Online Communication Shifts in the Midst of the Covid-19 Pandemic: A Case Study on Snapchat

Qi Yang[1], **Weinan Wang**[2], **Lucas Pierce**[2], **Rajan Vaish**[2], **Xiaolin Shi**[2], **Neil Shah**[2]
[1] Massachusetts Institute of Technology
[2] Snap Inc.
yangqi@mit.edu, {wwang, lpierce, rvaish, xiaolin, nshah}@snap.com

Abstract

The Covid-19 pandemic has created large shifts in how people stay connected with each other in lieu of social distancing and isolation measures. More and more, individuals have turned to online communications as a necessary replacement for in-person interaction. Despite this, the research community has little understanding of how online communications have been influenced by the offline impacts of Covid-19. Our work touches upon this topic. Specifically, we study research questions around the impact of Covid-19 on online public and private sharing propensity, its influence on online communication homophily, and correlations between online communication and offline case severity in the United States. To do so, we study the usage patterns of 79 million US-based users on Snapchat, a large, leading mobile multimedia-driven social sharing platform. Our findings suggest that Covid-19 has increased propensity to privately communicate with friends, while decreasing propensity to publicly share content when users are out-and-about. Moreover, online communications have observed a marked decrease in baseline homophily across locations, ages and genders, with relative increases in cross-group communications. Finally, we observe that increased offline positive Covid-19 case severity in US states is associated with widening gaps between across-state and within-state communication increases after the onset of Covid-19, as well as marked declines in public sharing. We hope our findings drive further interest and work on online communication changes during pandemics and other extended times of crisis.

Introduction

The Covid-19 pandemic has created seismic shifts in people's lives, with profound economic, social, family, work, and school disruptions. To flatten the curve of Covid-19 cases and mitigate negative social and economic impacts, governments have put in place numerous restrictions on constituents regarding limits to in-person interaction, enforced self-quarantine, and social distancing measures. With these practices in place, friends, families, and colleagues have been forced to suddenly adopt or augment new or existing communication modalities, with most work, school and social communications happening exclusively online in many parts of the world to this day (Koeze and Popper 2020).

Despite the massive and wide-reaching shift in interpersonal interactions being moved to online ones, we as a community still have little knowledge about how people's online communication habits via social platforms have changed as a result of a sudden and critical externality, like Covid-19. For example, who are people talking to, and where? How has the severity of the pandemic influenced the interaction behavior? One might expect that the manner, content, and intent of communications might have changed substantially as well, due to concern for friends and family, reaching out in periods of isolation, dealing with mental health struggles, and more. Research on communication behaviors during pandemics to date, including early work on Covid-19, mainly focus on risk assessment and misinformation (Strekalova 2017; Larson 2018; Bursztyn et al. 2020); several prior works also studied communication on social media during natural disasters (Metaxa-Kakavouli, Maas, and Aldrich 2018; Palen and Anderson 2016a) like hurricanes and tropical storms, with attention to the utility of social platforms for information spreading and relief effort organization in the extreme short-term. However, none of the prior works touch upon interpersonal online communication shifts due to such a jarring externality.

Given the size and scope of the impacted population, building understanding of this topic is crucial: many individuals around the world have felt the apparent impacts of in-person restrictions and their implications for human interaction. However, there are numerous challenges in facilitating study of pandemic impact on online communications, making it challenging or impossible to study in the past: Firstly, pandemics are naturally rare events, severely limiting the timeframe of their investigation and would-be investigators. Secondly, previous pandemics have not occurred in the modern heyday of online and social platform-driven communications. Thirdly, investigation of online behaviors is challenging without appropriate data access and large enough scope of study.

In this work, we take advantage of a confluence of factors which allow us to overcome these issues, and enable our investigation of Covid-19's impact on private and public online communication patterns in the United States via Snapchat. Snapchat is a highly popular multimedia ephemeral messaging platform which launched in 2011, and has 238 million daily active users across the world (Snap

Inc. 2020). It offers functionality for both public and private sharing of *Snaps* (short images or videos), through *Direct Snaps* (private, one-to-one communications) and *Story Snaps* (broader audience, either to all a users' friends or to all users on the platform). In our study, we examine online communication habits as proxied by both modalities.

Our work aims to address three key research questions to better understand shifts in online communication patterns before and after the onset of Covid-19:

1. **How has Covid-19 impacted online private and public sharing propensity?** We find that post-onset Covid-19 engagement is higher for online direct communication and lower for online geo-based public communications. Temporal analysis shows the change is sudden for most states in the US, and the difference is statistically significant for all states.
2. **How has Covid-19 influenced homophily in online communications between users?** We find that social distancing measures have reduced effects of homophily and induce increases in inter-state/gender/age online geolocation-based public communications.
3. **Are changes in online communication patterns correlated to the severity of offline Covid-19 impact?** We find that the number of Covid-19 cases in different states is not correlated to increases in communication frequency, but is positively correlated to differences in within-state and across-state communication metrics.

We took a quantitative approach to study these questions. To this end, we analyzed the communication patterns of over 79 million US-based Snapchat users from February to May 2020. In the remainder of the paper, we investigate these questions and provide detailed answers and discuss implications based on our analyses. We hope our work helps elucidate how Covid-19 has influenced changes in online communications as a function of in-person restrictions, and can help inform design improvements for social platforms as a result.

Background and Related Work

We discuss prior work in four areas: social media's role in times of crisis, communication during Covid-19, homophily in social media, and background about the Snapchat platform and prior work on Snapchat. Our work interfaces with each of these aspects.

Social Media's Role in Times of Crisis

Due to its unpredictable and negative nature, an ongoing crisis often produces a high amount of uncertainty and anxiety among the public during a short period, potentially resulting in large scale damages (Coombs 2014). Prior work suggests that during these crises, social media usage increases (Ulvi et al. 2019), and the interplay between social media and crises has led to numerous prior works on the broad theme of crisis informatics (Palen and Anderson 2016b). Several works (Ulvi et al. 2019; Goolsby 2010; Hiltz, Diaz, and Mark 2011) investigate the role of social media in information dissemination, coordination and public awareness about crises and natural disasters. (Imran et al. 2015) also surveys mechanisms for information extraction and distillation from social media in times of crisis. Other works touch on cultural comparisons of communication in critical times: (Ding and Zhang 2010) studies compares institutional communications during the 2009-2010 H1N1 flu outbreak, while (Welhausen 2015) discusses intercultural risk communications through data visualization during the 2014 Ebola outbreak in West Africa. (Metaxa-Kakavouli, Maas, and Aldrich 2018) finds that online social ties play a critical and previously underestimated role in natural disaster preparedness. Higher levels of bridging and linking social ties correlate strongly with evacuation propensity. While these prior works mainly focus on public or health authority responses, and information distillation and distribution via social media in times of crisis, none study the fundamental, characteristic changes in underlying communication patterns on social media brought about by large-scale social distancing and new communication norms, as our work aims to.

Communication during Covid-19

With the ongoing Covid-19 pandemic, people have increased their social media usage to seek information about the pandemic according to surveys (Wiederhold 2020). The widespread effects of Covid-19 have led to several recent initiatives in studying its interplay with social media use: Kim (2020) collected comments from Korean social media to analyze negative emotions and societal problems during Covid-19. Lin, Liu, and Chiu (2020) used Google Keyword Search frequency to predict speed of Covid-19 spread in 21 countries/regions. Singh et al. (2020) characterizes Twitter conversation around Covid-19, and indicates that online conversation about the virus leads new cases geographically. Several prior works have also studied misinformation around Covid-19: (Depoux et al. 2020) remarks upon the rapidity of the panic and spread of misinformation. (Huynh et al. 2020) studies how Covid-19's risk perception in Vietnam is heavily mediated by baseline and geographical social media use. (Pennycook et al. 2020) found that many people disseminated false information related to the virus because they failed to reason appropriately if content was true or false before sharing, and that propensity to share was misaligned with people's ability to judge accuracy. (Brennen et al. 2020) indexes many common false claims about Covid-19 circulating on social media, and notes that the majority are misinformative (improper context, misleading) rather than disinformative (fabricated or imposter content).

Homophily in Social Media

Homophily is the principle that contact between similar people occurs at a higher rate than among dissimilar people. This principle has implications in information diffusion, grouping and community formation, online exposure and more (McPherson, Smith-Lovin, and Cook 2001). (Catanzaro, Caldarelli, and Pietronero 2004; Krivitsky et al. 2009; Shah 2020) study incorporation of homophily as a first-class citizen in network and graph modeling. (Guacho et al. 2018; Akoglu, Chandy, and Faloutsos 2013; Pandit et al. 2007) exploit homophilic principles in networks to detect misbehaving users and misinformative articles. Recently, several works (Bessi et al. 2015; Gillani et al. 2018; Kumar and Shah 2018) discuss homophily's role in echo-chamber formation, opinion polarization and misinformation spread in social media. Homophily can also lead to between-group segregation of interpersonal relations in teams (Lau and Murnighan 1998), and can occur due to formative effects and preferential selection (Currarini, Jackson, and Pin 2009). While our work does not directly model homophily, it empirically studies variation in homophilic effects pre and post onset of Covid-19.

The Snapchat Platform

Snapchat is a popular, mobile multimedia-driven social messaging platform, introduced in September 2011. As of July 2020, Snapchat has roughly 238 million daily active users and enjoys widespread use (Snap Inc. 2020). Snapchat enables users to create short image or video snippets, called *Snaps* which can be both narrowcast (directly shared privately with friends) as *Direct Snaps*, or broadcast (made publicly visible either to all friends or all other users on the platform) as *Story Snaps*. Juhász and Hochmair (2018) found that Snapchat users are more likely to share Snaps to everyone on the platform that are taken in highly trafficked areas, such as tourist hotspots or urban centers. Snaps can be further modified with geolocation-based filters, augmented reality (AR) lenses, stickers and more, adding metadata and context to the content (Verstraete 2016; Rios, Ketterer, and Wohn 2018). Snaps are ephemeral: Direct Snaps persist only until the recipient views them, after which they are deleted. Story Snaps are appended to a user's *Story* timeline, and automatically deleted 24 hours after posting. Several prior works study these features and associated user engagement: Bayer et al. (2016) notes that Snapchat users associate Snapchat communications with increased trust in the audience, and reduced self-curation due to ephemerality; this is unlike other platforms where content is pervasive and retained indefinitely, and promotes full curation of a singular external online profile (Uski and Lampinen 2016). Katz and Crocker (2015); Habib, Shah, and Vaish (2019); Juhász and Hochmair (2018) note that users' individual sharing decisions on Snapchat in private versus public spheres are influenced by various contextual factors associated with identity, activity, location and time of sharing. (Lamba and Shah 2019; Kaghazgaran et al. 2020) characterize and statistically model consumption behaviors of Direct Snaps and Story Snaps, respectively. Several works (Tang et al. 2020; Saha et al. 2021) also propose approaches to model ad response and user churn phenomena on Snapchat.

Data

We utilized rich engagement data spanning the time between February 15 to May 13 from Snapchat. We designate March 11 as the date threshold for partitioning our study period into pre and post Covid-19 timeframes, given that March 11th was the day on which the World Health Organization (WHO) declared the outbreak to be a pandemic[1]. Notably, March 11th also corresponds to the timeframe that many states started adopting stay-at-home orders. We do not consider dates prior to Feb 13th or post May 13th due to access limits (limited availability prior, and limited access post). We gathered user engagement metrics for 79 million Snapchat users.

We focused the study population to those whose primary/home location is in the United States, and have stayed in their home location before the pandemic (as observable from Snapchat usage logs). To achieve this goal, we first filtered our candidate population to users who were active around January 15, on which day the Center for Disease Control (CDC) reported the first Covid-19 case in the United States. We further selected users whose geolocation was in the U.S. (considering 51 "states," including the 50 conventional states, and DC) for all of our analyses. Lastly, we limited selection to users who had consistently reported locations in the same state in the past 1, 7, 30, and 90 days, thereby removing effects from "visitors" who were only active in the United States for a short period of time but active in another country for the majority. As a result, we obtained a representative set of residents for each state. Importantly, we held this population consistent across pre and post Covid-19 periods to limit exogenous effects from different samples (new or resurrected users).

Additionally, for Direct Snaps, we specifically gathered de-identified sender and recipient user ID, to evaluate the flow of conversations. Note that one Snap could be (privately) sent to multiple recipients, and create several parallel information flows. We also collect user attributes like (self-reported) gender and age for both users in the pair to evaluate homophilic tendencies, and evaluate location based on IP address. For Story Snaps, we specifically gathered de-identified poster user IDs, and attributes indicating whether the Story Snap met our location-based criteria.

Privacy, Ethics and Disclosure. Our work uses sensitive data from Snapchat. It is conducted within Snapchat, and reflects our commitment to user privacy. Our analysis relies on de-identified data, and throughout the work, we discuss only aggregated metrics across a large cohort of users.

Methods

To quantify the impact of Covid-19 on online engagement habits, we conducted several empirical comparisons of engagement before and after the onset of Covid-19; we call these "pre Covid-19" and "post Covid-19," informally.

To answer RQ1, we analyzed several key daily metrics on public and private sharing to test if social distancing affected private and public communications differently on Snapchat. For public communications, we considered metrics based on Story Snaps (SS), due to their broadcast functionality:

1. **Total SS from the state.** We considered SS from senders whose home location was in the state.

[1] https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen

2. **SS per poster.** We considered the ratio of total SS posted to posters to control for fluctuations in the number of posters.

Notably, we limited our focus to those SS which utilize geolocation-based overlay filters, AR lenses and other modifiers indicating that the user is out-and-about, and publicly sharing at a location of interest, further contrasting with private sharing norms. This is consistent with (Juhász and Hochmair 2018)'s findings on sharing while in public being characteristically different from that in private. Enforcing the above condition allows us to focus on the set of SS which are not taken when the user is "at home." For private sharing, we considered metrics based on Direct Snaps (DS), due to their one-to-one narrowcast functionality:

1. **Total DS from the state.** We considered all DS from senders whose home location was in the state. This was the most common and straightforward way to track online engagement.
2. **DS per sender.** We divided total DS by number of users to derive each user's DS change on average.
3. **Recipients per DS sender.** We considered the number of recipient users engaged with per DS senders. Number of recipients demonstrated the size of social network.
4. **DS tie strength.** We define DS tie strength as the ratio between DS sent and the number of recipients. Higher DS tie strength indicates indicates tighter, or more concentrated engagement. Note that we define this notion as a node-level descriptor, rather than an edge-level one.

The above metrics are key indicators for user engagement, reflecting both raw and normalized quantities, as well as engagement concentration. We calculated these metrics on a daily basis, and then aggregated them based on pre and post Covid-19 timeframes. We used two-sample t-tests to determine whether measured quantities in the pre and post Covid-19 timeframes differed, and if those differences were statistically significant. We choose t-tests assuming normality in the means of the pre and post daily metric quantities, and roughly equal variances. We further corrected the p-values with the Benjamini-Hochberg (BH) procedure for controlling False Discovery Rate (FDR) in multiple testing (Benjamini and Hochberg 1995). Moreover, to analyze the temporality of these changes, we also conducted change-point detection on the time series to test if engagement changed abruptly or gradually. For the purpose, we used the Pruned Exact Linear Time (PELT) search method (Killick, Fearnhead, and Eckley 2012) to determine the existence and location of the change point.

To answer RQ2, we gathered user demographic descriptors like location, gender and age for both source and destination users across many communication pairs, to gauge shifts in the pre and post Covid-19 settings. We hypothesized that in-person distancing measures would qualitatively impact the types of online communications rather than just the quantities, and use distributions of these user properties across pairs to evaluate the variation in homophily (communications between alike-users) in the pre and post periods. Specifically, for each of the primary factors we studied, including location (within state versus across state), age (same age group versus different age group), and gender (same gender and other gender), we used two-sample t-tests on the daily metric quantities to measure the difference and significance of the pre and post Covid-19 periods, and draw conclusions reflecting whether lockdowns encourage users to interact more with others that are similar or dissimilar.

For RQ3, we aimed to evaluate the relationship between offline severity of Covid-19, and online communication differences in public and private settings. Specifically, we collected statistics on Covid-19 cases in different states from (The COVID Tracking Project 2020), and used linear regression to evaluate the correlations between the two.

Results

Below, we discuss our findings for each of the 3 RQs.

Public and Private Sharing Propensity (RQ1)

First, we analyzed how private and public sharing propensities shifted with the onset of Covid-19 and associated distancing measures.

Private Sharing. We compared pre and post Covid-19 user engagement in several metrics based on Direct Snaps (DS), which has the one-to-one narrowcast functionality as discussed in details in *Methods*.

First, we calculated the mean of total DS for each state in the US, post Covid-19, as shown in Figure 1. The figure clearly shows that post Covid-19 private (DS) engagement is substantially higher for all states, ranging from a 9.7 to 25.5 percentage increase state-wise. Two-sample t-tests for each state also demonstrate $p < .05$ after BH correction, further confirming significant inequalities in the sample means in the pre and post Covid-19 periods. Recall that since our pre and post metrics are evaluated over a fixed user population, the normalized quantity (DS per sender) is also significant across pre and post periods. The geography of these changes is shown in Figure 2a.

Moreover, DS per sender and recipient per sender also have a significant increase after the onset of Covid-19. Post Covid-19 means are 8.1 to 24 percent higher for DS per sender, and 1.5 to 6.0 percent higher for recipient per sender, with more geographical detail in Figure 2b-2c. Two-sample t-tests for each state also demonstrate $p < .05$ after BH correction, which suggests that on average, the private communication volume and the social communication network size increases for each user.

To further investigate if the increase in DS is attributed to all friends, or just top contacted friend, we considered tie strength (DS per recipient). We found that post Covid-19 means are 5.3 to 18.6 percent higher, and the increment is significant for all states, with $p < .05$ after BH correction. Figure 2d illustrates the geographical change on the map. This result demonstrated that on average, users are deepening their friendships with all friends. In short, in-person distancing measures led to substantial online private communication increases.

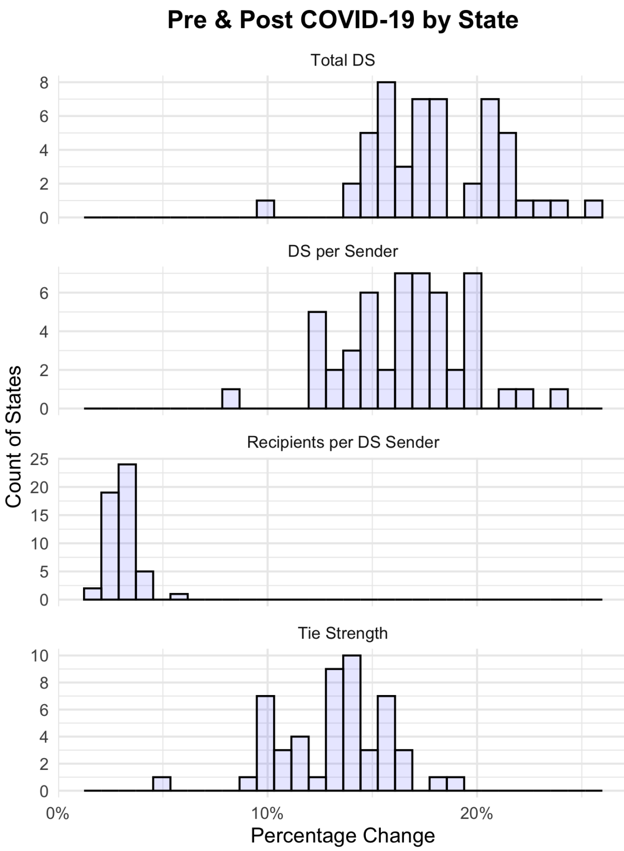
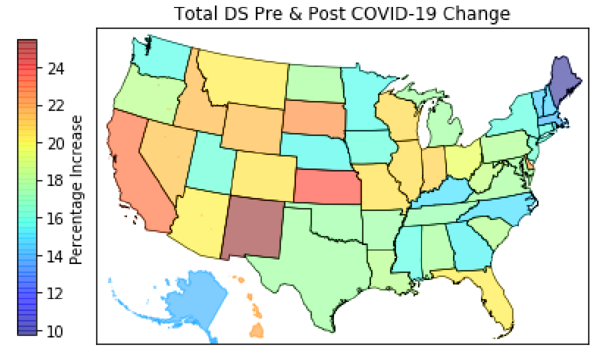


Figure 1: Percentage changes in private sharing (DS) across all the US states for several metrics indicate that online private sharing substantially increases (all $p < .05$).

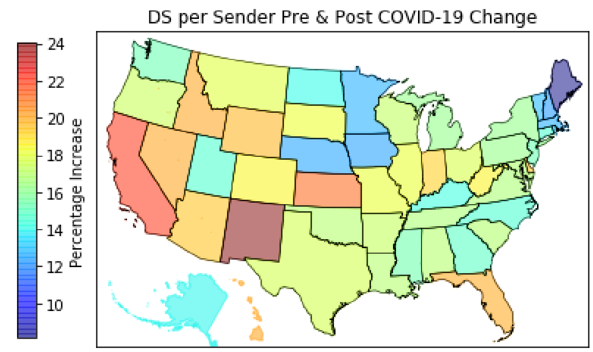
Geographical snapshots in Figure 2 show similarities across Total DS, DS per sender, and tie strength in highest increment states, with KS, CA and NM consistently showing highest % changes in these quantities, and ME with consistently low % changes. We observe some differing trends in Recipients per sender, which conveys more about communication breadth rather than depth like the other metrics; here, HI has the highest increase, perhaps owing to its disconnected status from the mainland.

Lastly, we performed change-point detection on each metric over the joint pre and post time period to test if engagement changed abruptly or gradually, as shown in Figure 3. We found that overwhelmingly, 49 states experienced a sudden surge in total DS and DS per sender, 47 in recipients per DS sender, as well as 48 in tie strength.

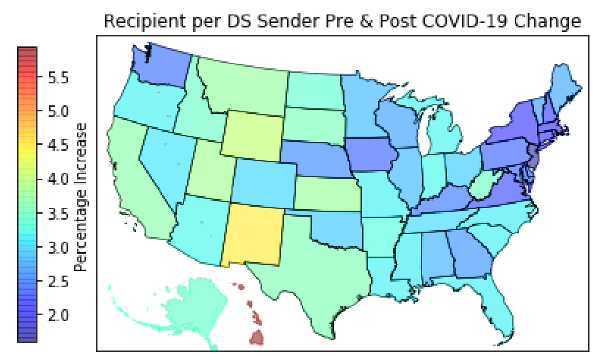
Public Sharing. We also compared pre and post Covid-19 user engagement as measured by location-based Story Snaps (SS). Figure 4 shows the relative percentage change in total SS for each state in the US, post Covid-19. Clearly, these metrics drop consistently across states, ranging from a -78.98 to -35.31 percentage decrease, indicating the limited mobility of users and desire to share content publicly due to distancing and isolation measures. Two-sample t-tests



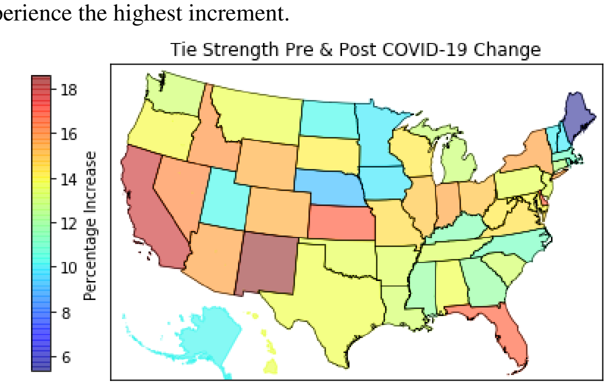
(a) ME, MA, and AK have the lowest, whereas CA, KS, and NM experience the highest increment.



(b) ME, NH, IA have the lowest, whereas KS, CA, and NM experience the highest increment.



(c) NJ, MA, and CT have the lowest, whereas WY, NM, and HI experience the highest increment.



(d) ME, NE, and IA have the lowest, whereas KS, CA, and NM experience the highest increment.

Figure 2: % changes in private sharing (DS) on the US map.

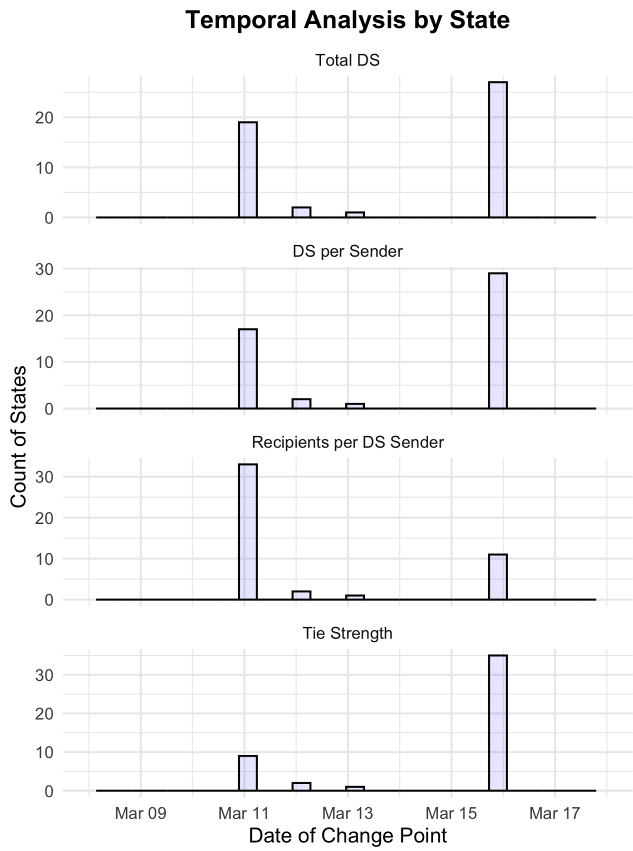


Figure 3: Change point detection in private sharing (DS) across all US states for several metrics indicate that online private sharing experienced a surge for most states.

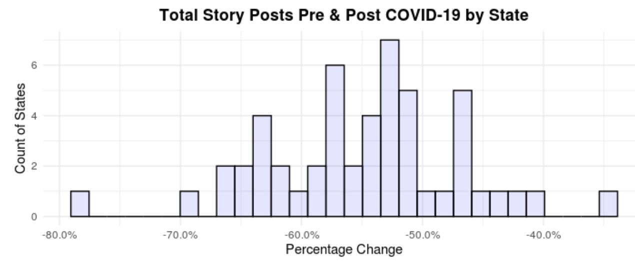


Figure 4: Percentage changes in public posting (SS) across all US states indicate that online location-based public sharing substantially decreases (all $p < .05$); post Covid-19 means are -78.98 to -35.31 percent lower.

for each state again confirm the significance of these effects, with all $p < .05$ after BH correction. These effects were also apparent in normalized per-poster quantities (not shown due to space constraints). Moreover, a sudden change point was found in all 51 states: 45 states on March 16, and 6 states on March 21, indicating an abrupt and significant change in the public engagement.

Variation in Homophilic Tendencies (RQ2)

Next, we consider how homophilic communication tendencies (baseline rates of within-group communications) shifted post Covid-19. We analyzed communication pattern changes in three aspects: location (within state vs. across states), age (within age-group vs. across age-groups) and gender (within gender vs. across genders). We use variations in the DS tie strength to reason about change in homophily, by comparing changes in the difference in within-group and across-group DS tie strengths (we offer further discussion on the use of this metric below, for *Location*).

Location. First, we considered within-state and across-state private communications. Before considering the DS tie strengths, we considered the two metrics contributing to the ratio: private communication volume (DS sent) and social network size (recipients per sender). As previously discussed, Figure 1 shows that private (user-user) communication volume increased for all states, but does not indicate the manner of this increase.

Thus, we evaluated the total DS by state, pre and post Covid-19, broken down by across-state and within-state communications. Figure 5 shows the absolute increase in total private sharing (DS) of within-state (red) and across-state (blue). In general, private communication quantities (DS) increased both within-state (red) and across-state (blue), and for most states, raw within-state communication volume increases outpaced across-state volume increases, due to their large baseline propensity (the majority of communications pre Covid-19 were within-state). However, upon considering the relative social network sizes of within-state and across-state groups, we see a different portrayal of the effect: Figure 6 shows that the across-state recipients per sender (blue) increased substantially for all states, while the within-state quantity (red) increased only for some states but decreased for others. So, while the volume of DS sent increased more within-state, the recipients per sender increased more across-state.

To study these contrasting indications more carefully, we consider the DS tie strength ratio, which more concisely summarizes the depth of the relationship that a sender has with a set of recipients. In this case, we considered both the across-state and within-state DS tie strengths (adjusting the numerator and denominator to account only for across-state and within-state interactions, respectively). Interestingly, 44 states have higher increase in across-state than within-state tie strength, of which 22 of the increases are significantly higher ($p < .05$) when evaluate with a two-sample t-test with BH correction. This suggests that while both within and across-state tie strengths increased post Covid-19 (suggesting deepening relationships), users actually deepened across-state relationships *moreso* than within-state ones. We conclude that although relationships within-state also deepened post Covid-19, larger deepening of across-state relationships suggests a relative reduction of location-based homophily from the pre-Covid baseline.

Age. Next, we considered communications between users within and across age-groups. We consider users partitioned into 5 age groups: 13-17, 18-20, 21-24, 25-34 and 35 plus.

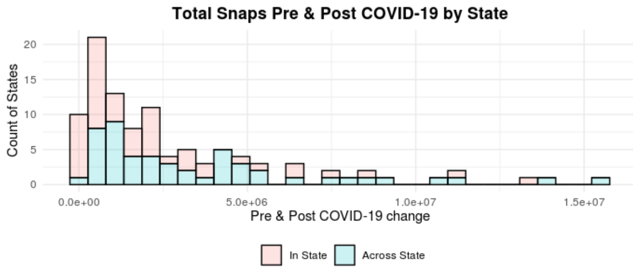


Figure 5: Raw increase in total private sharing (DS) of within-state (red) and across-state (blue). Within-state DS increases outsizes across-state DS increase for most states.

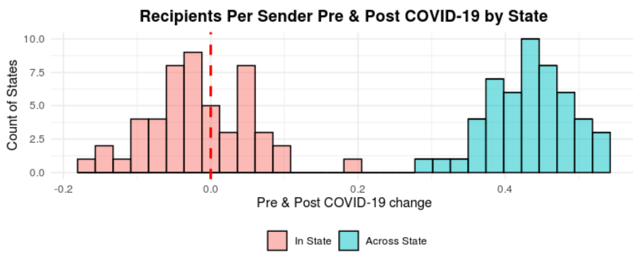


Figure 6: Social network size (measured by recipients per sender) consistently grows for across-state communications (blue), compared to mixed effects for in-state communications (red), indicating a reduction of location-based homophily and promotion of cross-location diversity.

We use the DS tie strength metric to measure the communication intensity between the groups.

Figure 7 shows absolute increases in DS tie strength post Covid-19 between associated sender and correspondent (receiver) users, computed by subtracting the pre Covid-19 from the post Covid-19 metric. Darker shades of blue denote larger increases. Note that conditioning on each sender age group (vertical), the darker cells are those corresponding to different age-grouped users. For example, considering 13-17 aged senders, the communication increase was 0.534 within age-group, compared to the much more substantial 2.975 increase to 25-34 correspondents. Likewise, the communication increase from 35 plus age-group senders was 0.072 within-group, compared to 0.482 to 13-17 correspondents. Note that these quantities are deltas in the normalized DS tie-strength ratio; a 1 unit change is extremely large, indicating that on average, users send 1 more DS to all of their friends. Two-sample t-tests with BH correction confirmed that tie strength means were significantly different across all age groups pre and post Covid-19 (i.e. quantities in all cells of Figure 7 are significant). Moreover, we also conducted two-sample t-tests with BH correction to evaluate the difference-in-difference measurements (Lechner et al. 2011). Specifically, we considered the difference across age groups, in difference in DS tie strength pre and post Covid-19 between one's own age-group and other age-groups (i.e. cells on the diagonal, compared to cells off the diagonal), to evaluate e.g. whether the increases in tie strength from

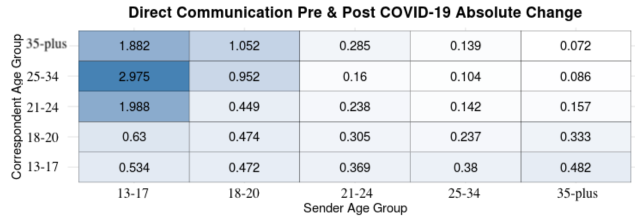


Figure 7: Absolute increases in private sharing (DS tie strength) between different age groups pre and post Covid-19 indicate reduction in age-group homophily. Users deepen communications both within and across age-groups, and seemingly moreso in the latter setting.

	13-17	18-20	21-24	25-34	35+
13-17		✓	✓	✓	✓
18-20				✓	✓
21-24	✓			✓	✓
25-34	✓		✓		✓
35+	✓	✓	✓	✓	

Table 1: Two-sample t-test significance results on the difference of difference in DS tie strength between "within age-group" and "across age-group" categories. Most results indicate across age-group increases are significantly different (✓ indicates $p < .05$) larger than within age-group ones.

13-17 senders to 13-17 correspondents are indeed significantly different to 18-20 correspondents (in other words, to evaluate whether the across-group effect is larger than the within-group effect). Table 1 shows the significance results (✓ indicates $p < .05$) for these difference-in-difference experiments, clearly indicating the effect is present and observable for the majority of age-group pairs. These results altogether suggest a reduction of age-group homophily from the pre Covid-19 baseline.

Gender. Thus far, we observed that the post Covid-19 period marks an observed reduction of homophily from both the location and age lens. We next consider whether this effect holds across gender as well, using within and across-gender DS tie strengths.

We found that in general, users have deeper DS tie strengths with the opposite gender pre Covid-19 (i.e. males have larger tie strengths to females on Snapchat, and vice versa). Post Covid-19, all 4 groups (MM, MF, FM, FF) saw statistically significant ($p < .05$ after BH correction) increases in tie strength (i.e. quantities in all cells of Figure 8 are significant). Moreover, communication with the opposite gender increased even more than with the same genders for both female and male senders (comparing cells in the same vertical). We conducted two-sample t-tests to evaluate the difference-in-difference (difference across gender, in difference in DS tie strength pre and post Covid-19) between one's same gender and the opposite gender. We found that the increases in tie strength for MF is indeed significantly larger than that for MM, and likewise increase in FM is sig-

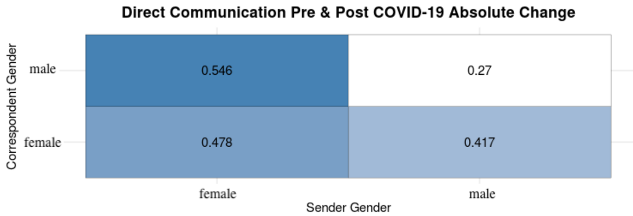


Figure 8: Absolute increases in private sharing (DS tie strength) between different gender groups pre and post Covid-19 indicate reduction in gender-group homophily. Users deepen communications both genders, and seemingly moreso with the opposite gender.

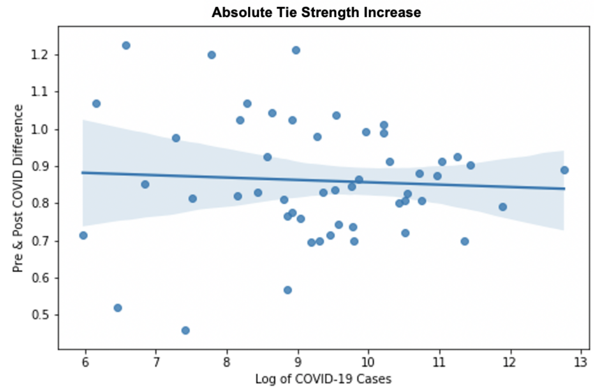


Figure 9: Offline Covid-19 case severity is not significantly correlated with online private sharing (DS) tie strength changes across states pre and post Covid-19.

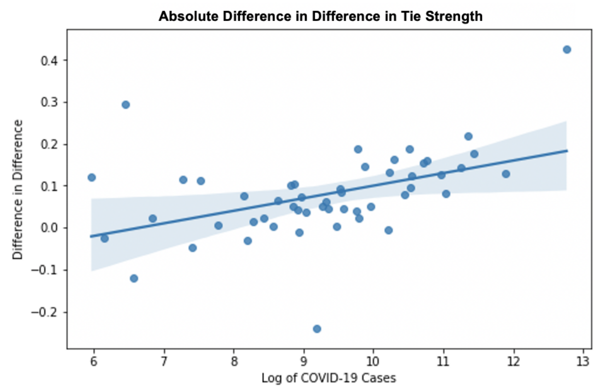


Figure 10: Offline Covid-19 case severity is significantly positively correlated with difference-in-difference (across-state minus within-state) measurements of online private sharing (DS) tie strength changes pre and post Covid-19. More Covid-19 cases is associated with larger margins between across-state and within-state tie strengths.

nificantly larger than that of FF (both $p < .05$).

Specifically, MF increased significantly more than MM, and FM significantly more than FF, indicating a further reduction of gender-based homophily in an already heterophilic regime.

Correlation with Covid-19 Case Severity (RQ3)

Private Sharing. Although Covid-19 is an international emergency, its impacts have been disparate across locations. In the US, while some states saw more severe outbreaks and announced early lockdown orders, some have not observed the same and are more "under control" from an offline (on-ground) standpoint. Those affected by more severe distancing measures may be communicating more or less online than others: thus, we study the relationship between offline severity and online impact.

We consider DS tie strength as the target metric for evaluating engagement depth. If the tie strength increases, it indicates that users send more DS to their social networks, and are thus communicating more closely with their friends. We use the positive case count on May 16, 2020 as a measure of offline/on-ground Covid-19 severity, i.e. higher positive case count implies higher severity. Technically, we use the logarithm of the measure (monotonically increasing with respect to the actual case count), due to its large scale.

Firstly, we evaluated whether the offline severity was correlated with increase or decrease in tie strength between pre and post Covid-19 periods. Figure 9 shows that two quantities are not significantly correlated, and that tie strength increases in all states occur despite (or without regard to) the case severity.

Next, we considered the difference-in-difference of tie strength for within-state communication and across-state communication measurements. Figure 10 shows a positive relationship between case severity and difference-in-difference measurement (across-state minus within-state), which indicates that increased case severity is correlated with increased communication across-state. A significant regression equation was found ($p < 0.001$) with $R^2 = 0.2$. The prediction of difference-in-difference of "tie strength" is equal to $-1.994 + 0.03(\log \text{ of Covid}-19 \text{ cases})$. In other words, for every unit increase in log of Covid-19 cases, there is a 0.03 increase in the difference-in-difference measurement in Snaps per recipient. This significant positive correlation not only evidences the association between case severity and online communication, but it also further substantiates that distancing effects contribute to a reduction of location-based homophily (as in our result for RQ2).

Public Sharing. While social distancing is positively correlated with the increase in private sharing, we also ask whether it correlates with the magnitude of the drop in location-based public sharing. Figure 11 shows a positive relationship between case severity and the drop in public story posting (SS), which indicates that increased Covid-19 case severity is associated with a larger drop in public sharing. A significant regression equation was found ($p < 0.01$) with $R^2 = 0.13$. The prediction of percentage drop in story posting is equal to $53.3 + 5.5e{-}5(\log \text{ of Covid}-19 \text{ cases})$. In other words, for every unit increase in log of Covid-19 cases, the drop in story posting grows by $5.5e{-}5\%$.

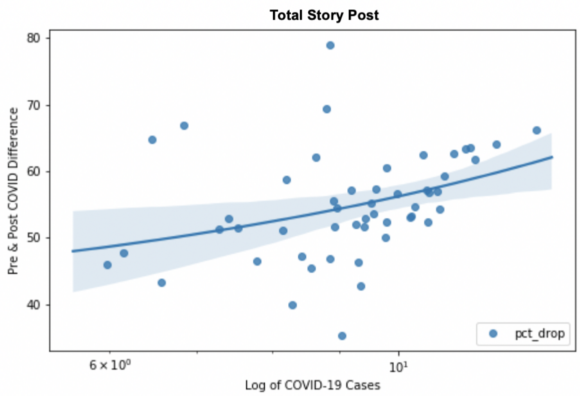


Figure 11: Offline Covid-19 case severity is significantly positively correlated with drops in online public sharing (SS) pre and post Covid-19. More Covid-19 cases is associated with larger reduction in public sharing activity.

Discussion and Conclusion

In this work, we quantified the impact of Covid-19 on online engagement habits through various angles. First, we found that post-onset Covid-19 engagement is higher for private sharing and lower for location-based public sharing. This finding reflected that after the onset of Covid-19, due to stay-at-home orders and other quarantine policies, people reduced their time outside and increased their communications on social media platforms. In particular, as Juhász and Hochmair (2018) found that Snapchat users share public Snaps from highly trafficked areas, such as tourist hotspots or urban centers. We postulate that when these centers temporarily closed, it affected the foot traffic negatively, and further intensified people's discomfort with being in popular public areas, therefore decreasing location-based public sharing. This crisis highlights the particular strengths of social media especially when in-person interactions are limited. Since many people cannot connect with their friends and family in person, for the time being (and potentially longer), social media has become an even more dominant means of maintaining valued connections.

Second, we found that lockdowns temporarily reduced effects of homophily and induced increases in across-state/gender/age-group online communications. As social distancing measures were put into place, most relationships became effectively the same "online distance" (just a Snap) away. Moreover, people realized how important it is to stay in contact with their friends and family. Both of the reasons presumably led to increased diversity in online communications. For instance, due to the rise in Covid-19 cases, students returned home from colleges and stayed connected with friends across the country. Due to stay-at-home orders, people couldn't meet family members and relatives (of all age groups and gender) regularly, so they relied on online communication to check on each other. There are many such instances that got people to step out of their usual social circle and bond with those outside it. Presumably, our observations capture both a flattening of multiple social circles (that many previously have been a mixture of in-person and online interactions) for preservation of routine communication and deepening, as well as a resurrection of previously less accessible relationships (far-away friends, colleagues and relatives).

Third, we concluded that the number of Covid-19 cases in different states is not correlated to the increase in private sharing frequency, but is significantly positively correlated to the difference in private sharing metrics of within-state and across-state communications. Summarily, Covid-19 cases do not affect increment in private sharing directly, but rather, the more severe pandemic is in a state, the more the reduction in location-based homophily is. Moreover, Covid-19 cases is also positively correlated with the reduction in location-based public sharing. These are surprising and interesting results. This is likely due to the effectiveness of stay-at-home orders and other social distancing measures. Overall, since the number of cases, testing access, response measures and enforcements varied from state to state in the US, the effects of the Covid-19's severity on communication patterns were not uniformly consistent. Generally, we believe that higher case-severity likely corresponds to higher public panic and more stringent distancing and isolation restrictions, leading to stronger effects where observed.

Limitations. The conclusions of this work are limited to US population, and may not generalize to other countries who had differing pandemic responses and lesser degrees of distancing. Moreover, our work was conducted using data from Snapchat, which offers a significant, but not comprehensive view of online communications – Snapchat's user population skews younger, and female, for example. Additionally, our study time period is subject to limitations of platform data access, integrity (i.e. user-misreported information) and availability, suggesting the value of more longitudinal work on the persistence, increase or decrease of the observed effects with the evolving response to Covid-19. Furthermore, our analysis with respect to on-ground case severity is impacted by the inconsistencies and challenges in measurement, detection and response imposed by external factors. Lastly, our analysis does not differentiate results across states. Future work can build on ours by carefully positioning findings with respect to diverse policy responses; the lack of to-date standardized data to quantify such complex policies and their public adherence makes this task non-trivial, and ripe for future work.

Impact. Overall, our work contributes to a more profound understanding of how Covid-19 has, and continues to influence online human behaviors, and moreover how online platforms provide an alternative foundation to support human connection during the pandemic. Notably, humans are inherently social. In times of physical distancing, they are inclined to compensate their social needs via online measures. Our findings shed light on how Covid-19's on-ground impact and associated in-person distancing and isolation measures influenced communication volumes and propensity differences in private and public sharing behaviors, variation and reduction in homophilic baselines, and correlate to the magnitude of such shifts. We hope that our study provides valuable and timely insights for other researchers, and inspires further work on longer-term impacts and communication changes as a result of a post Covid-19 world.

References

Akoglu, L.; Chandy, R.; and Faloutsos, C. 2013. Opinion fraud detection in online reviews by network effects. *ICWSM* 13(2-11): 29.

Bayer, J. B.; Ellison, N. B.; Schoenebeck, S. Y.; and Falk, E. B. 2016. Sharing the small moments: ephemeral social interaction on Snapchat. *Information, Communication & Society* 19(7): 956–977.

Benjamini, Y.; and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)* 57(1): 289–300.

Bessi, A.; Petroni, F.; Del Vicario, M.; Zollo, F.; Anagnostopoulos, A.; Scala, A.; Caldarelli, G.; and Quattrociocchi, W. 2015. Viral misinformation: The role of homophily and polarization. In *WWW*, 355–356.

Brennen, J. S.; Simon, F.; Howard, P. N.; and Nielsen, R. K. 2020. Types, sources, and claims of Covid-19 misinformation. *Reuters Institute* 7: 3–1.

Bursztyn, L.; Rao, A.; Roth, C.; and Yanagizawa-Drott, D. 2020. Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2020-44).

Catanzaro, M.; Caldarelli, G.; and Pietronero, L. 2004. Assortative model for social networks. *Physical review e* 70(3): 037101.

Coombs, W. T. 2014. *Ongoing crisis communication: Planning, managing, and responding*. Sage Publications.

Currarini, S.; Jackson, M. O.; and Pin, P. 2009. An economic model of friendship: Homophily, minorities, and segregation. *Econometrica* 77(4): 1003–1045.

Depoux, A.; Martin, S.; Karafillakis, E.; Preet, R.; Wilder-Smith, A.; and Larson, H. 2020. The pandemic of social media panic travels faster than the COVID-19 outbreak.

Ding, H.; and Zhang, J. 2010. Social media and participatory risk communication during the H1N1 flu epidemic: A comparative study of the United States and China. *China Media Research* 6(4): 80–91.

Gillani, N.; Yuan, A.; Saveski, M.; Vosoughi, S.; and Roy, D. 2018. Me, my echo chamber, and I: introspection on social media polarization. In *Proceedings of the 2018 World Wide Web Conference*, 823–831.

Goolsby, R. 2010. Social media as crisis platform: The future of community maps/crisis maps. *ACM Transactions on Intelligent Systems and Technology (TIST)* 1(1): 1–11.

Guacho, G. B.; Abdali, S.; Shah, N.; and Papalexakis, E. E. 2018. Semi-supervised content-based detection of misinformation via tensor embeddings. In *ASONAM*, 322–325. IEEE.

Habib, H.; Shah, N.; and Vaish, R. 2019. Impact of Contextual Factors on Snapchat Public Sharing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13.

Hiltz, S. R.; Diaz, P.; and Mark, G. 2011. Introduction: Social media and collaborative systems for crisis management.

Huynh, T. L.; et al. 2020. The COVID-19 risk perception: A survey on socioeconomics and media attention. *Econ. Bull* 40(1): 758–764.

Imran, M.; Castillo, C.; Diaz, F.; and Vieweg, S. 2015. Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)* 47(4): 1–38.

Juhász, L.; and Hochmair, H. H. 2018. Analyzing the spatial and temporal dynamics of Snapchat. In *AnaLysis, Integration, Vision, Engagement (VGI-ALIVE) Workshop*.

Kaghazgaran, P.; Bos, M.; Neves, L.; and Shah, N. 2020. Social Factors in Closed-Network Content Consumption. *CIKM* .

Katz, J. E.; and Crocker, E. T. 2015. Selfies— selfies and photo messaging as visual conversation: Reports from the United States, United Kingdom and China. *International Journal of Communication* 9: 12.

Killick, R.; Fearnhead, P.; and Eckley, I. A. 2012. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* 107(500): 1590–1598.

Kim, B. 2020. Effects of social grooming on incivility in COVID-19. *Cyberpsychology, Behavior, and Social Networking* .

Koeze, E.; and Popper, N. 2020. The Virus Changed the Way We Internet. *The New York Times* URL https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html.

Krivitsky, P. N.; Handcock, M. S.; Raftery, A. E.; and Hoff, P. D. 2009. Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. *Social networks* 31(3): 204–213.

Kumar, S.; and Shah, N. 2018. False information on web and social media: A survey. *arXiv preprint arXiv:1804.08559* .

Lamba, H.; and Shah, N. 2019. Modeling dwell time engagement on visual multimedia. In *KDD*, 1104–1113.

Larson, H. J. 2018. The biggest pandemic risk? Viral misinformation. *Nature* 562(7726): 309–310.

Lau, D. C.; and Murnighan, J. K. 1998. Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of management review* 23(2): 325–340.

Lechner, M.; et al. 2011. *The estimation of causal effects by difference-in-difference methods*. Now.

Lin, Y.-H.; Liu, C.-H.; and Chiu, Y.-C. 2020. Google searches for the keywords of "wash hands" predict the speed of national spread of COVID-19 outbreak among 21 countries. *Brain, Behavior, and Immunity* .

McPherson, M.; Smith-Lovin, L.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 27(1): 415–444.

Metaxa-Kakavouli, D.; Maas, P.; and Aldrich, D. P. 2018. How social ties influence hurricane evacuation behavior. *Proceedings of the ACM on Human-Computer Interaction* 2(CSCW): 1–16.

Palen, L.; and Anderson, K. M. 2016a. Crisis informatics—New data for extraordinary times. *Science* 353(6296): 224–225. ISSN 0036-8075. doi:10.1126/science.aag2579. URL https://science.sciencemag.org/content/353/6296/224.

Palen, L.; and Anderson, K. M. 2016b. Crisis informatics—New data for extraordinary times. *Science* 353(6296): 224–225.

Pandit, S.; Chau, D. H.; Wang, S.; and Faloutsos, C. 2007. Netprobe: a fast and scalable system for fraud detection in online auction networks. In *WWW*, 201–210.

Pennycook, G.; McPhetres, J.; Zhang, Y.; Lu, J. G.; and Rand, D. G. 2020. Fighting COVID-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention. *Psychological science* 31(7): 770–780.

Rios, J. S.; Ketterer, D. J.; and Wohn, D. Y. 2018. How users choose a face lens on Snapchat. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 321–324.

Saha, K.; Liu, Y.; Vincent, N.; Chowdhury, F. A.; Neves, L.; Shah, N.; and Bos, M. W. 2021. AdverTiming Matters: Examining User Ad Consumption for Effective Ad Allocations on Social Media. *CHI* .

Shah, N. 2020. Scale-Free, Attributed and Class-Assortative Graph Generation to Facilitate Introspection of Graph Neural Networks. *KDD Mining and Learning with Graphs* .

Singh, L.; Bansal, S.; Bode, L.; Budak, C.; Chi, G.; Kawintiranon, K.; Padden, C.; Vanarsdall, R.; Vraga, E.; and Wang, Y. 2020. A first look at COVID-19 information and misinformation sharing on Twitter. *arXiv preprint arXiv:2003.13907* .

Snap Inc. 2020. Snap Inc. Q2 2020 Earnings. https://investor.snap.com/events-and-presentations/events.

Strekalova, Y. A. 2017. Health Risk Information Engagement and Amplification on Social Media: News About an Emerging Pandemic on Facebook. *Health Education & Behavior* 44(2): 332–339. doi:10.1177/1090198116660310. URL https://doi.org/10.1177/1090198116660310. PMID: 27413028.

Tang, X.; Liu, Y.; Shah, N.; Shi, X.; Mitra, P.; and Wang, S. 2020. Knowing your FATE: Friendship, Action and Temporal Explanations for User Engagement Prediction on Social Apps. *KDD* .

The COVID Tracking Project. 2020. The COVID Tracking Project. URL https://covidtracking.com/.

Ulvi, O.; Lippincott, N.; Khan, M. H.; Mehal, P.; Bass, M.; Lambert, K.; Lentz, E.; and Haque, U. 2019. The role of social and mainstream media during storms .

Uski, S.; and Lampinen, A. 2016. Social norms and self-presentation on social network sites: Profile work in action. *New media & society* 18(3): 447–464.

Verstraete, G. 2016. It's about time. disappearing images and stories in Snapchat. *Image & Narrative* 17(4).

Welhausen, C. A. 2015. Visualizing a non-pandemic: Considerations for communicating public health risks in intercultural contexts. *Technical Communication* 62(4): 244–257.

Wiederhold, B. K. 2020. Social media use during social distancing. *Cyberpsychology, Behavior, and Social Networking* 23(5): 275–276.