User-attempted Algorithm Control on TikTok: Disability Awareness and Blackout Day

Ekaterina Fedorova

ECON 191 - Topics in Economic Research

University of California, Berkeley

Abstract

This paper examines attempts to exert control over the TikTok application's curation algorithm in order to signal-boost content from the disabled community. Using a difference in differences model comparing the disabled community to other identity/interest communities before and after Disability Awareness and Blackout Day, it finds that the blackout day has a statistically insignificant impact on trends of viewership relative to other comparable groups. These results indicate evidence of a potential failure in the attempted signal-boost of the Dec 3rd Disability Awareness Day.

1 Introduction

For many people, particularly in 2020, social media forms the basis of a very significant amount of their interaction and information consumption. As many of the platforms that are currently popular have existed for several years, most, especially Twitter, have a fairly high amount of literature analyzing a wide variety of empirical questions as well as presenting important descriptive statistics of various communities and user bases. The newest of these platforms, TikTok, however, has very little existing literature despite being the second most downloaded app of 2019 [1] and more recently, the most downloaded app of 2020 [2]. TikTok is a popular, mainly mobile, application owned by company ByteDance on which users can post up to 60 second videos. These videos can be edited prior to posting within the app, allowing users to add filters, music, text, stickers, among other effects. Each video is posted with a description (though one could presumably leave this blank) where creators can choose to include text, mentions to other users, and hashtags. As with most currently popular social media, other users may then choose to interact with the content by liking, commenting, or sharing. Unique to TikTok, there are also duet and stitch functions which allow users to create new tiktoks in response to others. These functions can serve to increase content engagement. Although stitching is still a very new function on TikTok, a previous study has previously examined the duet function [6] and found duetting was one interaction tactic used between Republican and Democrat communities on the platform.

Community formation on TikTok is an oft referenced aspect of the app amongst users and social media reporting focused websites. Within TikTok user circles and in content on the platform itself, these communities are usually called "sides of TikTok" and while these sides are not exclusive (i.e., you can be on many sides at once), phrases like "I'm on the frog side of TikTok," or even more simply, "I'm on Frog TikTok," are said both in real-life and within content on the app. One can find many articles or quizzes online to determine which side of TikTok one should be on. To be on a particular side of TikTok, depending on the nature of said side, could imbue users with a sense of belonging and for this reason I compare them to communities. Because sides of TikTok come about as a result of watching similar content, it can be presumed that said content is likely watched for some commonality. Additionally, many sides of TikTok are identity-based (LGBTQ TikTok, Disabled Tiktok, Teachers of TikTok, etc) and are essentially extensions of these communities, but within the application specifically. In an actual sense, to be on a particular side of the platform means that the user is frequently recommended tiktoks of a particular type. Essentially, to be on a side of TikTok is to be in a group of users that the algorithm deems likely to want to see and or engage

[1] https://sensortower.com/blog/tiktok-revenue-downloads-2019

[2] https://www.socialmediatoday.com/news/tiktok-was-the-most-downloaded-app-in-2020-according-to-new-data-from-app/591910/

with said side. One of TikTok's key diverging characteristics (as compared to other popular social media platforms) is its particular curation algorithm. Upon opening the application, users are directed immediately to the "For You Page", a scrollable list of automatically playing tiktoks which TikTok themselves dub, "A personalized video feed based on what you watch, like, and share." [3] This algorithmic opacity is characteristic of social media platforms as has been well-documented on other sites [2,3], however unlike most (if not all) other popular social media platforms, human-algorithm interaction is seemingly extremely high. As shown in a previous study, hashtags referencing the For You Page (#foryoupage, #foryou, etc) are among the most popular tags on the app. [6] The extent of their use (in particular as an attempt to exert control over the algorithm) may imply a level of algorithmic awareness that is not present on other similar platforms. The very nature of the sides of TikTok indicates user-algorithm consciousness as many articles exist on how to "infiltrate the different sides of TikTok" [4] which cite methods specifically on how to influence the For You Page (app-based language for the curation algorithm) to show one more content they want to see. Despite the ability of users to only see content they follow (as there is a "Following" scroll); it is the For You Page that clearly receives far more attention.

Human-algorithm interaction on TikTok, as it relates to an exertion of control, could be broadly split into an individual and community level. On an individual level, one may try to control their footprint in order to influence what the algorithm shows them (liking, sharing, commenting, searching certain tags, etc.). An individual might also attempt to influence the algorithm to show their content more frequently to others. The main function of tags like #fyp that directly reference curation is in attempt to achieve algorithmic favorability. On the community level, certain groups and their allies often actively signal-boost content of a particular community in order for it to become algorithmically favored. This is attempted in a multitude of ways such as mass commenting (words as simple as "boost" or others completely unrelated to the video's content) or Blackout/Awareness Days as is the focus of this paper.

1.1 Motivation

The aim of this paper is to look at human algorithm interaction on TikTok (specifically exertion of control) as it manifests within communities rather than on the previously described individual level. Because the curation algorithm is such a central part of most users' experience on the platform, what types of videos are recommended have often come under heavy scrutiny. It is documented that internet algorithms of many types are often subject to biases that tend to hurt underserved groups and minorities. [1] Naturally, TikTok's algorithms are no exception. In 2019, TikTok came under fire for actively suppressing disabled

[3]https://apps.apple.com/us/app/id835599320
[4]https://www.hercampus.com/school/susqu/how-infiltrate-different-sides-tiktok

users' (among others) videos [5] by having moderators prevent content from being otherwise treated normally by the curation algorithm. This choice was justified as an attempted to prevent bullying, but anger at this bias within and on behalf of disabled and other communities continued. Outside of the official confirmation of the moderation issues in 2019, anecdotal evidence cited of algorithmic bias (there is still a lack of research in this sphere) between users of the app can be seen quite frequently. Bias on the algorithmic scale obviously tends to be against a community level rather than an individual and the desire to counter it is one driver of entire communities attempting to exert force on their own algorithmic favorability. A major and arguably most discussed way of doing so is by means of Blackout and or Awareness days. Usually choosing a day that is meaningful to the community outside of the internet (ex: Autism Awareness Day), TikTok users will attempt to promote this day as one of uplifting said community's voices. The goal is to limit posting outside of the community (hence the term blackout) while signal-boosting (liking, commenting, following, viewing) community content. Information of such days are usually spread through videos created prior to the date, though comments, and often elsewhere on the internet. Such blackout days have been held (to varying degrees of success) in support of Black creators, Jewish creators, Indigenous creators, Disabled creators, and others.

On Dec 3rd, International Day of Persons with Disabilities, a Disability Awareness Day was held by many users of the disabled communities of TikTok. [6] This day was meant to boost disabled creators' voices and was the subject to prior videos by creators as well as comments on other videos (both prior and during the blackout as a request to respect the day). However, was this attempt at algorithm control successful at boosting creators? Using data of all Tiktoks created in the 2 months prior and in the week following (unfortunately, due to time constraints these pre and post periods are uneven), I examine a difference in differences regression across time periods comparing tiktoks of the disabled community to that of another identity community with similar trends prior to Disability Awareness Day, the LGBTQ community. The analysis of this data indicates that there are statistically insignificant interactions of the treatment (Disability Awareness Day) and the status of a tiktok being one within the disabled community.

The sections of the rest of this paper will proceed as follows: Section 2 serves a literature review of the previously conducted studies of relevant work, Section 3 is a summary of my collected data and empirical strategy, Section 4 is a discussion and report of results, Section 5 and 6 concludes my study of the topic, while finally Section 7 presents all tables and figures referenced.

[5] https://www.bbc.com/news/technology-50645345

[6] https://www.allure.com/story/international-day-of-persons-with-disabilities-celebrate-support

2 Literature Review

As discussed in the earlier section of this paper, there are extremely few studies done on the platform of TikTok specifically. In conducting this literature review, I was able to find only one, published earlier this year. Serrano, Papakyriakopoulos, and Hegelich examine the political discourse present on the application using similar methods to those done with other platforms, namely Twitter. [6] Collecting TikToks by crawling #democrat and #republican hashtags, the team analyzes key differences in topics discussed, engagement, hashtags used, and inferred user characteristics (from facial recognition done on video). They find that while democratic users are generally younger than their republican counterparts, the majority of users in the dataset in general are below 40 years old. In terms of content engagement, while republicans are more likely to create duets to content within the community, democratic users did more duets with users of a different ideology. Although their choice of specific communities differs significantly, the data of this paper is similarly gathered across hashtags in hopes of representing selected communities. Both their data collection and subsequent ethical standards have significantly influenced methodology of this study.

There exists a broad range of literature on the subject of human-algorithm interaction across a variety of other platforms. Of these, most relevant to this paper is a study from Burrell, Kahn, Jonas, and Griffin considering users' understanding of algorithms on Twitter [3]. The authors aim to examine what are the broad reasons why users may reference the twitter algorithm directly in their tweets in order to better understand where users see themselves in relation to this algorithm.

The team constructed a data set by taking tweets from the Twitter API that use the phrase "Twitter algorithm" (and all forms of capitalization) and further cleaned this data by omitting deleted, privated, and otherwise inaccessible tweets. Of the cleaned dataset, each of the tweets were read in full and viewed in context, then subsequently coded by 2 independent coders. The final 10 categories fell within 4 meta-categories, tweets seeking or offering explanation, tweets complaining, tweets managing a relationship, and tweets exerting control over the platform algorithms. The bulk of the analysis of the paper goes on to look at the differences in proportions of tweets within each code and the implied purposes of each category as well as whether these functions can be extended to users' understanding of the twitter algorithm of as a whole. With a special focus on control because of the large part of tweets it makes up, the paper concludes with an examination of the different methods of "gaming" the algorithm. A significant portion of my paper's language surrounding user-algorithmic interactions such as the exertion of control and gaming come from an application of this study, but transferred into the context of TikTok.

Another examination of users' algorithmic interaction on similar platforms comes from Wu, Pedersen, and Salehi who examine the crafting of algorithmic personas on Youtube. [2] Unlike on Twitter and even TikTok, where the line of creator and content-consumer is fairly easily blended into one idea of a "user", Youtube has a far more defined split between those who simply watch content and those who create it. Focusing on creators, the paper looks to define how they make sense of the algorithm. Data was gathered through the use of wiki surveys (distributed through youtuber online communities) and direct interviews asking specifically about how hobbyist youtubers see the algorithm's curation. Following participants' ranking (and adding) of personas, analysis conducted on the wiki surveys through All Our Ideas concluded that most important were the personas of Agent, Gatekeeper, and Drug Dealer, Agent being the promotive side of algorithm, Gatekeeper being a representation of that which determines whether content will get views, and Drug Dealer being the curation's choices in terms of creating an addiction to the platform.

Much like the previously referenced paper, this paper also seeks to categorize the different functions users see in the curation algorithm, but importantly this study works with human reaction data rather than statistics directly from the platform itself. Although this allows more insight into the psychological view users have of the algorithm, this is simply not scalable for platforms like Twitter or TikTok where "content-creator" is a far looser definition. Furthermore, unlike other platforms, Youtube does not have an easily accessible hashtag-based system, the presence of which on TikTok and Twitter makes exerting control on the algorithm easier.

Looking at multi-platform human-algorithm interaction, Jonas and Burrell examine in a different paper the consequences of algorithmic bias by region on more general internet users.[1]. Hoping to create a fairer system, the paper seeks to understand the impact of "good" and "bad" use-pattern classification on those in regions which has been historically discriminated against and find a major need for a reconsideration of existing security systems to remove such industry-wide bias. The authors use a qualitative approach by conducting both participant observations and in-depth interviews. Industry/community conferences were attended in order to understand the underlying rationale of the existing system. Naturally, due to this methodology, there are not significant statistical conclusions one can take from this analysis aside from the soft confirmation of wide-spread biased regional blocking on online shopping and dating sites (the particular case studies of the paper).

While the question my paper hopes to answer does not concern regional algorithmic classification, what is "promotable" and what is not is essentially at the crux of communities attempting to promote themselves by getting around the algorithm on TikTok. In creating awareness or blackout events, communities who do them with some exertion of control hope to see their content classified as more promotable by the curation algorithm.

Finally, there is a myriad of studies examining community formation across Twitter and Reddit. In 2011, one of the earliest studies of Twitter hashtag communities, Bruns and Burgess examined the use of the tags in the formation of Ad Hoc publics. [5] While the TikTok platform differs significantly in its use of hashtags (in that interaction with them carries significant weight for a user's subsequent FYP recommendations), this paper lays a basic foundation for the considering of hashtags on Twitter as modes of community building generally. Wood, Glover, and Buitelaar investigate the possibility of inference of community based on patterns of speech. [4] For their Twitter data, the team draws on posts under the pro-anorexia community of the platform. In order to do so, 2 seed tags were used from which other popular tags within posts were gleaned. All tags were manually examined in order to determine their viability leading to some tags being flagged or removed. This process serves as a basis for the creation of my community data set which follows a similar process however for the disability and lgbtq communities respectively.

All in all, there is a distinct lack in the literature of analysis specifically of TikTok data compared to other social platforms (though this is likely due to novelty rather than a lack of data. Additionally, no literature I have found addresses the exertion of control over algorithms on a community level in the form of an event such as an awareness day. While, analysis has been conducted on individuals and their reactions/attempts to game the algorithm, it might be expected that such interactions would not hold on the community level.

3 Data and Empirical Strategy

3.1 Data Collection

Data used in this paper was collected using an unofficial TikTok API by David Teather [7]. Unfortunately, due to the nature of the TikTok API, methods by which data could be collected were fairly limited in scope. While it seems that there are a nearly unlimited number of calls for data possible when collecting by hashtag, it is nearly impossible to do so by user. Prior to such data collection issues, the original goal of this paper was to consider data by users and look at their videos through time, however even with a set list of users to collect, beyond a maximum of approximately 65 users this did not work.

Instead, this paper collects all tiktoks within hashtags determined to be representative of the disability community on the app as well as those of the general lgbtq community (i.e.: not inclusive of tiktoks only tagged of one particular sexuality or gender identity). In order to get a selection of tiktoks representative of those within the disability community as a whole, I started with the most popular hashtag referencing specifically the disabled side of tiktok #disabilitytiktok. As of the 16th of December 2020, this hashtag had 198.2 Million views. On the platform it is fairly common to tag tiktoks that are intended for some imagined audience of a given community to be #—tiktok. Due to the nature of these community tags,

[7] https://github.com/davidteather/TikTok-Api

the vast majority are videos self-identifying as a member of the chosen community or looking to directly engage with said community. The next most popular tag directly referencing a disabled side of tiktok was #disabledtiktok with 115.1 Million views. (See Appendix for views of each hashtag used) Beyond these, I found that any next form of the word disability combined with the word tiktok (or simply tok, as is sometimes popular) was not of a similar or greater popularity. The only tags of greater popularity alluding to a community around disability were #disability and #disabled. Although these do not include the word TikTok as a direct allusion to the disability side of the platform, upon looking through content in this tag and considering potential reasons for tagging this (that the creator, subject, or viewer would be disabled), I scraped all the content from these 2 tags as well. The final collected dataset (before cleaning) comprised of all videos available from #disabilitytiktok, #disabledtiktok, #disabled, and #disability.

The data collection process of tags for the lgbtq community was also fairly similar, if not easier because the use #—tiktok posed no issue. In this case #lgbtqtiktok worked effectively as the "seed" hashtag from where I found the most viewed lgbtq tiktok community tags were #lgbtqtiktok, #lgbttiktok, #queertiktok, and #gaytiktok. Although traditionally #gaytiktok might imply a specific sexuality, in the case of tiktok, "Gay TikTok" is one of the most famous sides of TikTok and is used to refer to far more than exclusively a homosexual community. A review of the tag manually yielded the same conclusions suggesting that all the videos available across these 4 tags could reasonably be from the same community.

Although I would have liked to collect data more randomly (particularly by randomly sampling users), simply collecting data from individuals proved to be quite tenuous though the API. On top of this issue, random selection of either user or tiktok ID was not reasonable as users could only be collected by username, while the method by which tiktok IDs were assigned to each video was too opaque in order to replicate for randomness.

3.2 Data Limitations

All collected data (across both communities) was scraped on December 16th, 2020. This is important to note because, particularly for older tiktoks, this means that all viewership and engagement data is as of that date. Both tiktoks from Oct 12th, 2020 and Sept 15th, 2020, for example, would have their stats as of Dec 16th in this dataset. Additionally, intuitively, one may think it more likely that older tiktoks may be no longer public, meaning a data set may be "more incomplete" when it comes to tiktoks from 5 months ago versus tiktoks from 3 days ago. Naturally this timing of scraping creates some issues (discussed at length later in the paper), however this would have been an unfortunate problem of any collection method and is essentially unavoidable so all analysis of this data in this study will keep in mind these natural limitations.

Furthermore, as explained in the following paragraph, analysis of this data will be only from a fairly small select timeframe which should limit some of the issues in comparing data from drastically different time periods.

It is also important to note that it is not entirely clear if videos scraped from each hashtag encompass fully all videos in a given hashtag for all time or it is some other case. Using #disabled as an example, this uncleaned dataset of all videos I could scrape from the tag in the API was 4,324. In scraping data, a problem I had until I made sure to scrape the maximum number of tiktoks for every hashtag, was that the method of the API seemed to give me more popular (i.e. more viewed) tiktoks first. Once I was able to scrape all within the hashtag this was no longer a problem, but perhaps "all" in this case is not from the creation of the app, but rather all within a certain time limit. The oldest tiktok in this raw data set was from approximately September 2018 while its most recent was from the same day of download (December 16th) and the set overall included many tiktoks just recently created still at 10 views or under. Based on the breadth of more recent data and its inclusion of even little-viewed videos, for the remainder of this paper I will presume the data I have is "complete". As an additional precaution, for all analysis, I used Tiktoks exclusively from October 1st 2020 onwards. Based on all observed facts my dataset should include all public tiktoks created under each hashtag within the selected date range.

3.3 Data Summary

Upon collection, fairly extensive cleaning of data was required. In order to limit the effect of timed collection explained above, I kept tiktoks only from Oct 1st through Dec 13th. Although Dec 13th would have been fairly shortly before collection, I wanted to include as many post-treatment data points as possible. The standard time of creation of each tiktok was set in seconds so I binned dates to be periods of two days (48 hours). Within the cleaned data set, period 1 consists of Oct 1st 12amPST through Oct 2nd 11:59pmPST and so on. The finalized timestamps range from period 1 to period 37 where period 1 is as noted and period 37 is Dec 12th 12amPST through Dec 13th 11:59pmPST. As Disability Awareness Day was Dec 3rd and dates used here are California time, I took Dec 2nd through Dec 3rd to represent the moment of treatment. Due to the nature of time zones, this is not an entirely perfect representation of the day, however as the uncleaned dataset was in seconds and my mapping to Oct 1st is unlikely to be exactly correct second for second in the first place, it should be unlikely that it throws off the analysis so significantly. Disability Awareness day is period 32 within the data. Naturally, in order to DiD, treatment dummies were added as well as interaction terms.

The finalized and cleaned data set includes username, video length, count of likes, count of shares, count of comments, count of plays, timestamp (periods of 48 hours), treatment dummy variable, and interaction. Username represents the unique id of each user and was kept for user fixed effects in place of the more ideal user-based data, video length represents the length of a given tiktok ranging from 0 to 60s as per limitations of the app, various counts of user engagement can be used to represent the "success" of a tiktok, and finally both the treatment dummy variable as well as the interaction were added in order to complete the chosen empirical method described in the next section. See Tables 1 and 2 for more extensive summary statistics including means, standard deviations, and etc.

3.4 Empirical Strategy

Assuming that if Disability Awareness Day were effective in signal-boosting disabled creators generally and, by extension, the disabled community as a whole, one would see consistent (lasting) effects to community tiktoks' statistics, namely views, I use this as a treatment with periods prior to period 32 being those before the treatment. In theory, it could the case that Disability Awareness Day is not a lasting treatment, but this is outside the scope of my question which focuses specifically on whether there are lasting effects. From a human perspective, an awareness day is only truly beneficial if people continue being aware. Furthermore, one might expect that one at least stays "aware" for some time following such an awareness day. I assume most disabled activists themselves would also picture a "signal-boost" to not be exclusively confined to the single assigned day so for this reason I take treatment to be continuous following awareness day. Further extensions lightening this assumption would be interesting to explore as well.

In order to make a decent comparison for the basic DiD model I considered a variety of potential control groups to compare with the treated community. TikTok is host to many sides as has been extensively discussed and I first considered potentially using an interest-based community, but as a disability is an identity characteristic rather than simple hobby, interest, or job, this did not seem fitting. As a result, I settled on considering the LGBTQ community as a similarly identity-based community that has an internal commonality that is not a choice for community members. In order make sure this comparison would be valid outside of this more theoretical justification, I compared the trends of these 2 communities prior to treatment (periods 1 though 31). See Figure 1. In order to track these trends, I grouped the datasets by period and used the average view count of that period as that period's y-value. The red dashed lined on period 32 represents the beginning of treatment (Disability Awareness Day). Figure 1 indicates that prior to period 32 these 2 communities do not have the same exact patterns, but if we consider certain inconsistencies to be the result of white noise, it seems reasonable to presume the time trend of these graphs is relatively similar. Ideally, I would have had more data in order to compare these trends more thoroughly (as outliers in each particular period clearly cause quite a bit of spiking from day-to-day), however as already detailed, most possible data collection was unfortunately fairly limited.

Taking view count as the dependent variable, time fixed effects, a disabled community dummy variable (as there is no need for fixed effects considering I look at only 2 communities) and an interaction term, I used the following basic regression model:

$$views = \beta_1 * treatment + \beta_2 * Interaction + TimeFE \tag{1}$$

In addition to this very simple model, as the ideal situation would have been to have data by users, I also ran the model to include fixed effects by user. Across the aggregated dataset of both communities there are 4987 unique users despite the aggregated set containing a total of 15562 observations. Very clearly, consistent user trends like certain users being particularly popular and posting in the community often would greatly affect any such regression.

$$views = \beta_1 * treatment + \beta_2 * Interaction + TimeFE + UserFE \tag{2}$$

In addition to this main empirical study, although I could not easily test the viewership of the disability community to other communities without gathering more data, it was fairly simple to also take a look at other measurements of a tiktok's success within the same communities. Figures 2, 3, and 4 are graphs done in the same way as Figure 1 in order to test whether it seemed that disability TikTok shared parallel trends with LGBTQ TikTok in count of likes, comments, and shares as well. As indicated by figures 3 and 4, it seems that comments and shares are not similar enough to consider them parallel trends, at least not nearly to the extent of views and somewhat likes. A main problem with these additional measurements of success is that they are on average far lower in any given TikTok and data is more volatile (as compared to count of views). The range of both Figures 3 and 4 are orders of magnitude lower than Figure 2 and especially Figure 1. Because the trends are far less parallel, I did not conduct a DiD regression with comment count or share count, however I did do so with the count of likes. Here I used the same regression models as for views, but naturally with likes as the new dependent variable:

$$likes = \beta_1 * treatment + \beta_2 * Interaction + TimeFE \tag{3}$$

$$likes = \beta_1 * treatment + \beta_2 * Interaction + TimeFE + UserFE \tag{4}$$

Additional analysis considering different comparable communities would be very interesting to consider in the future. In particular, it would be beneficial to expand beyond an investigation of exclusively two communities so as to add more robustness to the results. As well as similar regression models, but in which one uses different measurements as the dependent variables for example engagement measured through like to comment ratio or other non-traditional methods.

4 Results and Discussion

4.1 Results

Table 3 column 1 shows the results of the first, slightly simpler regression model. Here the coefficient of the interaction is significant while the coefficient of the treatment dummy variable is not. In a vacuum (imagining the results of column 2 did not exit), these conclusions would be extremely valuable. In this case what this would be saying would be that there is a significant difference in change over time between the 2 communities, in particular that it is negative. This would essentially be the result opposite of what is hoped for by the very existence of a Disability Awareness Day. Rather than tiktoks from the disabled community experiencing an increasing trend in viewership greater than a comparable community, LGBTQ in this case, or even rather than lack of evidence for a major difference between the 2 communities' trends pre and post treatment, this kind of coefficient would imply that somehow the passing of Disability Awareness Day may have caused a dip in popularity even when relative to other groups.

Instead, however, there is the 2nd column which takes the same regression and adds in user fixed effects. As discussed in the data section, as compared the number of tiktoks themselves, the number of unique users is lower. For this reason, along with the intuitive idea that particular users would have higher viewership if they already, for example, have an established audience, it is very important to take into account these user fixed effects. As a result, the model concludes that there is no consistent long term change in viewership trends relative to comparable communities for the disability community as a result of Disability Awareness Day. Or rather, as a result of this December 3rd 2020 Disability Awareness Day.

4.2 Additional Results - Likes

Interestingly Table 4 column 1 indicates something quite different going on with likes. Here the treatment is very statistically (negatively) significant indicating a difference between the two communities pre-treatment which is not surprising because regardless of trend, at a glance it does seem from Figure 2 that the magnitude of likes within the disabled community is consistently lower than that of the LGBTQ. Additionally, the interaction is mildly significant, but all things considered, given the questionable parallel trends, this is not something I would find of note.

Upon adding in user fixed effects in Table 4 column 2, the significant treatment coefficient becomes immediately insignificant, definitely implying that the previous result was likely the result of an omitted variables bias instead of any real conclusion. The interaction decreases in magnitude slightly thus increasing its statistically significance, but again given the more tenuous parallel trends here, there is little of note to conclude.

4.3 Discussion

To elaborate on the ending note of section 4.1, one important takeaway from this particular difference in differences regression (for views) is that the external validity here is something to question before reaching more generalized conclusions about the utility of community-based awareness/blackout days. In terms of applicability to other facets of community-algorithm interaction, this study is quite limited because of the extremely small scope. It would be unreasonable to say just because of the lack of a causal link seen here between Disability Awareness Day 2020 and disability community viewership that this is something one might expect to see across the board for communities that have such days.

One problem, for example, may be a lack of knowledge about this particular event. As a TikTok user, this is an event I was aware of personally, but this in no way indicates to what degree it was truly widespread. Perhaps the issue in a lack of causal link lies not with a lack of effectiveness of such days on the algorithm overall. Another may be that blackout days are effective, but they require a scale that the disability community does not reach. This particular community is far from the largest or even most well-known identity community on the platform. Furthermore, within the month of December 2020 alone, there are at least two additional TikTok blackout days[8], it may also be the case that such movements are less effective in such saturated times. Regardless of the reason for this lack of causality link between the two, without further study it would be impossible to determine this without very significant assumptions.

A final issue lies within the study itself, or rather the issues caused by underlying data problems. As mentioned earlier in this paper, due to the nature of TikTok data collection, it is very easy to imagine ways in which success metrics (views, likes, comments, and shares) could be skewed. Data collected in one day accesses metrics of that day regardless of the date of the video's creation. A tikok from 25 days ago has likely reached stagnation in user engagement, however one from just 2 may not have. This places a large burden on the date of data assessment in cases when the data being looked at is incredibly recent. A redownload of the same tiktoks looked at in this study even on December 18th 2020 (2 days later) might

[8]https://www.distractify.com/p/tiktok-blackout

provide slightly different results. Unfortunately, this is not as easy to fix as looking at older data because due to the functionality of the API, this is usually much more difficult. Older data is far harder to access and also has the potential to have far more private content. Additional data issues faced in this study which may have affected results include general accessibility and, hand in hand, quantity problems. For the number of time periods used, the number of observations is very low and not technically the same through time. Rather than an individual user's videos through time as they go through a "treatment" this study looks at videos within a single community through time and assumes a certain consistency of these tiktoks based on their shared place in the community. This is a rather strong assumption to make.

5 Conclusion

This paper investigates the currently little-explored world of social media data specifically from TikTok. As a social media platform using algorithms in a quite a novel way, TikTok paves the way for interesting questions surrounding human-algorithm interactions specifically in the ways in which creators, viewers, and communities choose to attempt to exert their control over what is recommended and how their own content is recommended to others. As TikTok has come under fire for their algorithmic practices in the past, more communities attempt to mass signal boost through blackouts and awareness days as a way to get around systemic algorithmic bias. I collected data from two distinct communities on the application (disability TikTok and LGBTQ TikTok) and used the recent Disability Awareness Day on December 3rd 2020 in order to consider a treatment that would theoretically affect viewership and general content engagement only on disability TikTok. A difference in differences regression using the 2 communities and 37 periods of time (each of which are 48 hours) indicates that there is no causal link between change in viewership trends of disability TikTok relative to comparable-trend LGBTQ TikTok and the December 3rd Disability Awareness Day. Further research is needed to understand how such a conclusion could be generalized (if at all) to other community-based algorithmic exertions of control.

6 Ethical Considerations

In investigating this question of whether Disability Awareness Day is an effective treatment via DiD, I have looked at and collected large amounts of tiktoks in both the disability as well as the LGBTQ community. While all of these tiktoks are public and can easily be viewed by anyone the app, the aggregation and systematization of large amounts of this data, especially considering the popularity of TikTok among youth as well as vulnerability status of chosen groups, could be seen as a overstep in privacy. In order to maintain data protection, none of the identifiable data has been included in this paper or shared with others. In the Appendix readers can find communities tags I considered in writing this paper as well as their general tag viewership and date accessed.

7 Tables and Figures

Table 1: Descriptive statistics: Disability Community Oct 1st - Dec 13th

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Length of Video	8,178	29.914	19.790	0	13	53	60
# of Likes	8,178	8,185.736	53,274.150	5	74	2,442.8	1,600,000
# of Shares	8,178	210.722	2,397.029	0	1	26	128,400
# of Comments	8,178	134.684	834.208	0	5	62	31,100
# of Plays	8,178	60,229.040	352,303.300	22	429.2	18,600	11,300,000

Table 2: Descriptive statistics: LGBTQ Community Oct 1st - Dec 13th

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Length of Video	7,384	20.113	16.242	3	9	24	60
# of Likes	7,384	16,744.720	97,022.360	2	249	3,629	4,700,000
# of Shares	7,384	452.435	3,659.373	0	1	46	218,500
# of Comments	7,384	309.960	1,716.876	0	7	101	55,100
# of Plays	7,384	78,117.660	385,819.700	6	1,335.8	21,425	16,800,000

Table 3: DiD Regression (LGBTQ): Dec 3rd Disability Awareness Day as Treatment

	Dependent variable:	
	Video Plays	
	(1)	(2)
Awareness Day (treatment)	2,651.630	−2,666.322
Interaction	−50,124.260***	−35,984.750
User Fixed effects?	no	yes
Time Fixed effects?	yes	yes
Observations	14,958	14,958
R^2	0.041	0.391
Adjusted R^2	0.038	0.095
Residual Std. Error	371,682.200 (df = 14920)	360,526.400 (df = 10074)
F Statistic	16.664*** (df = 38; 14920)	1.322*** (df = 4884; 10074)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4: DiD Regression (LGBTQ): Dec 3rd Disability Awareness Day as Treatment

	Dependent variable:	
	Video Likes	
	(1)	(2)
Awareness Day (treatment)	−6,624.552***	−768.177
Interaction	−5,420.042*	−9,373.844**
User Fixed effects?	no	yes
Time Fixed effects?	yes	yes
Observations	14,958	14,958
R^2	0.030	0.402
Adjusted R^2	0.028	0.112
Residual Std. Error	78,323.570 (df = 14920)	74,841.770 (df = 10074)
F Statistic	12.181*** (df = 38; 14920)	1.387*** (df = 4884; 10074)
Note:		*p<0.1; **p<0.05; ***p<0.01

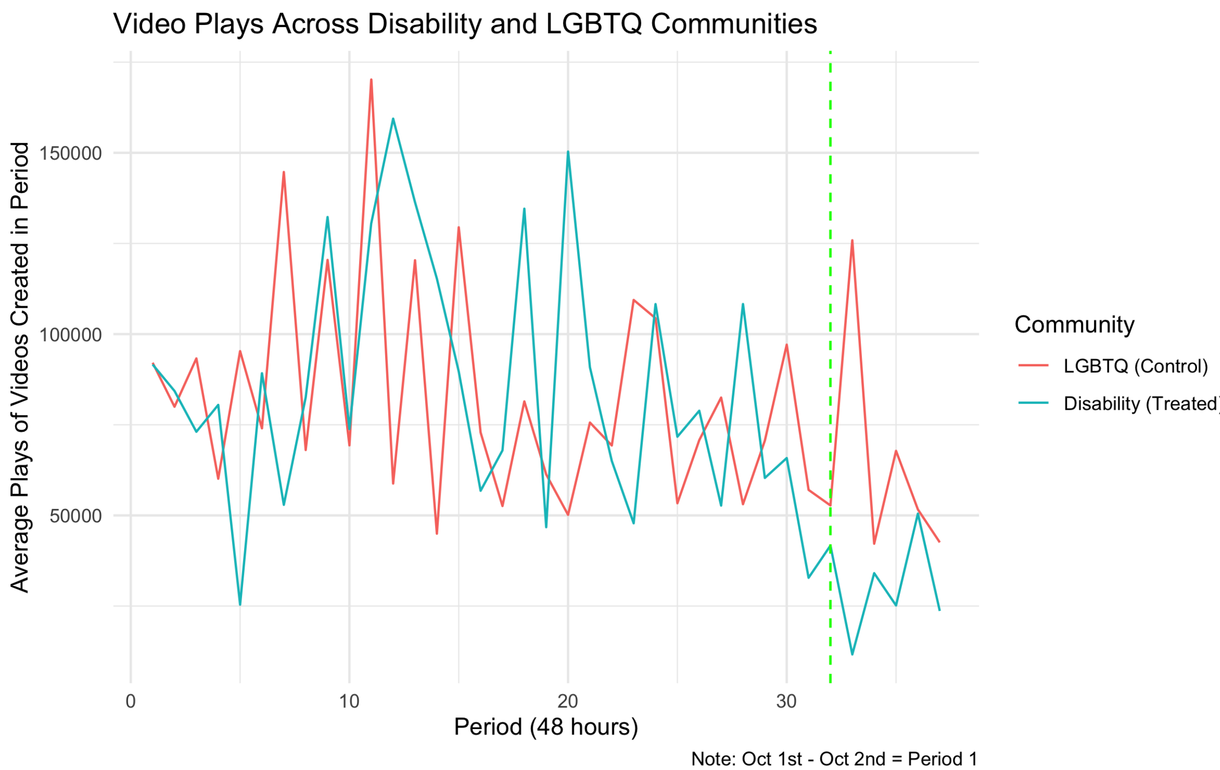

Figure 1

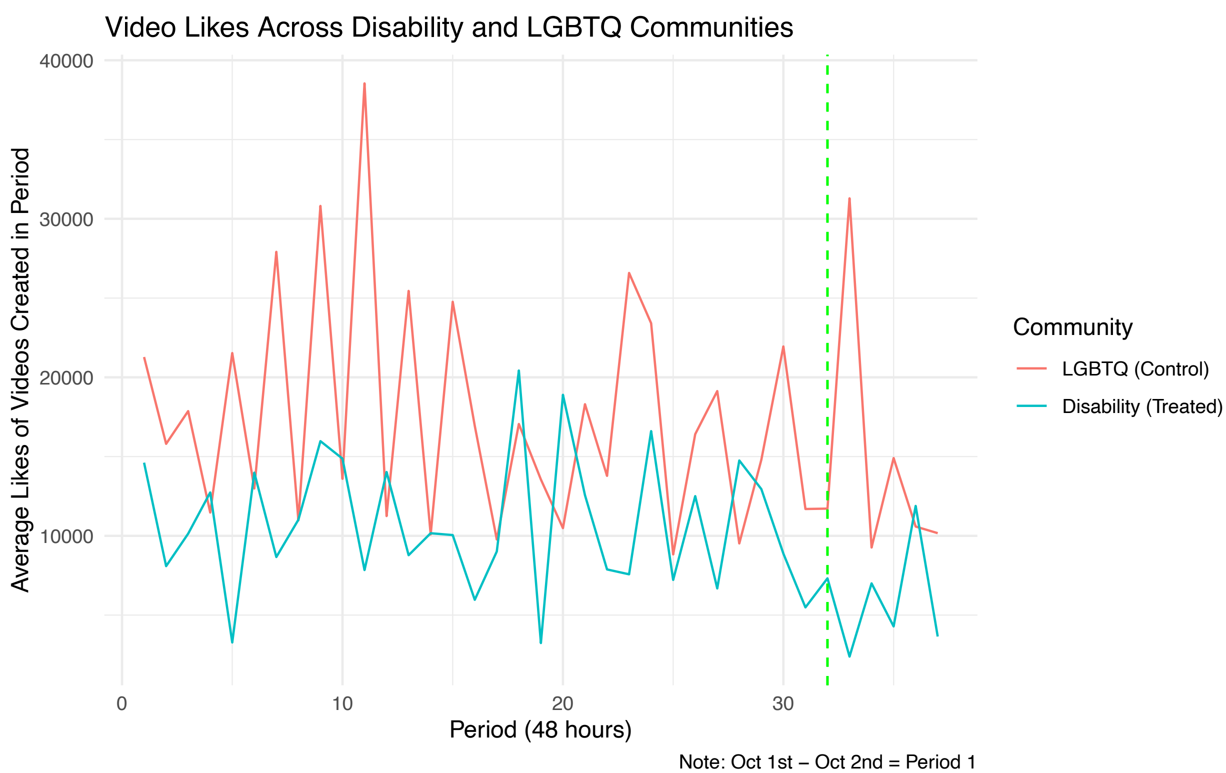

Figure 2

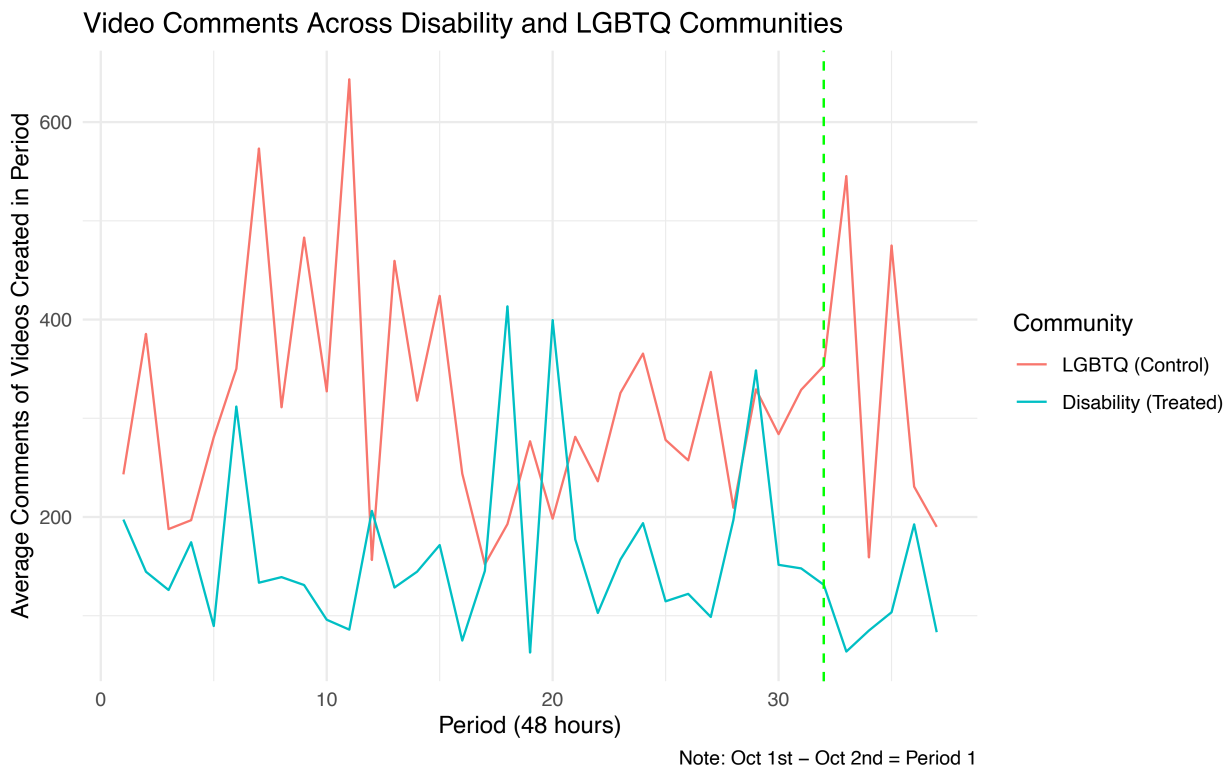

Figure 3

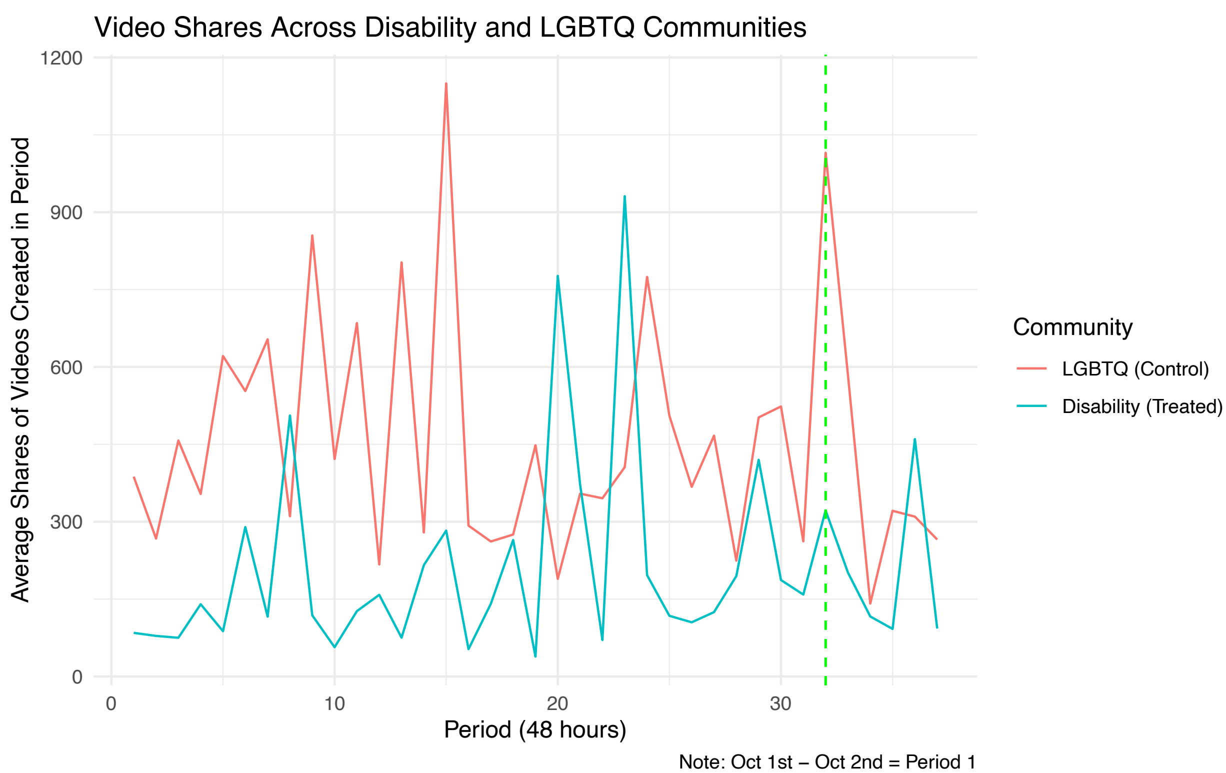

Figure 4

8 References

1. Jonas, Anne and Jenna Burrell. 2019. Friction, Snake Oil, and Weird Countries: Cybersecurity Systems Could Deepen Global Inequality through Regional Blocking. Big Data and Society, 6, 1. Jan 2019. [PDF]

2. Wu, Eva Yiwei, Emily Pedersen, Niloufar Salehi. 2019. Agent, Gatekeeper, Drug Dealer: How Content Creators Craft Algorithmic Personas. In proceedings of the ACM Computer-Supported Cooperative Work and Social Computing (CSCW) conference. Austin, TX. [PDF]

3. Burrell, Jenna, Zoe Kahn, Anne Jonas, and Daniel Griffin. 2019. When Users Control the Algorithms: Values Expressed in Practices on Twitter. In proceedings of the ACM Computer-Supported Cooperative Work and Social Computing (CSCW) conference. Austin, TX. [PDF]

4. Ian D. Wood, John Glover, and Paul Buitelaar. 2020. Community Topic Usage in Online Social Media. Trans. Soc. Comput. 3, 3, Article 14 (August 2020), 21 pages. DOI:https://doi.org/10.1145/3377870 5. Bruns, Axel and Burgess, Jean. (2011). The Use of Twitter Hashtags in the Formation of Ad Hoc Publics.

6. Juan Carlos Medina Serrano, Orestis Papakyriakopoulos, and Simon Hegelich. 2020. Dancing to the Partisan Beat: A First Analysis of Political Communication on TikTok. In ¡i¿12th ACM Conference on Web Science¡/i¿ (¡i¿WebSci '20¡/i¿). Association for Computing Machinery, New York, NY, USA, 257–266. DOI:https://doi.org/10.1145/3394231.3397916

7. Obaida Hanteer, Luca Rossi, Davide Vega D'Aurelio, and Matteo Magnani. 2018. From interaction to participation: the role of the imagined audience in social media community detection and an application to political communication on Twitter. In ¡i¿Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining¡/i¿ (¡i¿ASONAM '18¡/i¿). IEEE Press, 531–534.

8. Amin Salehi and Hasan Davulcu. 2018. Detecting antagonistic and allied communities on social media. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '18). IEEE Press, 99–106.

9 Appendix

Table 5: Tag Counts In Disability and LGBTQ Communities

	Hashtag	Views	Date
1	disabledtiktok	$115,100,000$	12/16/20
2	disabled	$816,100,000$	12/16/20
3	disability	$1,300,000,000$	12/16/20
4	disabilitytiktok	$198,200,000$	12/16/20
5	queertiktok	$71,600,000$	12/16/20
6	gaytiktok	$2,900,000,000$	12/16/20
7	lgbtqtiktok	$79,500,000$	12/16/20
8	lgbttiktok	$109,700,000$	12/16/20
9	lgbtiktok	$43,300,000$	12/16/20