

Through Thick and Thin: Reexamining r/The_Donald's Quarantine Using Network Methods

Zachary Philip-Taylor Brown

A thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Arts

University of Washington

2024

Committee:

Zack Almquist

Steve Pfaff

Program Authorized to Offer Degree:

Department of Sociology

©Copyright 2024

Zachary Philip-Taylor Brown

University of Washington

Abstract

Through Thick and Thin: Reexamining r/The_Donald's Quarantine Using Network Methods

Zachary Philip-Taylor Brown

Chair of the Supervisory Committee:

Zack Almquist

Department of Sociology

The subreddit r/The_Donald was a popular political forum that supported the Trump campaign and presidency through stimulating talking points, combating online detractors, and generating memes utilized by right-wing figures. After subreddit users continuously disregarded site rules regarding threats of violence, Reddit administrators first restricted and then banned the community in 2019 and early 2020. Previous efforts to understand how this sudden change to communication pathways led to declines in activity on r/The_Donald in the long run as well as shorter run drops in hate speech on subreddits that community members turned to after losing access to their previous home. This approach misses the importance of social ties between prominent members of the r/The_Donald community and cannot answer how enmeshed these accounts remained after the restriction. Utilizing social network analysis techniques, I analyze how the interactions between frequent r/The_Donald commenters changed after the quarantine. I find that recurring users were more densely connected in the three weeks after the quarantine than they were before it, that users posted in similar subreddits in July that they had in June, and that pre-quarantine in-degree was highly predictive of post-quarantine in-degree. In this paper I employ a network analysis framework to describe this important shift in platform activity, lending credence to the importance of

communication ties on social media sites and posing new questions about the future of online moderation and the resiliency of new forms of social politics.

Introduction

Over the past two decades Reddit has remained a stalwart in a social media landscape littered with abandoned apps and fluctuating levels of user activity. The site is home to more than 100,000 active sub-forums on virtually any topic, each with an engaged community that adds up to more than 70 million daily unique visits (Reddit, Inc 2023). Recently, Reddit has become known for initiating the formation of cohesive communities motivated around novel toxic political ideologies, from the harassment of #Gamergate targets utilizing r/KotakuInAction (Barnes 2018) to the consolidation of core QAnon beliefs on r/greatawakening (Ohlheiser 2021). Since 2016 there has been a noticeable rise in political polarization on the site, often driven by the influx of new users seeking out content that aligns with their political beliefs (Waller and Anderson 2021). One of the more influential of these online communities was r/The_Donald, a subreddit created in response to the presidential campaign announcement of Donald Trump in 2015 (Koebler 2016). The forum rapidly gained popularity, serving as a marketplace for memes influencing the policy direction of both the Trump presidential campaign and presidency while also contributing to the Pizzagate and Seth Rich conspiracy theories (Gilmour 2017). After repeated run-ins with Reddit over content that encouraged violence and harassment campaigns against elected officials, site administrators took action on June 26th, 2019, quarantining the subreddit (Robertson 2019). Similar to efforts used previously with forums that frequently violated content moderation policies, the quarantine prompted users with a warning about the content they were about to view if they attempted to access r/The_Donald, stopped showing posts in user news feeds and search requests, and limited

the ability for subreddit admins to profit from digital advertisements. This intervention gradually led to the decline of r/The_Donald, with a mirror of the site being created by a dissatisfied sect at TheDonald.win (later Patriots.win) and the eventual banning of the rump original subreddit just over a year later (Tiffany 2020).

The quarantine of r/The_Donald presents an interesting case for examining the impact of external intervention on the durability of social ties and the maintenance of online political communities. Reddit moderation efforts were a sudden change event that likely had a direct impact on community activity on r/The_Donald. Previous examinations of this event have focused on the decline in hate speech associated with user exodus after moderation (Trujillo and Cresci 2022) or user-level impacts on posting behavior and account continuation (Trujillo and Cresci 2023), but no attempts have been made to date at examining the cohesiveness of the social community of r/The_Donald before and after the quarantine. Past work has determined that distinguishable social roles are visible on Reddit when examining content creation behavior of individual users in a network framework (Buntain and Golbeck 2014) and that network attributes have an association with the survivability of subreddit communities (Tsugawa and Niida 2021). Other researchers have extolled the virtue of using comment interaction networks to identify like-communities on Reddit (Kanti Baowaly, William Kibirige, and Singh 2022) and identified that similarity in posted content has led to the formation of a group identity on r/The_Donald (Gaudette et al. 2021). Given the continued importance of online social spaces in the formation of group political identities, there is an opportunity using the quarantine of r/The_Donald to examine whether the attempted site moderation was as productive at extinguishing existing network ties between community members as it was at reducing overall interactions with the forum.

This paper aims to utilize social network analysis methods to assess the impact of the quarantine of r/The_Donald on the subreddit's users. While it has been suggested through previous efforts that

walling off the community led to a reduction in overall interaction and decline in user activity, this author contends that the resilient and continued connections between influential forum actors is an alternative measure to examine the impact of the attempted moderation. I build on the existing literature by redefining the sample population to more accurately capture the entirety of the r/The_Donald community before and after the quarantine; instead of utilizing a narrowly defined set of “core users”, I gather all comments posted by users who remained active in both the pre- and post-quarantine periods. To then further empirically test the effectiveness of the quarantine, I propose the following research questions:

1. Did the attempted moderation lead to meaningful shifts in the construction of the r/The_Donald commenter network?
2. Was there a shift in the relative frequency of subreddits commented in after the quarantine for previous r/The_Donald commenters?
3. Were powerful agents in the r/The_Donald network able to maintain their influence even if user activity shifted to different areas of Reddit?

These questions will help to determine the extent to which the core network of r/The_Donald users survived the quarantine and whether users were able to maintain their prior standing in the social network even if they shifted their posting to other active subreddits. It has been suggested by previous works that the attempted moderation led to an immediate decline of the r/The_Donald community; this paper aims to contribute to this knowledge with insights from network analysis techniques and the added ability to track how actors within the community may have influenced one another.

The results suggest that a more nuanced assessment of the success of the quarantine of r/The_Donald may be necessary. While I provide no conclusive evidence that the restrictions failed,

the evidence does support the conclusion that there were limited shifts in the underlying network construction of the r/The_Donald community, that posters continued sharing comments proportionally in different subreddits and r/The_Donald at similar rates, and that influence in the pre-quarantine network was fairly predictive of influence in the post-quarantine network. Therefore, further evidence is necessary over potentially longer timeframes than the three-week period analyzed in this study to determine how the quarantine impacted the social interactions of core members of the r/The_Donald community. This paper supports the use of comment-to-comment constructed networks as a tool for understanding online social interactions and posits interesting suggestions about the rigidity of new online political formations.

Background

Political Communities on Social Media Platforms

The rise of Internet communication channels and social interactions through various web platforms over the past three decades has remained of interest to sociologists. Historical organs of social collaboration outside of the home have diminished in the modern era (Putnam 2001), with the Internet increasingly serving as a medium for interpersonal communication. Bennet (2012) provided an early example of identifying the increased personal expression that the Internet enabled. His work examined how the rise of online communities coincided with a drive towards individualism and market ideology in American culture and the pursuit of groups that aligned with one's personal beliefs as opposed to a broader higher level. Though mainly concerned with liberal social movements online organized around environmentalism or worker's rights, Bennet's work can also be utilized to examine how our ideological beliefs drive us to seek out others with similar viewpoints.

The relative influence of social media on social patterns has continued to be the subject of fierce debate. It has been surmised by some that most democratic regimes around the world still rely on centrist voters as the base of their support and that online fringe political groups, though vocal, have limited influence (Zhuravskaya, Petrova, and Enikolopov 2020). Having this bastion of support, however, does not demur more radical ideologies from gaining a following in online spaces. Research has suggested that on Facebook it is our own ideological drive, not a hidden algorithm or proprietary sorting method, that determines the type of political content that we experience (Bakshy, Messing, and Adamic 2015). In this study by Bakshy et al, there was a significant level of ideological homophily in Facebook friend networks that led to increased exposure to news with a similar political position. Kidd and McIntosh (2016) take a nuanced view on the matter arguing that both social movements require more than just conversational chatter in corporate-governed platforms to find success but that we shouldn't discount the increased ability of these tools to connect disparate individuals and to create new organized bodies. Chayko (2014) points to the disappearing boundaries between our social lives in the real world and online to stress the importance of social media tools as a way to connect populations. While some research suggests that the actions that we take online are of a different type from those we make in-person (Kruse, Norris, and Flinchum 2018), it seems clear that there is a meaningful social activity taking place online and that some of that communication is geared towards political organizing.

There remains a robust debate about how well online organizing reflects and results in offline political participation. An early examination by Bennett (2008) found that though rare, there had been early social movements online incorporating some political activism, such as a protest in the game *World of Warcraft* against virtual class disparities, but that Internet users had grown more attached to networked activism online that shared specific aspects of their political identities. This has been shown to stand in more recent studies as well: Freelon et al (2020) disentangle the

different strategies that right- and left-wing ideologues utilize to push their online activism towards the physical world. They describe how left-wing individuals are more likely to partake in offline protests, while right-wing users tend to work cooperatively with existing power systems to reframe online and offline conversations. In an attempt to determine precisely how different the online domain is from the real world, Zúñiga et al (2019) quantified social media and face-to-face social capital, finding that they were empirically distinct and exhibited different behavioral patterns. This suggests that while there may be some differences in how social organization takes place online, that there are still meaningful political interactions that transpire on the Internet and that social scientists should pay attention to the domain to capture the leading edge of social politics.

Right-Wing Organizing Online and r/The_Donald

With the added accessibility that the Internet brings to forging connections with like-minded individuals, there has been an increase in the study of fringe and radical political ideologies. An early investigation comparing neo-Nazi groups and environmentalists in online spaces by Wojcieszak (2009) found that for both groups individuals were more open to expressing their desired political positions in online spaces than with their interpersonal connections. Since then, more research has been dedicated to understanding how specifically the fringe right organize and spread their ideology online, for both ideological and material reasons. While some papers have suggested that less extreme information seekers tend to avoid entering discussions of politics online for fear of the content that they might run into or abuse that they endure (Kruse et al. 2018), others have found that social media sites open up channels of communication that haven't existed prior and encourage the flow of information (Duncombe 2019), including in radical political spheres. Wahlström and Törnberg (2021) argue that online spaces now allow radical political movements to transcend local political borders, connecting anti-immigration radicals both across

Sweden and around Europe in a way that previously would have kept anger at the town or county level. The added complexity of cross-national communication pipelines has enabled formerly disparate ideologues to connect with others of similar viewpoints through catered communities and lend a hand in the further radicalization of radical-friendly observers. Grover and Mark (2019) examined r/altright, a subreddit that rose out of the push for a more insular, darker conservatism that coincided with the candidacy and presidency of Donald Trump and found, through textual analysis, that behavior on this particular forum showed more concerning signs for potential radicalization than other similar political subreddits around the same time period. With community norms that allowed for the open embrace of white identity politics and a juxtaposition between the in-group of true believers in the alt-right cause and those who opposed them, Grover and Mark echo a claim made earlier by De Koster and Houtman (2008) that as a group's political opinions become more stigmatized offline, the more likely they are to find a home in online communities. The online activities of right-wing political groups can also have offline impacts. While ties to organized communities online may not necessarily increase the vote share of right-wing parties or candidates in large democracies (Boulianne, Koc-Michalska, and Bimber 2020), it does appear that users engaged in these spaces tend to spread polarizing news sources (Schulze 2020) and have a penchant for violent language, sometimes rallying real-world support for movements like #StopTheSteal which eventually produced the insurrection on January 6th, 2021 (Chen, Lukito, and Koo 2023).

The fringe right has utilized various online platforms to increase their reach and mold a cohesive political identity. Blee et al (2023) suggest that the election of President Barack Obama motivated a new wave of previously conservative men to engage in discussions online that could be criticized in the real world, while more recently white supremacists and neo-Nazi groups have used applications like Discord to promote outreach to the younger generation. Daniels (2018) sees the

Internet as a place where the political right as a forum to craft and pass down their ideals, forging new community icons such as “Pepe the Frog” with a unique meaning to members of the in-group. These symbols can be used to recruit others into the fold and the feelings that they invoke can be pivotal in maintaining a cohesive identity even in the relatively loose associations formed online. Latif et al (2018) relates this to Collins (2005) and the production of emotional energy, suggesting that the familiar emotions produced when interacting with visual media, music, and themes associated with a political identity can connect members even when separated by cyberspace. These elements have helped right-wing political groups gain a foothold on various social media platforms and helped to expand their influence in the offline world as well.

Reddit has not been spared the ascendance of organized right-wing politics in online spaces. *r/The_Donald* has been studied previously as one such example of an influential forum. Horta Ribeiro et al (2021) examined how users migrated away from the subreddit after the ban to both other political subreddits and through username matching to different websites altogether, comparing the *r/The_Donald* community with that of *r/Incels*, a semi-political forum for individuals distressed at their inability to find sexual partners that features some of the misogynistic and chauvinistic undertones. They found that while overall activity declined, both groups showed a cohesiveness and sense of community that suggested their continual survival into the future. Another study (Chandrasekharan et al. 2022) examined the impact of the quarantine on the relative levels of misogynistic and racist comments on user posts on *r/The_Donald*, finding that while there were fewer new users after the moderation attempt, the levels of negative posts remained relatively unchanged and there were more posts than before in *r/The_Donald*. That paper and the work of Trujillo and Cresci (2022) and (2023) lay the foundations for this study. They suggest that overall user activity on *r/The_Donald* diminished after the implementation of the quarantine and that the presence of hateful content also began to decrease. While cautious about interpreting too

much about smaller user communities on the subreddit from the larger results, they contend that the moderation was successful in limiting the number of toxic submissions and comments on Reddit and reducing the exposure of outside users to this material. The earlier piece defines a group of “core users” or individuals posting at least once a week in the 30-week pre-quarantine period that they identify, while the later analyzes these users exclusively. Both pieces mention the potential social network ties that these users have created through their interaction, but by reporting on user activity at the aggregated community level miss the opportunity to elaborate further on them. Instead, they leave open the avenue for future research that involves a more elaborate definition of a community and reincorporates users who are not as frequent posters on the subreddit.

Social Ties through Comments

An early suggestion by Stieglitz and Dang-Xuan (2013) explored different ways through which the emerging social media landscape could be analyzed by researchers, with the goal of identifying the best methods to capture the ongoing social processes that were taking place in a realm where data was easily accessible and millions of users were logging on each day. One of their ideal forms of analysis was to take a structural approach towards analysis, focusing on a broad definition of “community” and attempt to describe it, either by identifying influential actors or by utilizing network theory to attribute interactions between users as nodes and edges. This line of analysis has been used in the proceeding decade by authors such as Deen and Pan (2022), who took advantage of retweets and replies to construct a community based around conversation about the aftermath of the 2020 Presidential Election. Using this approach, they (surprisingly) found a hub of activity in support of both Joe Biden and Donald Trump originating from a news aggregator in India. Murthy (2012) framed the interaction between online users in a more sociological way, informed by

the interactional activities described by Goffmann (1981). They note that even though users on Twitter may not acknowledge the username providing the content that they are retweeting or responding to, there is still a meaningful social connection taking place and the diffusion of information from one source to another. This activity implies the embedding of language from one user into the lexicon of another; with Twitter this takes place in the direct form of a retweet literally sharing the words of the previous commenter. This relationship is more nuanced on Reddit, but a similar principle can be applied. Users on Reddit are responding to the thoughts and ideas that others have previously shared, and responding suggests that the replier is engaging with the articulated statements that came before them in the thread. This social relationship is used to originate the social network utilized in this paper.

The theoretical underpinnings of the method utilized in this paper have not been previously presented as a well-articulated schema to understand online social interactions, but they seem to be in use by other researchers interested in this discipline. Prior to use of comment-to-comment relationships, a common method used to examine Reddit communities has been the identification of cross-posting networks, or posts (either images, text, or links to other web articles) that have been shared from one subreddit into another subreddit. While this method works well to understand the proximity of subreddits to one another and can determine to some extent how users interact between them, it misses the nuance of how posters in the forum themselves act and how that may be different than the wider lens of the subreddit as the observational unit. Sawicki et al (2023) is a modern example of this approach, where they used cross-posting data to construct a network of subreddits connected to each other within a calendar year and compared their findings to previous studies on cross-posting. They found that most cross-posts originated from a small set of large, general interest subreddits that shared popular conversations amongst each other and to other corners of Reddit. More recently, it has become common to see other analyses take

advantage of the social connections forged between comment pairs. Buntain and Golbeck (2014) set the stage for this method of analysis by constructing a network of comments exchanged between users on the same subreddit to assess the roles of user accounts, whether they were an “Answer-Person” or a “Non-Answer-Person”. They then were able to validate the assignment of these roles based on posts that featured an IAmA situation, a population Reddit conversation starter where a user with a niche or interesting position answers questions about they’re job from the rest of the community (more commonly known today as an AMA or Ask Me Anything post). Fang et al (2023) is another recent work that takes advantage of user comment interactions to assess the changing nature of subreddit rules over time, associating especially changes in total degree with positive changes to rules that encouraged cooperation and restricted the types of posts allowed. Brambilla et al (2021) uses a slightly modified approach in which they created a two-mode network of submissions and comments to test how well a machine learning approach could be trained to determine the content of a Reddit comment. The comment-to-comment approach at building out the edges of a Reddit social network seems like an applicable method to test the research questions presented in this paper.

Online Moderation and Behavioral Changes

Previous social media moderation attempts have been a frequent subject of research into online responses to the restrictions they impose, as discussed earlier. Outside of Reddit, Seering et al (2017) examined the impact of a variety of tacit and explicit moderation styles common on twitch.tv, a video live streaming service popular with video game players and viewers, and found that methods meant to deter users from certain behaviors were generally effective. This could be considered the motivation for the Reddit quarantine imposed on r/The_Donald, as it was an explicit set of rules intended to deter other Reddit users from interacting with the site. Unfortunately,

previous evidence of the effectiveness of these types of measures has been mixed. While Chandrasekharan et al (2017) found that an earlier 2015 ban of two explicitly hostile subreddits, r/fatpeoplehate and r/CoonTown, was impactful in lowering the relative levels of hate speech on the site, a later piece by Shen and Rose (2019) suggests that when quarantines are utilized by Reddit that right-wing users tend to cry out against the supposed “censorship”, which may impact how other adjacent communities view present and future moderation attempts. Although prior evidence about the impact of the quarantine and ban of r/The_Donald has generally been positive, research on other Reddit moderation attempts is still ongoing.

Even when restrictive measures are put into place, past findings have also suggested ways in which online communities may be able to weather such changes and continue into the future. Tsugawa and Niida (2021) looked into common statistics of on an unweighted Reddit comment network to determine if any were associated with the growth or longevity of the communities examined. They found that while graph attributes were not associated with an increase in the number of users interacting on a particular subreddit, that the networks best fit for survival were sparse, not highly centralized, and had a large core component. Such communities are likely to be more resistant to demise following restrictions imposed by Reddit, which may apply to the r/The_Donald community. It is also understood that following the quarantine and later ban of r/The_Donald many users migrated to alternate websites set up to imitate the original forum. While beyond the scope of this study, previous efforts have found success in linking usernames across web domains (Trujillo and Cresci 2023) and other approaches have attempted to categorize the relative levels of user migrations between social media sites following the imposition of new restrictions by outside influences. Newell et al (2016) took a look at a flare up in tensions on Reddit in 2015 (which occasionally come to a head during major cultural moments and subreddits often protest by “going dark”; see Porter and Peters (2023) for more on the protest in response to the API changes that

killed Pushshift) and led to an exodus of users from the platform. Unsurprisingly, Reddit maintained its market share as the leading web forum in part due to the variety of content hosted on the site and the popularity of its leading subreddits; smaller outlets comparable to Patriots.win that arose from the demise of r/The_Donald were unable to attract and maintain a critical mass, resulting in nothing but a rump offshoot of the original forum participants. Previous findings on the efficacy of moderation attempts have been somewhat mixed and leave room for the investigation of r/The_Donald proposed by this paper.

Data

This paper utilizes the Pushshift repository of saved Reddit content. This multi-year effort enhances access to Reddit posts and comments dating back to 2005 by bypassing the request throttling of the official Reddit API and the inability to retrieve posts using before/after date structures. This repository has been utilized in multiple peer-reviewed papers and is renowned for the ability to extract historical data not obtainable via the official repository (but not without criticism, see Gaffney and Matias (2018) for more details). For more on the architecture of this dataset, see Baumgartner et al (2020). Last summer, Reddit adjusted public access to its API data stream following the rise of publicly trained large language models such as ChatGPT, restricting how many free calls clients can make and requiring moderator status to access archived data (Geurkink 2024). This change has disrupted the data pipeline that the Pushshift community relied on, and the site has since been restricted to allow access to registered moderators only. Third-party file repositories exist of the archival data, but the files utilized for this paper are available upon request.

The research questions presented in this study use the date of the subreddit quarantine, June 26th, 2019, as the intervention date, with users having little foreknowledge of the impending restrictions and exposure by every contemporary user of the forum. To establish the sample of users to collect data from, I gathered all comments submitted to r/The_Donald from June 1st through June 25th, 2019. Only comments (not submissions) were gathered due to the implied conversational interaction of the comment structure on Reddit; submissions may be an image or a link to a news article, while comments are in dialogue with each other and suggest a deeper connection between users. I used a Python script to gather comments by scanning the large .zst month files available from Pushshift for comments with a matching subreddit data field as the sought-after subreddit, r/The_Donald. This results in some 41,475 unique users with non-deleted accounts who posted a comment on the forum during this time. I then ran a similar process on all comment submissions from June 27th, 2019, through July 21st, 2019, an equivalent 3-week time frame after the quarantine was imposed to construct a dataset of all comments posted afterwards. These comment datasets were then merged and restricted to only users who posted a comment branching from another comment (excluding top-level replies to the submission itself) in both the pre- and post-quarantine periods, with a rule recursively applied to include accounts that are commenting on comments from other accounts in the sample. This ensures that no isolated accounts will be present in the completed network, simplifying network comparisons by requiring both the pre- and post-quarantine networks have the same node set. The final sample consists of 11,295 user accounts that were active in both time periods and commented in response to at least one comment of another user in the sample.

A key innovation of this paper is the novel conceptualization of key actors on r/The_Donald in the immediate proximity of the quarantine on June 26th, 2019. Instead of previous attempts to include all users posting more than a proscribed frequency over an extended period, this paper attempts to

capture the reaction of users actively engaged on the subreddit in a brief window before and after the quarantine. This approach includes the previously designated “core users” while expanding the population of interest to include less engaged and more peripheral actors who may view r/The_Donald frequently but post less often. It is suggested that this method of account inclusion will better capture the dynamics of the subreddit during this snapshot of time and is replicable to larger timeframes for future study.

With both the pre- and post-quarantine users gathered, I constructed a network for both time periods using the parent comment information found in the comment dataset. The parent comment ID was merged back with the comment list to attach the target account name, with the source account being the comment poster. The network is weighted by the number of times that one account replied to another and permits loops, where one poster replies to a comment that they previously made. There are 103,643 comments represented in 88,952 edges in the pre-quarantine network and 107,264 comments represented in 93,687 edges in the post-quarantine network. Additional node data points are added based on a summation of the included accounts’ other posting behavior: the total number of comments that users submitted on Reddit in the 3-week period, the proportion of those posts on r/The_Donald, and the subreddit with the account’s largest share of comments during the 3-week timeframe.

Methods

This paper supplies a combination of descriptive and inferential statistics interrogating both the pre- and post-quarantine networks to supply answers to the three primary research questions. To address the first research question, this paper initially examines network summary statistics and degree distributions of both the pre- and post-quarantine networks to begin to assess similarities

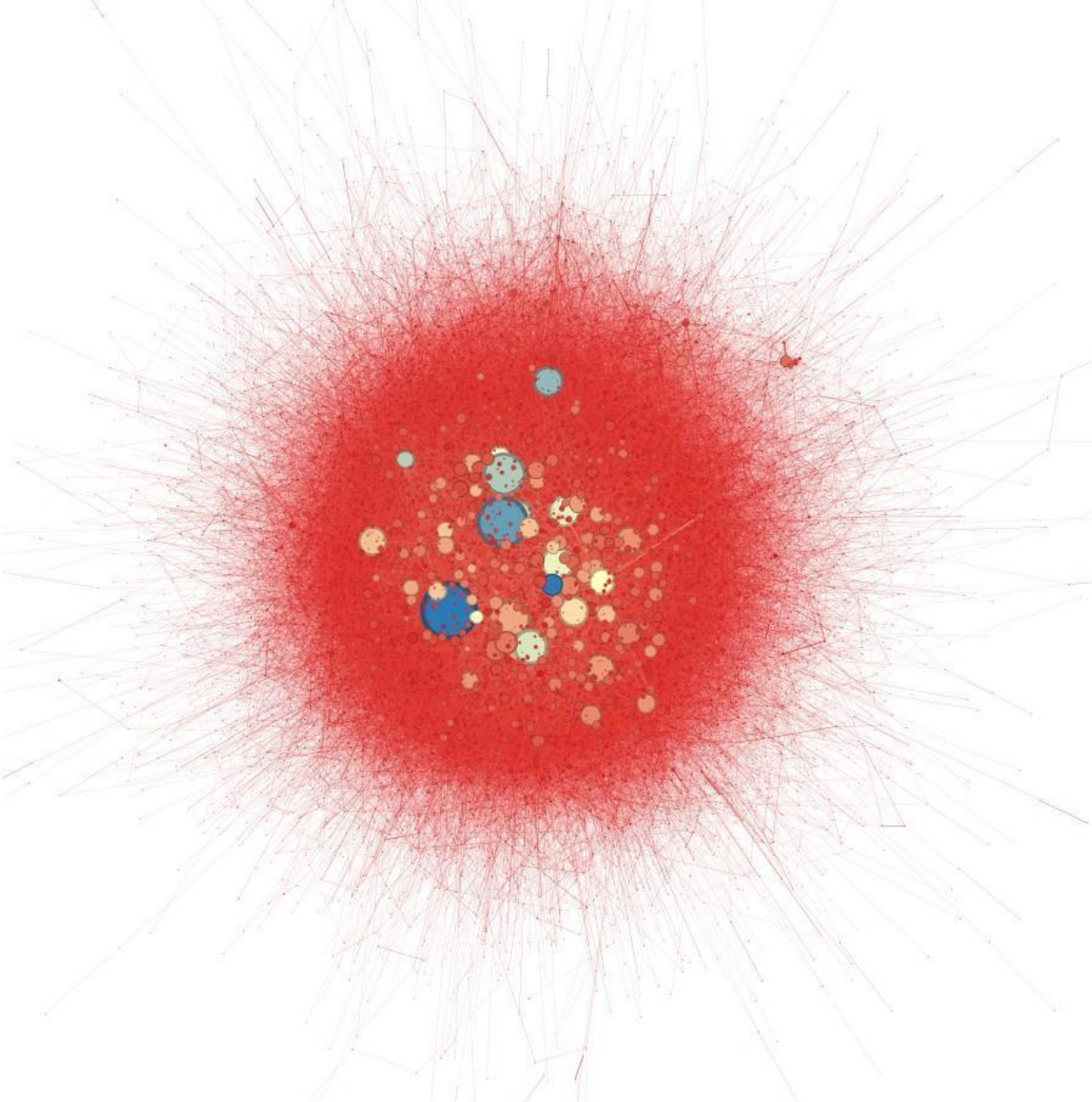
and differences between them before utilizing the DELTACON function (Koutra, Vogelstein, and Faloutsos 2013) from an R script to quantify the network similarity of the node sets. DELTACON is utilized due to consistent named node sets in both networks and the graph's directed construction; for a comparison of network similarity metrics, refer to Tantardini et al (2019). A high value stipulated through use of this algorithm would suggest a similar underlying network structure, with r/The_Donald users commenting on the posts of the same users in both time periods with about the same frequency, whereas a low value would suggest that users are replying to different accounts altogether. These methods will provide insights into whether and to what extent members of the r/The_Donald community altered their interaction patterns with other users in the initial weeks following Reddit's moderation attempt.

The second research question will be examined utilizing an initial qualitative scan of the summary tables for the most frequently commented-in subreddits in both the pre- and post-quarantine periods as well as scatterplots of the share of user comments in r/The_Donald over time. These results provide a broad overview of how users shifted their behavior in relation to other subreddits, but a more rigorous approach involves utilizing linear network regression to predict the proportion of comments in r/The_Donald in the post-quarantine network based on prior r/The_Donald comment share and the number of total comments that a user produced in the pre-quarantine timeframe. This model will help ascertain whether r/The_Donald users shifted their posting behavior away from the subreddit in the aftermath of the quarantine and whether this is associated with how active they previously were in the community and their overall activity level on Reddit. These analyses will shed insights into whether members of the r/The_Donald community meaningfully shifted their attention to other corners of Reddit when it became more difficult to access the forum and what types of communities they were likely to comment in more or less frequently after the quarantine.

The final research question of this paper will be assessed using a similar method as the second, first using a table of pre- and post-quarantine network degree and eigenvector centrality values to judge the relative increase or decline of influence for a subset of noteworthy nodes before proceeding with modeling, using a linear network regression, the post-quarantine in-degree of r/The_Donald users using the pre-quarantine in-degree, posting share during the pre-quarantine period in r/The_Donald, and pre-period total posts as covariates. In-degree is utilized in this model to capture the relative popularity of user posts in both time periods, as comments that are provocative, powerful, or in agreement with the beliefs of the community will have more interaction and more comments in return. The additional regression variables will help cast the post-quarantine in-degree as a reflection of how influential the accounts previously were as well as how active they were in r/The_Donald before the event. Upon completion, the results of this research question will indicate whether the influential figures in the pre-quarantine network retained their community influence after the moderation attempt and if the forums they continued to post in had an influence on their relative power in the post-quarantine network.

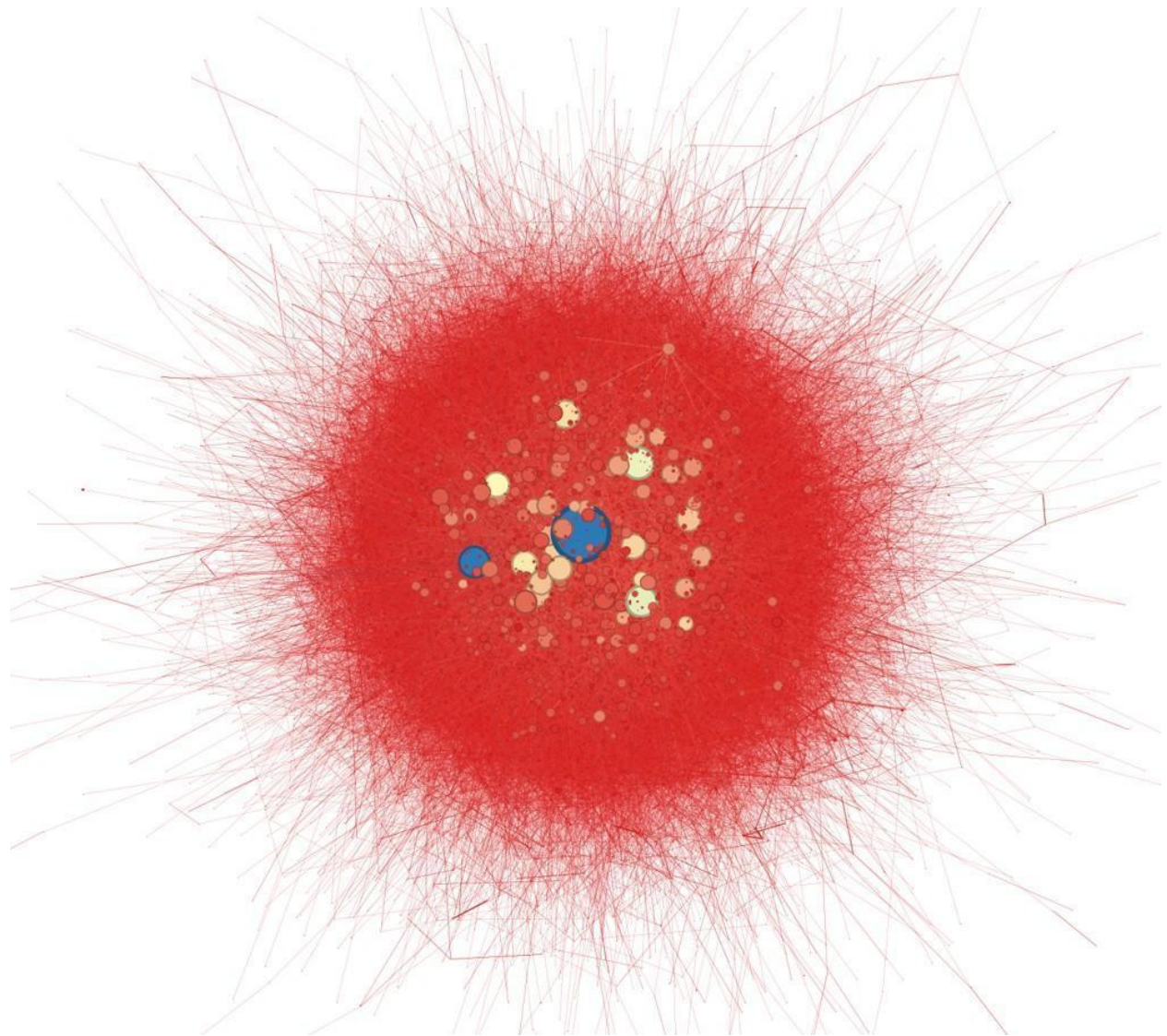
Results

Graph 1: Pre-Quarantine Network



The image above depicts the network of the r/The_Donald community before the quarantine, spanning the three weeks before restrictions were applied to the subreddit. Nodes are scaled based on their in-degree, representing the frequency of replies they received to their comments, with larger nodes indicating higher values. Additionally, nodes are colored according to their out-degree, which signifies the number of replies they sent to other's comments. Blue nodes represent more frequent commenters, while red nodes represent less frequent ones.

Graph 2: Post-Quarantine Network



The markedly similar image above is the network of the post-quarantine r/The_Donald comment community. Nodes are similarly scaled based on in-degree and colored based on out-degree.

Research Question 1

Table 1: Pre-Quarantine Summary Statistics

Network	Network_Size	Network_Density	Mean_Degree	Median_In_Degree	Median_Out_Degree	K_Core_Mean	K_Core_Median
Pre-quarantine	11295	0.0008053103	18.27118	3	3	8.066932	6

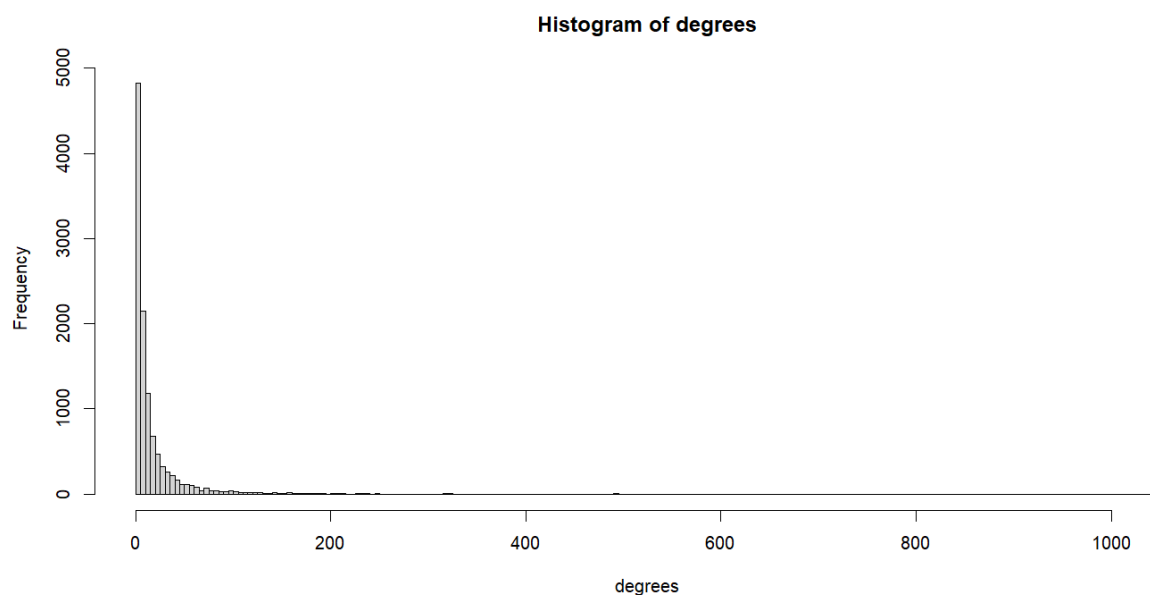
The table above presents summary statistics for the pre-quarantine network. It is important to note the difference between the average total degree and the median in- and out-degree: in this network, where all potential ties are accounted for and only comments between users are recorded as ties, the average and median in- and out-degrees are identical. However, there exists a discrepancy between a few users who post many comments and most users who only post less frequently, that is captured by comparing the two. Previous attempts to categorize “core users” of r/The_Donald based on a once-a-week posting frequency may overlook significant activity on the subreddit. In this network the k-core mean and median indicate a dense cluster of nodes at the center of the network, with many nodes interacting with popular users; this is reflected in the image of the graph previously presented. Despite this dense core, the network density remains low, with most users engaging with a few comments overall.

Table 2: Post-Quarantine Summary Statistics

Network	Network_Size	Network_Density	Mean_Degree	Mean_In_Degree	Mean_Out_Degree	K_Core_Mean	K_Core_Median
Post-quarantine	11295	0.0008330372	18.90491	3	3	8.502081	6

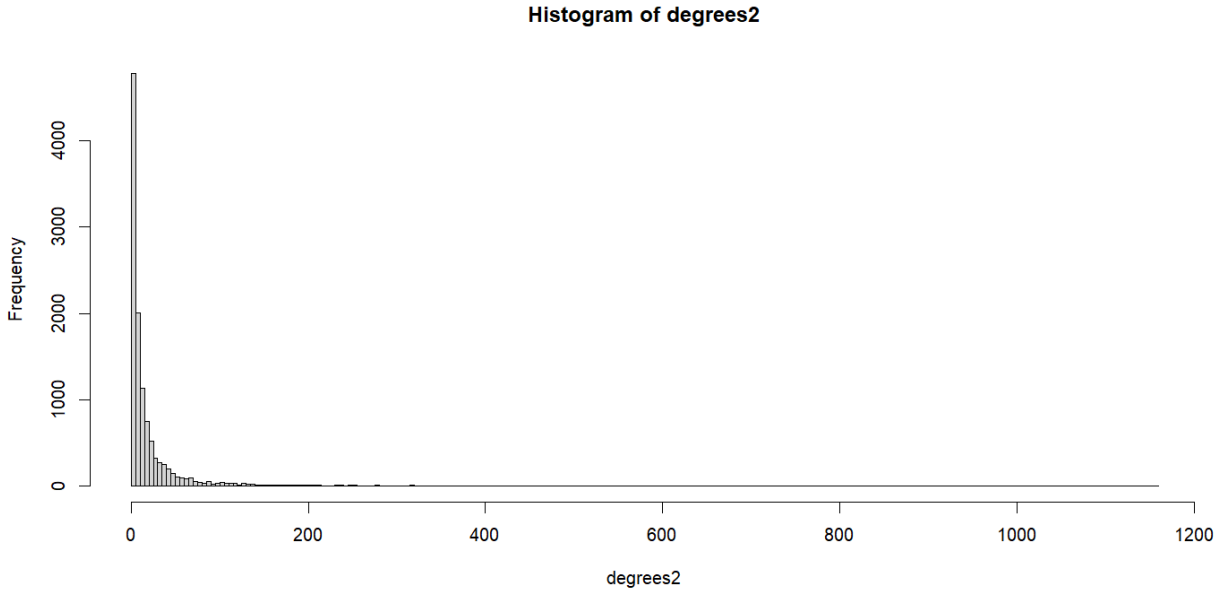
Table 2 presents the same information for the post-quarantine network. Although the median in- and out-degree values remain consistent between the two time periods, there is an increase in the mean total degree, implying a slight rise in network density. This observation suggests that the level of activity among this subset of r/The_Donald users experienced a slight uptick following the quarantine, resulting in a denser network. Moreover, the core-periphery structure observed in the earlier period persists after the quarantine, indicating a degree of similarity between the two networks.

Graph 3: Pre-Quarantine Degree Distribution



The above histogram of the total degree distribution for the pre-quarantine network highlights that most nodes in the network have very few ties, but some accounts have hundreds of combined replies and comments. This skewed distribution suggests that several accounts are tightly enmeshed with the entire community and are the focus of much of the conversation.

Graph 4: Post-Quarantine Degree Distribution



Similar to the distribution above, the total degree distribution of the post-quarantine network shows how very few nodes have more than a couple total connections, but there is an adjustment to the scale as one node has a total degree value of 1,156, compared to the max in the pre-quarantine network of 1,046 (incidentally, these were not the same user, with the pre-quarantine maximum belonging to u/boobiemcgoogle and the post- to u/advancedrescue).

Table 3: Pre-Quarantine Degree Distribution Comparison

Distribution	np	log-lik	AIC	BIC
Poisson	1	-208767.85	417537.70	417545.03
Geometric	3	-43466.90	86939.80	86961.79
NegBinom	3	-42087.85	84181.71	84203.71
Yule	3	-42913.00	85832.01	85854.00
GeoYule	4	-42145.45	84298.90	84328.22
Waring	3	-41323.35	82652.70	82674.69

The relative fit of the total degree distribution of both the pre-and post-quarantine networks to parameterized known distributions was calculated to infer some characteristics of the degree distribution. For the pre-quarantine network, the Waring distribution is the preferred model when utilizing the log-likelihood, the AIC, and the BIC, suggesting that it is the best fit for our degree distribution. This corresponds well with the elongated tail of the degree distribution and serves as a comparison point with the post-quarantine network.

Table 4: Post-Quarantine Degree Distribution Comparison

Distribution	np	log-lik	AIC	BIC
Poisson	1	-210889.72	421781.44	421788.77
Geometric	3	-43681.80	87369.61	87391.60
NegBinom	3	-42323.95	84653.90	84675.90
Yule	3	-43445.92	86897.84	86919.83
GeoYule	4	-42478.30	84964.60	84993.93
Waring	3	-41807.61	83621.22	83643.22

Table 4 above highlights the closest modeled fit for the post-quarantine network. While less similar to the Waring distribution than before, it is still the nearest match to the observed degree distribution of this network. This is again expected given the elongated tail of users with very high total degree values and matches with the prior distribution for the pre-quarantine network.

Table 5: DELTACON Metric

DELTACON_Method	Similarity_Score
Naive	0.2446724
Fast	0.4020928

The table above showcases the DELTACON similarity scores computed during a comparison of the pre-quarantine network to the post-quarantine network. This algorithm utilizes an edge list for both graphs to compare the frequency of connections between similar nodes at both time periods. A score of 0 would indicate totally dissimilar graphs, while a score of 1 would denote identical networks. To ensure accurate comparison, node IDs for the post-quarantine network were arranged into the order of the pre-quarantine network before running the function, given the separate storage order for both networks. The naïve method of calculation compares all pairs of nodes by ID to construct the similarity score while the fast method accounts for underlying network structure and has a sharper focus on influential nodes when determining the similarity score.

Given the core-periphery structure identified in both networks from both time periods, it is not surprising to observe a greater similarity between the two graphs with the fast method. Although a higher score would suggest greater similarity to the pre-quarantine graph, the post-quarantine network demonstrates an impressive reflection of the original network structure, considering the

assumption that posters are not likely to remember the accounts they previously replied to. This analytical metric suggests a significant continuity in the network structure after the attempted moderation than may have been anticipated.

An overview of the summary statistics of both the pre- and post-quarantine networks, comparisons of their degree distributions, fitting said distributions to known distributions, and calculating the similarity scores showcasing the relative fit of both graphs collectively suggest more underlying similarity between the two time periods than initially expected after the implementation of the quarantine. While the parameters set on the construction of the network, with the requirement that accounts be active in both time periods, may impact the similarity to some extent, the congruence of all supplied statistics suggests that, at least initially, not too many changes occurred in overall comment patterns on r/The_Donald.

Research Question 2

Table 6: Pre-Quarantine Most-Frequented Subreddits

Subreddit	x
The_Donald	597642
AskReddit	92684
unpopularopinion	47517
politics	43393
worldnews	26459
teenagers	26294
dankmemes	20300
nba	19318
pics	19137
news	17490
memes	16254
funny	15953
conspiracy	15717
MGTOW	13396
AmItheAsshole	12708

The figure above highlights the most popular subreddits that members of the r/The_Donald pre-quarantine network were commenting in during the three weeks prior to Reddit moderation.

Expectedly, r/The_Donald holds down the most comments, with more general political subreddits like r/politics and r/worldnews receiving far fewer comments. Many of the sites that are most

popular are similarly common for many Reddit users, like r/news, r/memes, and r/funny, and others like r/AmItheAsshole and r/nba have avid interest group followings. Interestingly, there is a significant overlap between the r/The_Donald and r/conspiracy, a forum dedicated to strange happenings around the world and their (mis)interpretation, and r/MGTOW, or Men Going Their Own Way, a misogynist subculture adjacent to incels that promoted male supremacy until it was eventually quarantined and then banned in 2020 and 2021, respectively (Thalen 2021).

Table 7: Post-Quarantine Most-Frequented Subreddits

Subreddit	x
The_Donald	562942
AskReddit	165193
dankmemes	99013
politics	69524
memes	68595
unpopularopinion	52485
AmItheAsshole	52107
Showerthoughts	43837
conspiracy	32095
teenagers	31420
PewdiepieSubmissions	28713
worldnews	24984
darkjokes	22180
nba	20613
Conservative	20350

Table 7 showcases the most popular subreddits for the post-quarantine r/The_Donald network to comment in during the three weeks following Reddit moderation. Interestingly, all of the most popular subreddits during this time period saw increases to their raw number of comments aside from comments to r/The_Donald, and this isn't reflected only at the top end of the distribution shown here: in total, 2,506,953 comments were posted by r/The_Donald users in the pre-quarantine period and 3,831,880 were posted after access was restricted. This surprising 52% increase in the overall number of comments posted suggests that r/The_Donald users active on the site just before moderation were more active on Reddit following the attempt to cordon them off. Many top subreddits from the pre-quarantine network find their way to the top of the charts in the later era as well, with the newcomers being r/Showerthoughts, a subreddit for random realizations, r/PewdiepieSubmissions, a fan forum for YouTuber Pewdiepie, r/darkjokes, for disturbing humor, and r/Conservative, a meeting place for the more mainstream right-wing of Reddit. It is beyond the scope of this paper, but a deep dive into the types of subreddits that the r/The_Donald community simultaneously inhabited based on similar userbases and how this shifted following the quarantine would yield additional insights. For now, there is no significant shift in the most popular subreddits r/The_Donald users commented in after the quarantine.

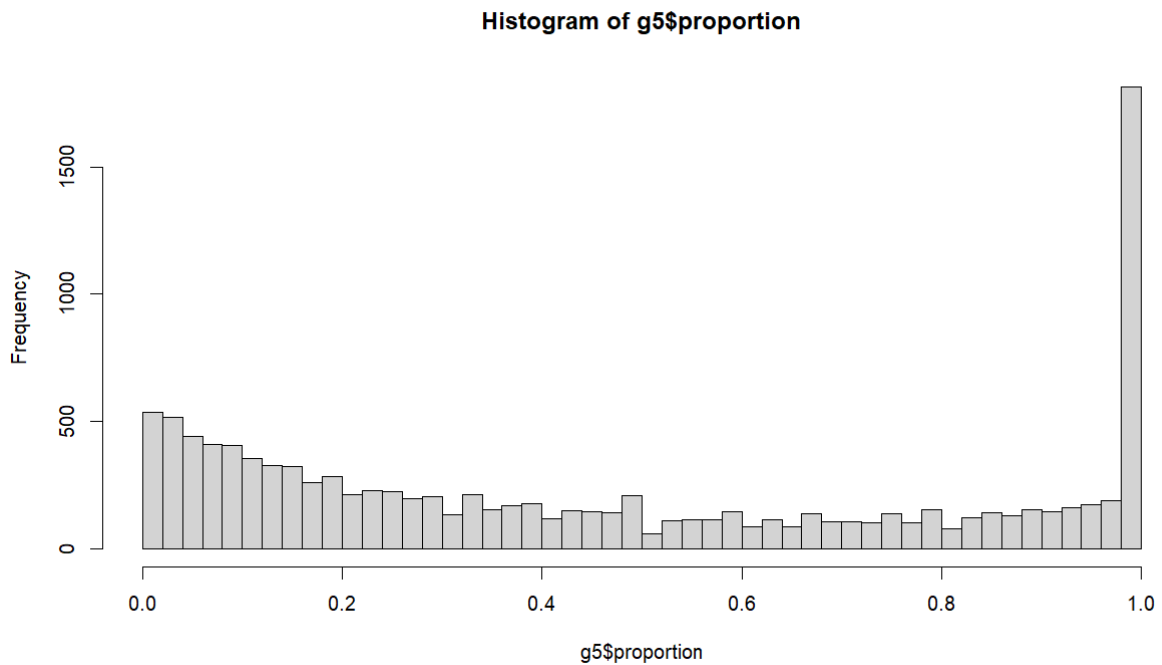
Table 8: Greatest Change in Subreddit Number of Comments

Subreddit	PercentChange	Subreddit	PercentChange
AskDocs	53.99485	blog	-0.9966997
WritingPrompts	34.03000	straightpride	-0.9913043
darkjokes	30.41643	LostLandsMusicFest	-0.9835165
SkincareAddiction	30.00833	LingerieGW	-0.9717514
HelpMeFind	26.29730	u_KaslanCorp	-0.9663677
PhotoshopRequest	22.61947	stlouisblues	-0.9322917
StrangerThings	21.77500	FreeKarma4You	-0.9218750
photoshopbattles	19.69767	BostonBruins	-0.9218203
cursedimages	18.88012	40plusGoneWild	-0.9175258
Market76	18.74157	wec	-0.8904110
MakeMeSuffer	17.89640	baldursgate	-0.8636364
nukedmemes	16.39091	FallenOrder	-0.8602941
Celebhub	16.16463	MinionMasters	-0.8598726
softwaregore	16.14824	warriors	-0.8567474
Advice	16.07198	ChernobylTV	-0.8506694

The table above highlights the subreddits with the greatest proportional change in the number of comments they experienced from r/The_Donald users following the site quarantine, restricted to subreddits that had more than 100 comments from r/The_Donald users in the pre-quarantine period (choosing this threshold is more art than science, smaller values were more likely to capture big swings in comments on pornographic subreddits while larger values were more generic subreddits with broader bases). Only one subreddit from the most popular list, r/darkjokes, makes the table above with a 304% increase in the number of posts. The rest tend to be advice subreddits, potentially highlighting how a viral post can attract comments from across different Reddit

communities. The subreddits that lost the greatest share of their comment activity range from almost going completely dark to hemorrhaging 85% of comment activity, and tend to be pornographic subreddits, video game communities, or the NHL team fandom at r/stlouisblues, who had just capped their last-to-first Stanley Cup run against the r/BostonBruins. Temporal phenomena do not motivate the majority of either list, with the only meaningful subreddit that gained influence due to an event was r/StrangerThings, which had a new season released during the post-quarantine period. It could be hypothesized that political subreddits would see an increased share of r/The_Donald migrants following Reddit's attempt to discourage activity on the forum, but this does not appear to be the case for the community analyzed in this paper.

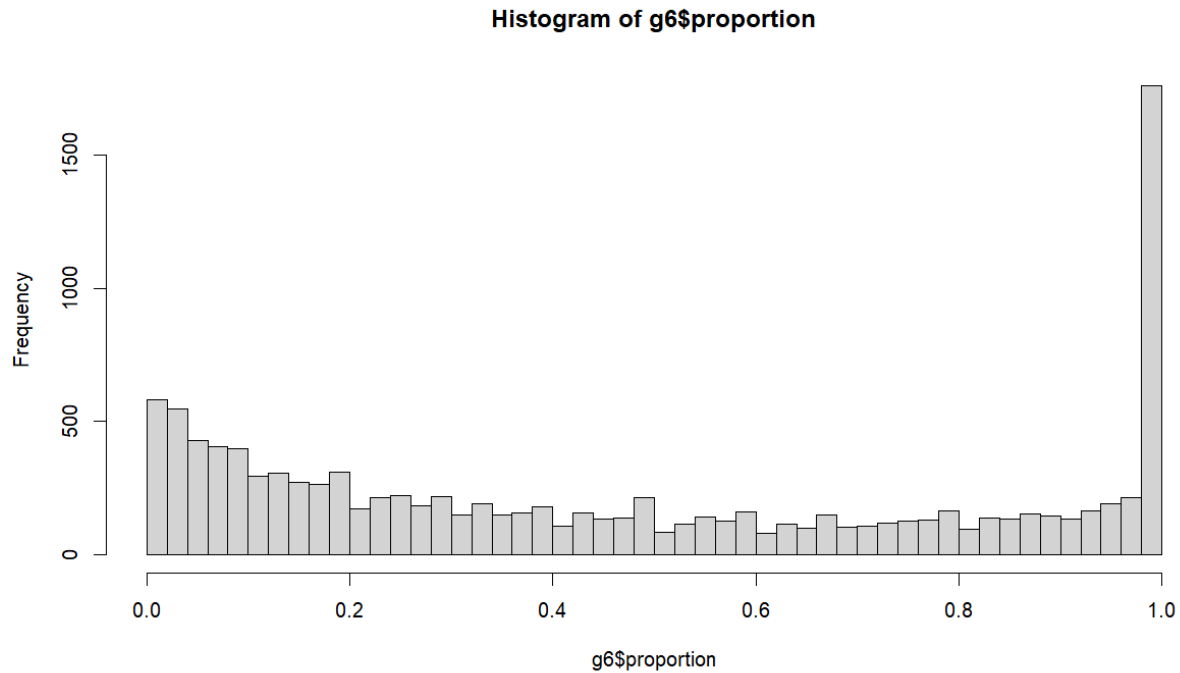
Graph 5: Pre-Quarantine r/The_Donald Share



The histogram above shows the distribution of r/The_Donald users in the pre-quarantine network based on the proportion of their overall comments placed in that subreddit during the earlier period. There is a local maximum in the left-most bin of the graph, indicating that just over 500

users posted on r/The_Donald with a fraction of their total comments on Reddit in June 2019, suggesting that they were only tangentially associated with that community, the relative proportional share of posts on the subreddit gradually decline through 50%, before staying at a relatively low frequency of users until the last bin indicating 100% of posts. This population dominates, with 1,644 users interacting solely with r/The_Donald. Without a detailed dive into the overall Reddit behavior of these users, it is difficult to determine whether these accounts are dedicated throwaway usernames commonly used across Reddit to share opinions, images, or thoughts outside of the main account of an individual or if they are accounts singularly focused on r/The_Donald. A qualitative scan into analyzing only users with more than 1 or 10 posts in r/The_Donald shifted the distribution even more to the right without diminishing the relative frequency of 100% commenters.

Graph 6: Post-Quarantine r/The_Donald Share



Graph 6 shows a similar distribution for the post-quarantine network, with the relative frequency of users with a low proportion of posts on r/The_Donald starting high before tapering off around 50% to a steady level and increasing again 1,561 100% commenters. The similarity of these two histograms suggests that most users on r/The_Donald were not shifting their posting behavior enough to impact the overall norms of the networked community. Similar to the analysis of the most popular subreddits that r/The_Donald users also commented in, the distribution of r/The_Donald posting shares also suggest that there were not major changes to the places where this community was posting, indicating a light impact of the quarantine measures on this cohesive group of users.

Table 9: Predicting Post-Quarantine r/The_Donald Share

```

OLS Network Model

Residuals:
      0%      25%      50%      75%     100%
-0.7780070 -0.2856148  0.1352996  0.3167838  0.3760694

Coefficients:
              Estimate      Pr(<=b) Pr(>=b) Pr(>=|b|)
(intercept)  0.6493704615  1.00   0.00   0.00
x1           0.0000910962  1.00   0.00   0.00
x2          -0.0257131745  0.01   0.99   0.01

Residual standard error: 0.3381 on 127565727 degrees of freedom
Multiple R-squared: 0.008589    Adjusted R-squared: 0.008589
F-statistic: 5.526e+05 on 2 and 127565727 degrees of freedom, p-value: 0

Test Diagnostics:

      Null Hypothesis: qap
      Replications: 100
      Coefficient Distribution Summary:

      (intercept)      x1      x2
Min      1627.123 -299.313 -278.104
1stQ     1743.112  -96.117  -74.539
Median   1811.346  -20.315  -4.625
Mean     1817.497  -15.248  -3.868
3rdQ     1883.432   78.833   79.038
Max      2071.130  235.647  224.607

```

The regression output summarizes the network linear regression run on a matrix of post-quarantine r/The_Donald comment proportions on independent matrices of the pre-quarantine total number of comments (x1) and r/The_Donald comment proportions (x2). All p-values for the intercept, x1, and x2 are significant with an alpha level of 0.05. This suggests that an increase in the share of posts in the pre-quarantine period likely leads to a slight increase in the proportion of comments in r/The_Donald during the later time period. Conversely, an increase in the share of earlier comments in r/The_Donald decreases the propensity to comment in the subreddit during the later time period. The values of the comment proportion variables scale from 0 to 1, so it is not surprising that the value of the regression coefficients is small, but more concerning for the overall

relationship between these chosen variables and the dependent variable is the miniscule R^2 value. At just 0.008, it suggests that very little variation in the observed values of post-quarantine *r/The_Donald* commenting shares is explained by fluctuations in these independent variables. This is in contrast to the relatively large F-statistic and p-value of 0 for the model overall, suggesting that there is a significant relationship between this set of variables. The interpretation of this paper is that while there is a degree of correlation between pre-quarantine total number of comments, *r/The_Donald* comment proportion, and post-quarantine *r/The_Donald* comment proportion, levels of change for most users may be too small to exhibit a stronger explanation for the variation seen in users who significantly changed their behavior. This would align with the earlier interpretation that for the community as a whole there was little average change to the subreddits that users were commenting on after the moderation attempt, with what little change there was occurring on the margins. The evidence provided by both the subreddit analysis and the regression results suggest that during the initial weeks following the restrictions imposed by Reddit, community members of *r/The_Donald* were not greatly impacted in their commenting behavior.

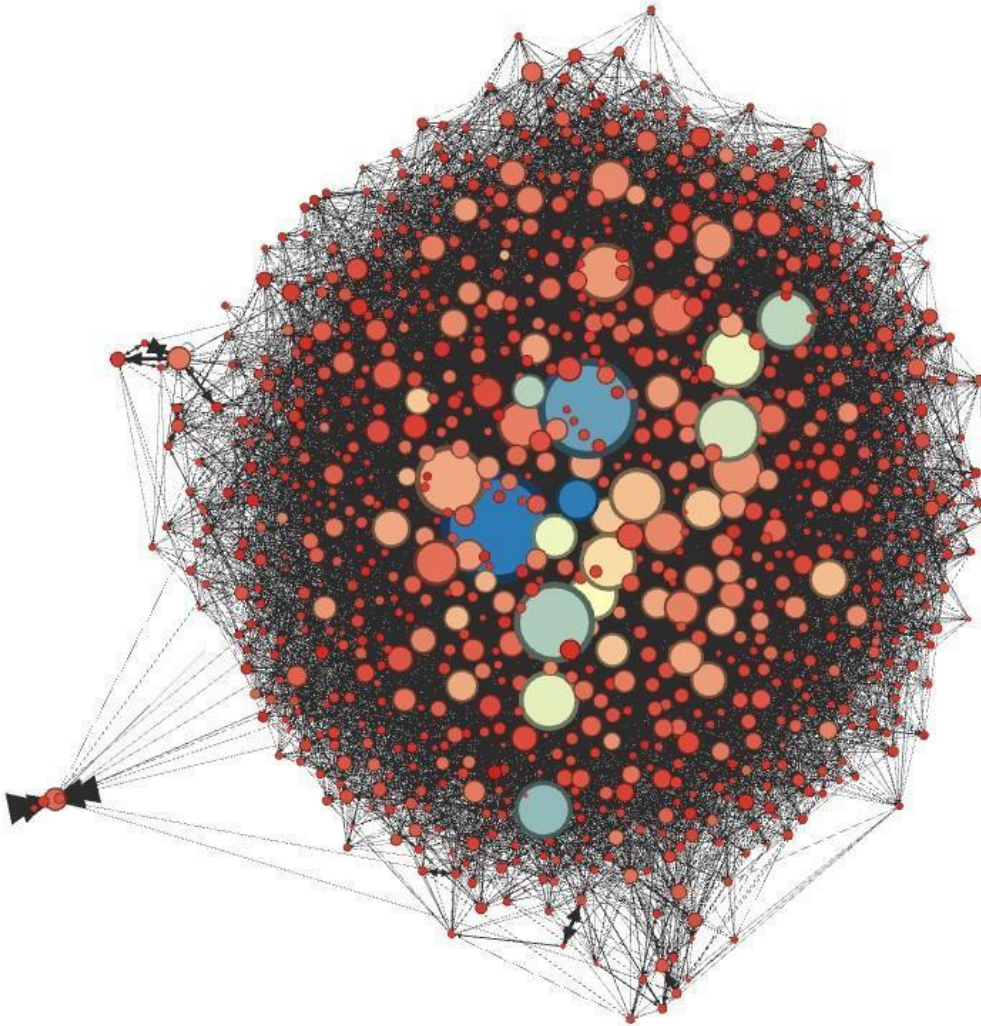
Research Question 3

Table 10: Top Nodes by Eigenvector Centrality, Pre- and Post-Quarantine

User	preQeigs	User	postQeigs
OhLollyLollyPop	0.2070639	OhLollyLollyPop	0.18969645
bobadobalina	0.1813040	ActivatedJoeBot	0.17875069
covfefe_crusader	0.1591841	TrumpBrickBot	0.17768338
Useful_Vidiots	0.1208884	RiverFenix	0.13089428
brotherjustincrowe	0.1203975	NewSouthernBelle	0.12456256
cursedcassandra	0.1164663	DiamondsOnTheFloor	0.12019297
DiamondsOnTheFloor	0.1150959	RuleoVicus	0.11176434
RiverFenix	0.1143510	bobadobalina	0.10551223
happycookie6	0.1092711	booblitchutz	0.10185079
SAW2TH-55th	0.1087368	Chernoobyl	0.09644064

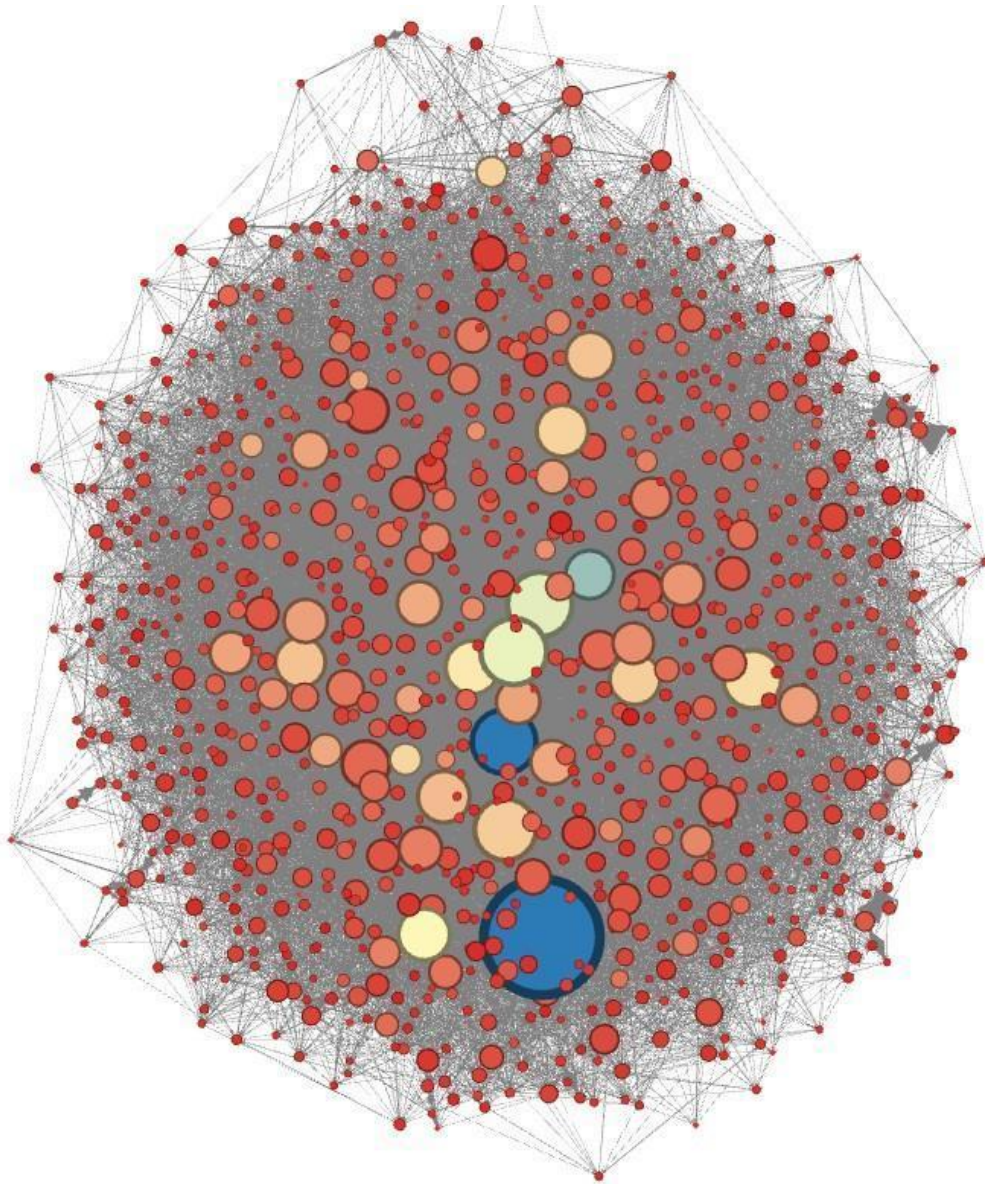
The table above highlights the pre- and post-quarantine eigenvector centrality for nodes in the r/The_Donald community. Eigenvector centrality assists in determining which actors in a graph hold influence not just due to their own ties to the rest of the network but also getting additional credit for being attached to other influential nodes themselves. As a measure of social influence, these users are both providing and receiving a large number of comments, as well as commenting on the posts of others doing the same. The top node in both time periods is u/OhLollyLollyPop, with others on the list at both time periods being u/bobadobalina and u/DiamondsOnTheFloor. This table suggests that there is significant turnover for the most influential individual accounts using this measure and suggests that additional looks into the powerful actors in the network may yield interesting results.

Graph 7: Pre-Quarantine Most Influential Nodes



The graph above zooms in on the most influential nodes according to total degree, filtering the original pre-quarantine network to those with a weighted degree value of 50 or more. This leaves just under 1,000 total nodes in the network, again colored based on out-degree and scaled based on in-degree. Even the smallest actors in this network commented or had their comments replied to dozens of times during the initial three-week period, leading to the dense core with only a few influential actors attracting attention away from the largest nodes. Note that most nodes are red, indicating a low out-degree value relative to the most prominent nodes.

Graph 8: Post-Quarantine Most Influential Nodes



This image of the post-quarantine network with the same filter stipulated for nodes with a weighted total degree value of 50 or more shows a similar structure to the previous graph, with a few large nodes near the center and a dense cluster containing nearly every other node in close proximity to them. The three blue-hued users are a bit further from the center than in the previous projection, indicating that although they may have a similar value of influence to the previous projection that

they may not be the center of conversation for as many users as in the initial time period. An additional quantitative look is needed to determine how degree values shifted for these influential figures and whether it was related to the frequency that they posted within r/The_Donald.

Table 10: Predicting Post-Quarantine In-Degree

```

Residuals:
      0%      25%      50%      75%     100%
-299.787017  -4.139804  -2.282268   1.397514  502.335749

Coefficients:
              Estimate      Pr(<=b) Pr(>=b) Pr(>=|b|)
(intercept)  3.4827348342  1.00    0.00    0.00
x1          -0.0001327859  0.48    0.52    0.75
x2          -0.2149715370  0.26    0.74    0.58
x3           0.6734030702  1.00    0.00    0.00

Residual standard error: 13.93 on 127565726 degrees of freedom
Multiple R-squared: 0.481      Adjusted R-squared: 0.481
F-statistic: 3.941e+07 on 3 and 127565726 degrees of freedom, p-value: 0

Test Diagnostics:

      Null Hypothesis: qap
      Replications: 100
      Coefficient Distribution Summary:

      (intercept)      x1      x2      x3|
Min          14.278 -297.634 -271.733 -204.101
1stQ         230.120  -79.252  -55.879  -83.022
Median       294.715  -27.966   -8.841  -25.789
Mean         310.815  -17.829    1.515   -5.311
3rdQ         398.687   42.033   74.560   53.137
Max          589.667  340.293  289.896  303.956

```

Table 10 above highlights the regression results testing the key tenet of research question 3 by running a network linear model predicting the post-quarantine in-degree of users in the r/The_Donald community using the pre-quarantine independent variables of total comments (x1), proportion of comments in r/The_Donald (x2), and early period in-degree (x3). With less significance to total comments and r/The_Donald share than the regression in Table 10, this model

places a heavy emphasis on the impact of the pre-quarantine in-degree of nodes in this community. The R^2 value suggests that almost half of the variation in observed values of post-quarantine in-degree can be accounted for using this model, with the F-statistic and associated p-value also lending support to utilizing this model to understand how actors maintain their influence. In an alternate simulation where actors had an identical value of in-degree in both the pre- and post-exposure networks the coefficient value for x_3 would be 1, indicating that accounting for the other independent variables does lead to a slight decline in the relative level of in-degree between the two periods, but this makes some intuitive sense: in this conceptualization, in-degree is based on other users commenting on a post of yours. Over a six-week period, it is unlikely that every post you make will have the same level of popularity and thus exposure and having a high value of in-degree in the early three-week period does not necessarily imply that you will also receive many replies later. Instead, this model serves as slight evidence to the hypothesis that the moderation attempt imposed by Reddit had a less severe impact on the interactivity of the r/The_Donald community than previously described, with connected users replying to each other at rates near the level they had previously.

Discussion

This paper aims to evaluate the impact of the subreddit quarantine imposed on r/The_Donald on June 26th, 2019, on the interactions among recent commenters on the site. By analyzing the social network constructed based on comment-to-comment ties that forum users had with each other in the three weeks before and after the restrictions were enacted, this study offers a nuanced perspective about the effectiveness of the moderation attempt and explores the dynamics of online communities. Summary statistics of both the pre- and post-quarantine networks suggest

minimal fundamental changes between the two time periods, and the DELTACON similarity score also lent credibility to the hypothesis that the quarantine had not done much to disturb the commenting connections between r/The_Donald users. The qualitative scan of the other subreddits that users were posting in during the same time period and the relative proportion of comments submitted to r/The_Donald indicated, if anything, a slight increase in the frequency of posts on that forum, with the network linear regression highlighting the relationship between past and present posting proportions in the subreddit without suggesting accounting for the variability seen in users who significantly changed their shares. The insights from the adjudication of the most influential nodes during the two time periods showed how there was some change in the top nodes by eigenvector centrality, but the network linear regression accounted for nearly half of the variability noted in the sample, due primarily to the predictive power of previous degree centrality for powerful nodes. In sum, the evidence presented here is more nuanced than previous research efforts that have proclaimed Reddit moderation a success. Utilizing a different sample population and methods that shift some focus onto frequent contributors to the site would yield a different bottom-line answer, but the takeaway from this paper has been that at least in the initial three weeks after the quarantine, it was not a clear victory for those wishing to watch r/The_Donald quickly fade to black.

The primary contribution of this paper lies in its alternative approach to analyzing close-knit communities that form on political subreddits, offering insights with broader implications for sociology, social media research, and our understanding of political identity formation. For sociology, this study is a verification of previous efforts that defined the social ties between users of online platforms as meaningful social interactions. Like in face-to-face conversations, the communications that we have online share a piece of us with those we choose to spend our time with, the same phenomenon also seen in constructing a networked social discourse of a subreddit.

The similarity between graphs of more than 10,000 nodes suggest that even in the relatively anonymous space of a subreddit, we are still producing ties with users we recognize and exchanging thoughts and ideas with others, even on a toxic political forum. For social media researchers, this piece highlights the utility of a social networks approach beyond just analyzing aggregate information about post numbers and the sentiment of text. Network regression models have the added assumption of non-independence baked in, crucial for understanding how users on a site like Reddit may influence each other and lead to the diffusion of new ideas. And maybe most prescient, this paper yields new insights about the formation of ideological ties in our new online era. Much attention has focused on the rise of the “alt-right”, the election of Donald Trump and the resurgence of vile views in online spaces. However, without understanding the origin of these viewpoints, the durability of the communities where they fester, and the effectiveness of means to combat them, researchers studying political identity formation and online radicalization in the modern era face significant challenges.

Limitations and Future Directions

While this paper offers a fresh perspective on a well-studied event in the history of social media, it encounters several limitations that affect the interpretation of its conclusions. Chief among these is the definition of the community as a contained social network, where only nodes who remained in the network at both time periods were included. Initially, all users who commented in either time period were considered for inclusion, resulting in a multitude of disparate, disconnected nodes added to the graph images and 0's imputed for the network regression analysis. However, it was determined that this approach could bias the regression results and limit the applicability of certain network similarity measures, such as DELTACON, which requires two identical nodesets to

function and is considered one of the more respected tools for categorizing change in network construction over time.

Additionally, the results of this study may be influenced by the chosen timeframe for both the pre- and post-quarantine networks. By focusing on the three-week period, this paper was able to examine users closest to the event itself, in contrast to prior examinations that stipulated a relative posting frequency for inclusion in the sample. Further investigation using longer time periods before and after the restrictions were implemented could provide a more comprehensive understanding of how users reacted to the quarantine. It is possible that initially, behaviors remained relatively normal before significant changes occurred. The current sample may not capture enough of this evolving behavior and there may be a short window that opens later in July to observe it before users begin fleeing the site.

Beyond these areas of additional work with this dataset, this paper suggests avenues for future research in both Reddit and the broader social media ecosystem. To further a comment network approach to analyzing online spaces, a larger study following the communities networked in a cohort of subreddits may yield insights into the dynamics of these types of networks and whether the findings of this study are unique to the experience of r/The_Donald following the quarantine. It may be that many subreddits experience a similar clustering of users who post solely on replies to posts of a particular subreddit, although this is considered more likely due to the taboo nature of the political positions held by the r/The_Donald community. Off Reddit, political community dynamics and identity formation could be studied in online environments like YouTube, which routinely discovers new content that they would like to moderate, such as anti-vaccine positions or Qanon. Subscribers to these beliefs will then often develop a covert language to discuss these topics whilst evading the automated content bots, which could yield interesting insights into how community usage of these terms changes with network interaction over time. Due to the pressing

nature of the political concerns stirred by the activities of r/The_Donald, it is likely that this domain will remain fruitful for social researchers in the near future.

References

- Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. 2015. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science* 348(6239):1130–32. doi: 10.1126/science.aaa1160.
- Barnes, Renee.Bennett, W. Lance. 2008. "Changing Citizenship in the Digital Age." Pp. 1–24 in *Civic life online: Learning how digital media can engage youth, The John D. and Catherine T. MacArthur Foundation series on digital media and learning*. Cambridge, MA, US: MIT Press.
- Blee, Kathleen M., Robert Futrell, and Pete Simi. 2023. *Out of Hiding: Extremist White Supremacy and How It Can Be Stopped*. Taylor & Francis.
- Collins, Randall. 2005. *Interaction Ritual Chains*. Princeton, NJ: Princeton University Press.
- Daniels, Jessie. 2018. "The Algorithmic Rise of the 'Alt-Right.'" *Contexts* 17(1):60–65. doi: 10.1177/1536504218766547.
- Freelon, Deen, Alice Marwick, and Daniel Kreiss. 2020. "False Equivalencies: Online Activism from Left to Right." *Science* 369(6508):1197–1201. doi: 10.1126/science.abb2428.
- Latif, Mehr, Kathleen Blee, Matthew DeMichele, and Pete Simi. 2018. "How Emotional Dynamics Maintain and Destroy White Supremacist Groups." *Humanity & Society* 42(4):480–501. doi: 10.1177/0160597618802552.
- Zúñiga, Homero Gil de, Matthew Barnidge, and Andrés Scherman. 2019. "Social Media Social Capital, Offline Social Capital, and Citizenship: Exploring Asymmetrical Social Capital Effects." in *Digital Politics: Mobilization, Engagement and Participation*. Routledge.
2018. *Uncovering Online Commenting Culture: Trolls, Fanboys and Lurkers*. 1st ed. 2018. Cham: Springer International Publishing : Imprint: Palgrave Macmillan.
- Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. "The Pushshift Reddit Dataset." doi: 10.48550/ARXIV.2001.08435.
- Bennett, W. Lance. 2012. "The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation." *The ANNALS of the American Academy of Political and Social Science* 644(1):20–39. doi: 10.1177/0002716212451428.
- Boulianne, Shelley, Karolina Koc-Michalska, and Bruce Bimber. 2020. "Right-Wing Populism, Social Media and Echo Chambers in Western Democracies." *New Media & Society* 22(4):683–99. doi: 10.1177/1461444819893983.
- Brambilla, Marco, Alireza Javadian, and Amin Endah Sulistiawati. 2021. "Conversation Graphs in Online Social Media." Pp. 97–112 in *Web Engineering*, edited by M. Brambilla, R. Chbeir, F. Frasincar, and I. Manolescu. Cham: Springer International Publishing.

Buntain, Cody, and Jennifer Golbeck. 2014. "Identifying Social Roles in Reddit Using Network Structure." Pp. 615–20 in *Proceedings of the 23rd International Conference on World Wide Web*. Seoul Korea: ACM.

Chandrasekharan, Eshwar, Shagun Jhaver, Amy Bruckman, and Eric Gilbert. 2022. "Quarantined! Examining the Effects of a Community-Wide Moderation Intervention on Reddit." *ACM Transactions on Computer-Human Interaction* 29(4):1–26. doi: 10.1145/3490499.

Chandrasekharan, Eshwar, Umashanthi Pavalanathan, Anirudh Srinivasan, Adam Glynn, Jacob Eisenstein, and Eric Gilbert. 2017. "You Can't Stay Here: The Efficacy of Reddit's 2015 Ban Examined Through Hate Speech." *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW):1–22. doi: 10.1145/3134666.

Chayko, Mary. 2014. "Techno-Social Life: The Internet, Digital Technology, and Social Connectedness." *Sociology Compass* 8(7):976–91. doi: 10.1111/soc4.12190.

Chen, Bin, Josephine Lukito, and Gyo Hyun Koo. 2023. "Comparing the #StopTheSteal Movement across Multiple Platforms: Differentiating Discourse on Facebook, Twitter, and Parler." *Social Media + Society* 9(3):20563051231196879. doi: 10.1177/20563051231196879.

De Koster, Willem, and Dick Houtman. 2008. "'Stormfront Is Like a Second Home to Me': On Virtual Community Formation by Right-Wing Extremists." *Information, Communication & Society* 11(8):1155–76. doi: 10.1080/13691180802266665.

Deen, Ahmed, and Po-Lin Pan. 2022. "Exploring Social Networks of #Election2020results and #BidenTransition on Twitter after the Presidential Election in the United States." *The Journal of Social Media in Society* 11(2):77–96.

Duncombe, Constance. 2019. "The Politics of Twitter: Emotions and the Power of Social Media." *International Political Sociology* 13(4):409–29. doi: 10.1093/ips/olz013.

Fang, Anna, Wenjie Yang, and Haiyi Zhu. 2023. "Shaping Online Dialogue: Examining How Community Rules Affect Discussion Structures on Reddit."

Gaffney, Devin, and J. Nathan Matias. 2018. "Caveat Emptor, Computational Social Science: Large-Scale Missing Data in a Widely-Published Reddit Corpus" edited by C. M. Danforth. *PLOS ONE* 13(7):e0200162. doi: 10.1371/journal.pone.0200162.

Gaudette, Tiana, Ryan Scrivens, Garth Davies, and Richard Frank. 2021. "Upvoting Extremism: Collective Identity Formation and the Extreme Right on Reddit." *New Media & Society* 23(12):3491–3508. doi: 10.1177/1461444820958123.

Geurkink, Sarah Gilbert and Brandi. 2024. "Why Reddit's Decision to Cut off Researchers Is Bad for Its Business—and Humanity." *Fast Company*. Retrieved April 9, 2024 (<https://www.fastcompany.com/91014116/reddit-researchers-bad-for-business>).

Gilmour, David. 2017. "4chan and Reddit Users Set out to Prove Seth Rich Murder Conspiracy." *The Daily Dot*. Retrieved February 19, 2024 (<https://www.dailydot.com/debug/reddit-the-donald-4chan-seth-rich-murder/>).

Goffman, Erving. 1981. *Forms of Talk*. University of Pennsylvania Press.

Grover, Ted, and Gloria Mark. 2019. "Detecting Potential Warning Behaviors of Ideological Radicalization in an Alt-Right Subreddit." *Proceedings of the International AAAI Conference on Web and Social Media* 13:193–204. doi: 10.1609/icwsm.v13i01.3221.

Horta Ribeiro, Manoel, Shagun Jhaver, Savvas Zannettou, Jeremy Blackburn, Gianluca Stringhini, Emiliano De Cristofaro, and Robert West. 2021. "Do Platform Migrations Compromise Content Moderation? Evidence from r/The_Donald and r/Incels." *Proceedings of the ACM on Human-Computer Interaction* 5(CSCW2):1–24. doi: 10.1145/3476057.

Kanti Baowaly, Mrinal, George William Kibirige, and Bikash Chandra Singh. 2022. "Co-Comment Network: A Novel Approach to Construct Social Networks within Reddit." *Computación y Sistemas* 26(1). doi: 10.13053/cys-26-1-4175.

Kidd, Dustin, and Keith McIntosh. 2016. "Social Media and Social Movements: Social Media and Social Movements." *Sociology Compass* 10(9):785–94. doi: 10.1111/soc4.12399.

Koebler, Jason. 2016. "How r/The_donald Became a Melting Pot of Frustration and Hate." *Vice*. Retrieved April 6, 2023 (<https://www.vice.com/en/article/53d5xb/what-is-rthedonald-donald-trump-subreddit>).

Koutra, Danai, Joshua T. Vogelstein, and Christos Faloutsos. 2013. "DeltaCon: A Principled Massive-Graph Similarity Function." Pp. 162–70 in *Proceedings of the 2013 SIAM International Conference on Data Mining (SDM), Proceedings*. Society for Industrial and Applied Mathematics.

Kruse, Lisa M., Dawn R. Norris, and Jonathan R. Flinchum. 2018. "Social Media as a Public Sphere? Politics on Social Media." *The Sociological Quarterly* 59(1):62–84. doi: 10.1080/00380253.2017.1383143.

Murthy, Dhiraj. 2012. "Towards a Sociological Understanding of Social Media: Theorizing Twitter." *Sociology* 46(6):1059–73. doi: 10.1177/0038038511422553.

Newell, Edward, David Jurgens, Haji Saleem, Hardik Vala, Jad Sassine, Caitrin Armstrong, and Derek Ruths. 2016. "User Migration in Online Social Networks: A Case Study on Reddit During a Period of Community Unrest." *Proceedings of the International AAAI Conference on Web and Social Media* 10(1):279–88. doi: 10.1609/icwsm.v10i1.14750.

Ohlheiser, Abby. 2021. "Reddit Bans r/Greatawakening, the Main Subreddit for QAnon Conspiracy Theorists." *Washington Post*, October 26.

Porter, Jon, and Jay Peters. 2023. "More than 7,000 Subreddits Have Gone Dark to Protest Reddit's API Changes." *The Verge*. Retrieved April 10, 2024 (<https://www.theverge.com/2023/6/12/23755974/reddit-subreddits-going-dark-private-protest-api-changes>).

Putnam, Robert D. 2001. *Bowling Alone: The Collapse and Revival of American Community*. 1. Touchstone ed. London: Simon & Schuster [u.a.].

- Reddit, Inc. 2023. "About- Reddit." *Reddit, Inc.* Retrieved February 18, 2024 (<https://www.redditinc.com/>).
- Robertson, Adi. 2019. "Reddit Quarantines Trump Subreddit r/The_Donald for Violent Comments." *The Verge*. Retrieved February 19, 2024 (<https://www.theverge.com/2019/6/26/18759967/reddit-quarantines-the-donald-trump-subreddit-misbehavior-violence-police-oregon>).
- Sawicki, Jan, Maria Ganzha, Marcin Paprzycki, and Yutaka Watanobe. 2023. "Reddit CrosspostNet—Studying Reddit Communities with Large-Scale Crosspost Graph Networks." *Algorithms* 16(9):424. doi: 10.3390/a16090424.
- Schulze, Heidi. 2020. "Who Uses Right-Wing Alternative Online Media? An Exploration of Audience Characteristics." *Politics and Governance* 8(3):6–18. doi: 10.17645/pag.v8i3.2925.
- Seering, Joseph, Robert Kraut, and Laura Dabbish. 2017. "Shaping Pro and Anti-Social Behavior on Twitch Through Moderation and Example-Setting." Pp. 111–25 in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17*. New York, NY, USA: Association for Computing Machinery.
- Shen, Qinlan, and Carolyn Rose. 2019. "The Discourse of Online Content Moderation: Investigating Polarized User Responses to Changes in Reddit's Quarantine Policy." Pp. 58–69 in *Proceedings of the Third Workshop on Abusive Language Online*, edited by S. T. Roberts, J. Tetreault, V. Prabhakaran, and Z. Waseem. Florence, Italy: Association for Computational Linguistics.
- Stieglitz, Stefan, and Linh Dang-Xuan. 2013. "Social Media and Political Communication: A Social Media Analytics Framework." *Social Network Analysis and Mining* 3(4):1277–91. doi: 10.1007/s13278-012-0079-3.
- Tantardini, Mattia, Francesca Ieva, Lucia Tajoli, and Carlo Piccardi. 2019. "Comparing Methods for Comparing Networks." *Scientific Reports* 9(1):17557. doi: 10.1038/s41598-019-53708-y.
- Thalen, Mikael. 2021. "Reddit Bans Notorious Anti-Feminist Subreddit 'Men Going Their Own Way.'" *The Daily Dot*. Retrieved April 9, 2024 (<https://www.dailydot.com/debug/mgtow-subreddit-banned/>).
- Tiffany, Kaitlyn. 2020. "Reddit Is Done Pretending The Donald Is Fine." *The Atlantic*. Retrieved February 19, 2024 (<https://www.theatlantic.com/technology/archive/2020/06/reddit-ban-the-donald-chapo-content-policy/613639/>).
- Trujillo, Amaury, and Stefano Cresci. 2022. "Make Reddit Great Again: Assessing Community Effects of Moderation Interventions on r/The_Donald." *Proceedings of the ACM on Human-Computer Interaction* 6(CSCW2):1–28. doi: 10.1145/3555639.
- Trujillo, Amaury, and Stefano Cresci. 2023. "One of Many: Assessing User-Level Effects of Moderation Interventions on r/The_Donald." Pp. 55–64 in *Proceedings of the 15th ACM Web Science Conference 2023*. Austin TX USA: ACM.
- Tsugawa, Sho, and Sumaru Niida. 2021. "Analyzing the Effects of Social Network Structure on the Growth and Survival of Online Communities in Reddit." *IEICE Transactions on Communications* E104.B(7):760–69. doi: 10.1587/transcom.2020CQP0006.

Wahlström, Mattias, and Anton Törnberg. 2021. "Social Media Mechanisms for Right-Wing Political Violence in the 21st Century: Discursive Opportunities, Group Dynamics, and Co-Ordination." *Terrorism and Political Violence* 33(4):766–87. doi: 10.1080/09546553.2019.1586676.

Waller, Isaac, and Ashton Anderson. 2021. "Quantifying Social Organization and Political Polarization in Online Platforms." *Nature* 600(7888):264–68. doi: 10.1038/s41586-021-04167-x.

Wojcieszak, Magdalena. 2009. "'Carrying Online Participation Offline'—Mobilization by Radical Online Groups and Politically Dissimilar Offline Ties." *Journal of Communication* 59(3):564–86. doi: 10.1111/j.1460-2466.2009.01436.x.

Zhuravskaya, Ekaterina, Maria Petrova, and Ruben Enikolopov. 2020. "Political Effects of the Internet and Social Media." *Annual Review of Economics* 12(Volume 12, 2020):415–38. doi: 10.1146/annurev-economics-081919-050239.