left-leaning videos have significantly less tweets but have more views on average. The same pattern holds for *Gun Control*. This is consistent with the previous finding that left-leaning videos are more effectively promoted (higher viral potential) with fewer number of tweets.

We can observe that higher numbers of followers does not imply higher numbers of tweets on average as witnessed by most of middle buckets. However, we note that the top three outliers in left-leaning videos in 40% bucket of *Follower count* boxplot for *Gun Control* corresponds to the top three outliers of left-leaning videos in the same bucket of *Tweet120*.

In terms of *View120*, the average trends support findings in Section 5.1 in that left-leaning videos in *Abortion* and *Gun Control* have higher views whereas left-leaning videos in *BLM* have lower views. Although view count is considered as an important measure of video popularity in general, no noticeable correlation is shown between relative engagement and *View120*. This is expected as engagement metric and popularity metric capture very different perspective of online attention (Wu, Rizoiu, and Xie 2019).

Appendix References

- Brodersen, A.; Scellato, S.; and Wattenhofer, M. 2012. Youtube around the world: Geographic popularity of videos. In *Proceedings of the International Conference on World Wide Web*.
- Guo, M.; Hwa, R.; Lin, Y.-R.; and Chung, W.-T. 2020. Inflating topic relevance with ideology: A case study of political ideology bias in social topic detection models. In *Proceedings of the International Conference on Computational Linguistics*.
- Hawkes, A. G. 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*.
- Kulshrestha, J.; Eslami, M.; Messias, J.; Zafar, M. B.; Ghosh, S.; Gummadi, K. P.; and Karahalios, K. 2017. Quantifying search bias: Investigating sources of bias for political searches in social media. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*.
- Ledwich, M.; and Zaitsev, A. 2020. Algorithmic extremism: Examining YouTube’s rabbit hole of radicalization. *First Monday*.
- Rizoiu, M.-A.; and Xie, L. 2017. Online popularity under promotion: Viral potential, forecasting, and the economics of time. In *Proceedings of the International AAAI Conference on Web and Social Media*.
- Rizoiu, M.-A.; Xie, L.; Sanner, S.; Cebran, M.; Yu, H.; and Van Hentenryck, P. 2017. Expecting to be HIP: Hawkes intensity processes for social media popularity. In *Proceedings of the International Conference on World Wide Web*.
- Serrano, M. Á.; Boguná, M.; and Vespignani, A. 2009. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*.
- Stewart, L. G.; Arif, A.; Nied, A. C.; Spiro, E. S.; and Starbird, K. 2017. Drawing the lines of contention: Networked frame contests within# BlackLivesMatter discourse. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*.
- Wu, S.; Rizoiu, M.-A.; and Xie, L. 2019. Estimating attention flow in online video networks. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*.
- Yan, M.; Lin, Y.-R.; Hwa, R.; Ertugrul, A. M.; Guo, M.; and Chung, W.-T. 2020. MimicProp: Learning to incorporate lexicon knowledge into distributed word representation for social media analysis. In *Proceedings of the International AAAI Conference on Web and Social Media*.
- Zhou, D.; Bousquet, O.; Lal, T.; Weston, J.; and Schölkopf, B. 2003. Learning with local and global consistency. *Advances in Neural Information Processing Systems*.