DISCUSSION PAPER SERIES

DP15593

SOCIAL DISTANCING DURING A PANDEMIC - THE ROLE OF FRIENDS

Michael Bailey, Drew Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel and Johannes Ströbel

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

SOCIAL DISTANCING DURING A PANDEMIC - THE ROLE OF FRIENDS

Michael Bailey, Drew Johnston, Martin Koenen, Theresa Kuchler, Dominic Russel and Johannes Ströbel

Discussion Paper DP15593
Published 22 December 2020
Submitted 12 December 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

SOCIAL DISTANCING DURING A PANDEMIC - THE ROLE OF FRIENDS

Abstract

We explore how social network exposure to COVID-19 cases shapes individuals' social distancing behavior during the early months of the ongoing pandemic. We work with de-identified data from Facebook to show that U.S. users whose friends live in areas with worse coronavirus outbreaks reduce their mobility more than otherwise similar users whose friends live in areas with smaller outbreaks. The effects are quantitatively large: a one standard deviation increase in friend-exposure to COVID-19 cases early in the pandemic results in a 1.2 percentage point increase in the probability that an individual stays home on a given day. As the pandemic progresses, changes in friend-exposure drive changes in social distancing behavior. Given the evolving nature and geography of the pandemic --- and hence friend-exposure --- these results rule out many alternative explanations for the observed relationships. We also analyze data on public posts and membership in groups advocating to "reopen" the economy to show that our findings can be explained by friend-exposure raising awareness about the risks of the disease and inducing individuals to participate in mitigating public health behavior.

JEL Classification: I0

Keywords: Social Networks, peer effects, COVID-19, Social distancing

Michael Bailey - mcbailey@fb.com
Facebook

Drew Johnston - drewmjohnston1@gmail.com
Harvard University

Martin Koenen - martin_koenen@g.harvard.edu
Harvard University

Theresa Kuchler - tkuchler@stern.nyu.edu
Leonard N Stern School of Business and CEPR

Dominic Russel - drussel@stern.nyu.edu
NYU Stern

Johannes Ströbel - jstroebe@stern.nyu.edu
Stern School of Business and CEPR

Acknowledgements
This research was facilitated through a research consulting agreement between the academic authors and Facebook. Bailey is an employee at Facebook.

SOCIAL DISTANCING DURING A PANDEMIC: THE ROLE OF FRIENDS*

MICHAEL BAILEY†
DREW JOHNSTON‡
MARTIN KOENEN§
THERESA KUCHLER¶
DOMINIC RUSSEL||
JOHANNES STROEBEL**

Abstract

We explore how social network exposure to COVID-19 cases shapes individuals' social distancing behavior during the early months of the ongoing pandemic. We work with de-identified data from Facebook to show that U.S. users whose friends live in areas with worse coronavirus outbreaks reduce their mobility more than otherwise similar users whose friends live in areas with smaller outbreaks. The effects are quantitatively large: a one standard deviation increase in friend-exposure to COVID-19 cases early in the pandemic results in a 1.2 percentage point increase in the probability that an individual stays home on a given day. As the pandemic progresses, *changes* in friend-exposure drive *changes* in social distancing behavior. Given the evolving nature and geography of the pandemic — and hence friend-exposure — these results rule out many alternative explanations for the observed relationships. We also analyze data on public posts and membership in groups advocating to "reopen" the economy to show that our findings can be explained by friend-exposure raising awareness about the risks of the disease and inducing individuals to participate in mitigating public health behavior.

JEL Codes: I0, D83, D85, H0
Keywords: Social networks, peer effects, COVID-19, social distancing

*This draft: December 12, 2020. This research was facilitated through a research consulting agreement between the academic authors and Facebook. Bailey is an employee at Facebook.

†Facebook, Inc: mcbailey@fb.com.
‡Harvard University: drewjohnston@g.harvard.edu
§Harvard University: martin_koenen@g.harvard.edu
¶Stern School of Business, New York University, NBER, CEPR: tkuchler@stern.nyu.edu.
|| Stern School of Business, New York University: drussel@stern.nyu.edu.
**Stern School of Business, New York University, NBER, CEPR: johannes.stroebel@nyu.edu.

1 Introduction

The ongoing COVID-19 pandemic has lead to an unparalleled global health crisis. As of November 14, 2020, more than 53 million reported cases and 1.3 million deaths worldwide have heavily burdened health systems across the globe. Given the threats posed by the disease, policy makers have imposed substantial restrictions on public life, and public health organizations have urged individuals to avoid nonessential travel, wear masks, and practice social distancing. However, despite significant evidence that these behavioral adjustments impede the transmission of the virus, there has been widespread variation in the extent to which individuals have adjusted their behaviors. As a result, understanding the different factors that determine individuals' decisions to adjust their behaviors to reduce the spread of the virus is central to designing an effective pubic health response to the COVID-19 pandemic.

One potentially important factor in explaining public health behavior during a pandemic is the information that individuals receive from their social networks about the seriousness of the disease. Such information can be complementary to (and may be more trusted than) communications from health officials or the news media. Especially close contacts, such as friends and family, may play important roles in shaping individuals' perceptions of the need to adjust their health behaviors. Indeed, prior studies show that friends can influence whether an individual becomes obese, smokes regularly, and chooses to get vaccinated (Christakis and Fowler, 2007, 2008; Sato and Takasaki, 2019).

In this paper, we study the effects of network exposure to COVID-19 cases on health behavior during the ongoing pandemic, focusing on the roles played by friends. We work with de-identified data from Facebook, a large online social networking service. These data allow us to analyze health behaviors and beliefs through providing information on (i) movement patterns, (ii) public posts on the platform, and (iii) membership in public Facebook groups. The first of these outcomes is a measure of social distancing behavior, while the latter two outcomes can provide information on beliefs about COVID-19 and related public health measures and can therefore shed light on potential mechanisms of the effects of social networks. Relative to other work that has explored social distancing behavior using cell phone location data, our data is unique in that it allows us to link these health behaviors to both individual-level demographics and information on social networks. It can therefore provide insights into health behavior and perceptions unavailable to other researchers, and thus has the potential to inform more effective public health interventions. In addition, the Facebook data is highly representative of the U.S. population and has been used extensively in prior research to study the determinants and effects of social networks (e.g. Bailey et al., 2018a, 2019a,c).

We begin by exploring aggregate time-series patterns in mobility, and find—consistent with prior work—that, during the early stages of the pandemic, U.S. Facebook users drastically reduced their mobility relative to before the outbreak of the pandemic. In mid-February, the probability of staying home averaged around 18% on a given day; by late March, this probability had increased to about 30%.[1] We find that highly educated users, women, older users, and users living in high-income areas reduced their mobility more than others.

[1]Our mobility sample only includes Facebook users who consented to sharing and storing their location information. We proxy for staying at home with staying within a single level-16 Bing tile, an area corresponding to about 600m x 600m at the equator. For more details regarding our mobility sample and variable construction, see Section 2.1.

We then look at the role of friend networks in shaping distancing behaviors. We find that users with greater friend-exposure to COVID-19 cases—that is, those who have relatively more friends living in areas highly affected by the virus—are substantially more likely to reduce their mobility. To isolate this relationship, we take both a static as well as a dynamic approach. In the static approach, which focuses on the early pandemic, we first classify users by whether their friend-exposure to COVID-19 is higher or lower than the median of users *within* their zip code as of March 15, right after President Trump declared a national emergency. The average movement patterns of the two groups look strikingly similar prior to the pandemic, before diverging strongly after the outbreak begins. Using an event study design that controls for the time-varying effect of a number of relevant user characteristics that might be correlated with friend-exposure to COVID-19 cases, we estimate that this above-median friend-exposure to COVID-19 early in the pandemic is associated with a 1 percentage point increase in a user's likelihood of staying at home on a given day during the pandemic. Interestingly, this effect persists through April and May, consistent with existing survey evidence showing that individuals in areas with early COVID-19 spread remained the most worried about the virus in later months (Blumenthal, 2020). Quantitatively, a one standard deviation increase in friend-exposure to COVID-19 cases is associated with a 1.2 percentage point increase in the probability an individual stays home on a given day. We also find that friend-exposure effects become larger with closer friendships. In addition, while the effects of friend-exposure tend to be larger for some users—in particular those with a college-degree and those living in higher income areas—sizeable positive effects can be observed for all subgroups considered.

One possible concern with our analysis is that it might not capture a causal effect of friend-exposure to the disease, but instead result from a correlation of friend-exposure with unobserved demographic characteristics that explain differential changes in mobility. For example, since early outbreaks were primarily concentrated in urban centers such as New York City and Seattle, it might be that people with friends in these places would have disproportionately reduced their mobility even without the outbreaks where their friends live. We think that such alternative explanations are unlikely to explain our results: besides numerous demographic and location-specific controls, we control for various friend characteristics, such as the average population density and income of places where friends live. This allows us to rule out the possibility that we just pick up differences between people who have friends in cities or high income areas and people who have more friends in rural areas or lower income places.

In addition, we conduct a number of robustness checks to rule out many alternative stories. First, we address concerns that friend-exposure to COVID-19 cases might be correlated with individuals' ability to work from home. To do this, we estimate the effects of friend-exposure to COVID-19 on movement patterns separately for weekends and weekdays. Our results are virtually identical, suggesting differences in the ability to work from home that are correlated with friend-exposure to COVID-19 cases are not driving the observed effects. Second, we limit our sample to users with college information on Facebook and estimate the effect of friend-exposure to COVID-19 with an added control for users' exact college. Our results remain essentially unchanged. As this specification compares individuals within the same zip code, of the same gender and age group, who attended the same college, it allows us to hold fixed many factors determining individuals' desire or ability to socially distance.

In addition to studying the effects of friend-exposure early on in the pandemic, we also estimate the effects of *changes* in friend-exposure to COVID-19 cases over a given month on *changes* in mobility during that month, instead of using friend-exposure at a *single* point in time as we do in our static baseline analysis. For example, we estimate the effect of the change in friend-exposure from late April to late May on the change in social distancing behavior over the same time. We find strong positive effects over all periods, even after controlling for changes in prior months. Because the geography and characteristics of users with the highest change in our exposure measure varies as the pandemic evolves (for example, in February, cases grew the fastest in New York City and Seattle, while in June, they grew the fastest in Oklahoma, Texas, and Arizona), this *dynamic* relationship between changes in friend-exposure to COVID-19 cases and changes in social distancing behavior allows us to rule out many unobserved effects driving our results. In particular, it is difficult to argue that, in every month, it would be exactly those individuals with friends in parts of the country with the fastest acceleration of the virus that would have independently reduced their mobility for reasons other than their friend-exposure.

Next, we explore the mechanisms through which friend-exposure to COVID-19 affects individual behavior using data from public user posts and group memberships, as well as disaggregated mobility and spending data. We first define a series of regular expressions to identify public posts related to the pandemic and whether the post supports social distancing guidelines and other restrictions imposed on public life. Similarly, we identify public groups advocating for a reopening of the economy ('Reopen-Groups'). We find that friend-exposure to COVID-19 cases increases individuals' propensity to post about COVID-19 and the probability that a post with a social distancing opinion will voice support for restrictions on public life. In addition, we find that greater friend-exposure to COVID-19 cases lowers the likelihood of an individual joining a Reopen-Group. Overall, these observations are consistent with our hypothesis that friend-exposure to COVID-19 cases raises awareness about the risks of the disease and induces individuals to participate in mitigating public health behavior. It also suggests that the effects come through an effect of individuals' *desire* for social distancing (and support of policies that enforce such distancing), rather than through a differential *ability* to engage in such behavior.

In the final part of the analysis, we combine zip-code level mobility data from Safegraph with the Social Connectedness Index (SCI) from Bailey et al. (2018b) to show that our qualitative results—that friend-exposure to COVID-19 affects social distancing—replicates with alternative, public data sources. In addition, we use these public data to better understand the mechanisms through which friend-exposure to COVID-19 cases affects social distancing. We first find that, at the zip code level, friend-exposure to COVID-19 results in substantial decreases in visits to restaurants, bars, and places related to the arts, entertainment, and recreation. In contrast, we observe no effects for visits to essential places such as food and beverage stores, or centers of health care and social assistance. This provides further evidence that the reductions in mobility driven by friend-exposure are a measure of health-driven social distancing behavior, rather than capturing the ability to work from home. We also use transaction data from Facteus to explore the effect of friend-exposure to COVID-19 on economic outcomes. We find that individuals living in zip-codes with high friend-exposure to COVID-19 substantially decrease their transactions at nonessential merchants (i.e. Starbucks) relative to others while not differentially changing their spending behavior on Amazon. This is again consistent with friend-exposure to COVID-19 increasing one's desire to avoid nonessential physical interactions.

This paper contributes to a recent literature that studies the relationship between social networks and social distancing behavior using aggregated data. For example, Charoenwong et al. (2020) use the SCI data introduced by Bailey et al. (2018b) to show that individuals living in US counties with more connections to China and Italy—two early hotspots of the COVID-19 pandemic—are more likely to reduce their mobility. Makridis and Wang (2020) also use the SCI to show that consumption decreases more in counties with more friend-exposure to COVID-19. We contribute to this emerging literature in several ways. Most importantly, because we measure mobility behavior at the individual level, we are able to further explore the relationship between mobility and certain demographics. In turn, this allows us to rule out various alternative explanations that might confound an observed relationship between social networks and distancing behavior at the aggregate level. The availability of individual-level demographic data also allows us to explore the heterogeneity of the effect of friend-exposure to COVID-19 on different individuals. In addition, we can exploit additional data on posts and group memberships that shine a light on the mechanisms driving these relationships.

More broadly, this paper contributes to a rapidly growing literature on the determinants of social distancing during the COVID-19 pandemic. Giuliano and Rasul (2020b) and Brodeur et al. (2020a) provide early overviews of this work. Prior work has investigated informational channels that drive distancing behavior: Simonov et al. (2020) and Bursztyn et al. (2020) find that exposure to news downplaying the seriousness of the pandemic decreased distancing behavior, while Tian et al. (2020) argue that international migration networks helped to convey information about the disease. The literature has also found that higher levels of civic capital (Giuliano and Rasul, 2020a; Barrios et al., 2020), trust in scientific knowledge (Brzezinski et al., 2020), trust in policy makers (Bargain and Aminjonov, 2020), and general trust (Brodeur et al., 2020b) are all associated with greater levels of social distancing. Related studies have investigated the role of political affiliation (Allcott et al., 2020b; Barrios and Hochberg, 2020). Not only are a state's own policies an important determinant of social distancing (Allcott et al., 2020a), but there are also important spillover effects across places (Holtz et al., 2020). Our results build on these prior works, using large-scale individual level data to provide new insights into the determinants of health behaviors during a pandemic.

The remainder of this paper is structured as follows. In Section 2 we describe the Facebook data and descriptive patterns of social distancing over time. Section 3 presents our primary analyses, exploring the effects of friend-exposure to COVID-19 on social distancing behavior. We discuss heterogeneity in the effects of friend-exposure in Section 4. Using data on public post, group membership behavior as well as public, disaggregated mobility and spending data, we provide evidence on the mechanisms of our findings in Section 5. We conclude in Section 6.

2 Data and Descriptive Statistics

We work with de-identified data from the global online social networking site Facebook. As of December 2019, Facebook had 248 million monthly active users and 190 million daily active users in the U.S. and Canada (Facebook, 2020). An independent 2016 study found that, among U.S. adults, usage rates are relatively constant across income groups, education levels, and race; usage rates were slightly declining in age (Duggan et al., 2016). Establishing a connection on Facebook requires the consent of both individuals, and a person can have at most 5,000 connections. As a result, Facebook connections are primarily between real-world acquaintances. Facebook networks therefore resemble real-world social networks more closely than networks on other online platforms where uni-directional links to non-acquaintances (e.g., celebrities) are common. Indeed, a number of prior studies show that Facebook networks predict many important real-world economic and social interactions including patterns of trade (Bailey et al., 2020a), patent citations (Bailey et al., 2018b), travel flows (Bailey et al., 2019b, 2020b), bank lending (Rehbein et al., 2020), social program participation (Wilson, 2019), investment decisions (Kuchler et al., 2020a), and disease transmission (Kuchler et al., 2020b). Information on individuals' Facebook friendship links can also help understand their product adoption decisions (Bailey et al., 2019c) and their housing choices (Bailey et al., 2018a, 2019a).

2.1 Sample Restrictions and Summary Statistics

Our analyses of mobility behavior are limited to a sub-population of Facebook users who have consented to sharing and storing their location; have active accounts; are 18 or older; live in the 50 U.S. States or the District of Columbia; and have between 100 and 1,500 U.S. friends. We restrict the analysis to ZIP Code Tabulation Areas (ZCTAs) with 50 or more users that meet all previous requirements. Overall, the sample of users that meet the above criteria includes 12.8 million individuals. The average ZCTA has 592 users, the median has 319, and the 10th percentile has 72 users. We do not require users to have location information in every week (for example, if their device was off) and thus observe information for about 7.2 million users per week.

Table 1 provides summary statistics describing the sample. Age is widely distributed, with the 10th percentile at 26 years and the 90th percentile at 63 years. 53% of the sample is female, and just over half has a college listed on Facebook. We also observe information about whether the user primarily accesses Facebook mobile from an iPhone and whether the user accesses Facebook from a tablet (such as an iPad). Around 25% of the sample primarily uses an iPhone and around half have a tablet.

[Table 1]

After mapping users to their presumed ZCTA of residence, we supplement our individual-level data with public data on median household income from the 2014-2018 American Community Survey (ACS). The median user in our sample lives in a ZCTA with a median household income of $54,000, not far from the true median household income of $53,958. The 10th and 90th percentiles are $36,160 and $88,096, respectively, close to the true population values of $34,658 and $89,355.

Measuring Mobility and Social Distancing. We measure mobility using user-level GPS data for individuals who have consented to sharing and storing their location.[2] Location data are aggregated using the Bing Maps Tile System, which defines a series of grids at different resolution levels over a rectangular projection of the world (Schwartz, 2018). We use level-16 Bing tiles, which are 600 meters x 600 meters at the equator. Based on these data, we construct two mobility indices: (i) whether a user remains in the same level-16 Bing tile throughout the day (which we will refer to as "staying at home") and (ii)

[2]These data are similar to those described in Maas et al. (2019) and used to create the Facebook Data for Good Mobility Dashboard, available at https://www.covid19mobility.org/dashboards/facebook-data-for-good/.

the total number of distinct level-16 Bing tiles visited on a given day.

[Figure 1]

Figure 1 shows daily values of our two mobility measures between early February and late May.[3] In Figure 1a, we see that in February and early March, between 15% and 20% of users stay at home on a given day, with recurring spikes on weekends.[4] Starting around March 16—the day on which a large number of schools and offices were closed in response to the pandemic—the probability of staying at home jumps to well above 20%. As restrictions tightened in the U.S., the measure continues to rise, exceeding 30% by March 23 and rarely falling under 30% throughout April. In May, as social distancing restrictions were eased across parts of the U.S., the series decreases steadily; however, the probability of staying at home remains elevated relative to the baseline period and never falls under 20%. Figure 1b looks at the average number of tiles visited and depicts a similar trend: at the beginning of the sample period, there are fairly consistent patterns in the average number of tiles visited; the measure drops substantially around March 16, reaches its lowest levels in late April, before increasing steadily throughout May. Overall, we see similar large and persistent changes in both mobility metrics over time. Thus, in the main body, we focus on the probability that a user stays at home as our primary metric. Our findings remain unchanged if we instead use the percentage change in tiles visited as our metric. The corresponding figures are presented in the Appendix.

2.2 Heterogeneity in Social Distancing Behavior by Demographics

In this section, we show how social distancing varies across demographic characteristics. Using Facebook data allows us to understand the relationship between *individual* factors—such as age, gender, and college degree attainment—and distancing behaviors. We also match users to their assumed place of residence and study variation in distancing by geographically aggregated measures of income and direct exposure to COVID-19 in their locations. In Table 2, we split our sample into demographic groups and present statistics on each group's baseline probability of remaining at home during a given day (during the period between February 25 - March 2) and the change in that probability between the baseline period and the period April 14 - April 20 as a measure of social distancing. (Appendix Table A1 shows corresponding statistics for the change in the average number of tiles visited.)

[Table 2]

Table 2 shows that, while older individuals already spent more time at home prior to the pandemic, they also changed their behavior more during the pandemic. The probability of staying at home for users older than 55 years increased from about 26 percent to about 41 percent; for users between 18 and 34 years old the probability of staying at home increased from about 14 percent to about 28 percent. This finding is consistent with the fact that COVID-19 poses a greater risk to older individuals, which may

[3] In all graphs in this section, we control for a change in the methodology of location data collection near the end of February. Specifically, we assume that the relationship between the levels of our metrics in early February and the levels in the week of February 24th matches the relationship over the same time periods in the SafeGraph data described in the Appendix. We do not make this adjustment in any other section, where we either use only data after the methodology change or estimate results using a difference-in-difference approach (where the methodology change had quasi-random effects across groups).

[4] These patterns are also reflected in Table 1: during the month of February, the probability of staying at home is around 18.5% averaged over all days. It is 19.5% on weekends and 17% on weekdays.

induce them to social distance more. Similarly, prior to the pandemic, female users spent about 20% of days at home, while males only spent about 16% of days at home. Despite these baseline differences, women increased their rate of staying home by 15.7 percentage points, versus 11.3 percentage points for men. This differential shift could be driven by gender-based differences in labor market participation or occupation, or by the increased childcare burden being borne by women during the pandemic, as noted in Boca et al. (2020) and Alekseev et al. (2020).

We also find that while users who list a college education spent less time at home during the baseline period (17.7% probability for college users vs. 19.1% for non-college users), they increased their probability of staying home by *more* than users without college education. This finding is consistent with the conclusions from Dingel and Neiman (2020), who note that jobs requiring high levels of educational attainment are less likely to be deemed "essential" and can more often be done from home. Individuals in areas with higher median incomes also spent less time at home before the pandemic, but social distanced more after the onset of the pandemic. The likelihood of spending a day at home for users in high-income ZCTAs increased by 16.7 percentage points compared to 12.8 percentage points and 11.5 percentage points for users in middle- and low-income areas, respectively.[5] Coven and Gupta (2020) find similar results in New York City, hypothesizing that wealthier individuals have more work flexibility and a greater ability to use expensive food delivery options to replace retail visits.

Finally, the table shows heterogeneity in social distancing behavior by local exposure to COVID-19 cases, as reported by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.[6] In the baseline period, areas that had high, medium, and low levels of cases show similar mobility patterns. During the pandemic, individuals in areas with worse COVID-19 outbreaks practiced more aggressive distancing: users in high-COVID counties increased their share of days at home by 16.8 percentage points, compared to a 10.9 percentage point increase for users in low-exposure areas. This is likely a result of both stricter local lockdown policies and differences in perceived threat. We present time series versions of these results in Appendix Figure A1; these figures highlight that the demographic differences in social distancing behavior discussed above arise in mid-March 2020, and persist through the end of May. In Appendix Figure A2 we plot the corresponding figures using our alternative metric of social distancing based on the average number of Bing tiles visited in a day. The conclusions using this alternative measure mirror those discussed above.

To explore whether these heterogeneities primarily reflect differences in the ability to work from home, columns 3-6 of Table 2 separately analyze mobility on weekends and weekdays (see also Appendix Figures A3, A4, A5, and A6.). Although essential work may be also required on weekends, the majority of workers do not work over the weekend. Hence, differences in the ability to work from home should have a much smaller effect on social distancing on those days. However, we find similar heterogeneities in mobility on weekdays, suggesting that differences in the ability to work from home are not the primary cause of the observed differences.

[5] We sort individuals into low, medium, and high tertiles (with the same number of users) based on their ZCTA median income. Low-income ZCTAs have a median income below $47,178, while high-income ZCTAs have a median income above $62,734.

[6] Counties in the bottom tertile had below 0.09 cases per 100k residents, while counties in the top tertile have above 0.47 cases per 100k residents.

3 Effects of Friend-Exposure to COVID-19 Cases on Social Distancing

In this section, we study the relationship between friend-exposure to COVID-19 cases and social distancing behavior. Specifically, we ask whether users whose friends live in areas with worse coronavirus outbreaks engage in more social distancing. We begin by describing our measure of friend-exposure, before outlining our empirical approach and our results.

3.1 Measuring Friend-Exposure to COVID-19 Cases

Our measure of friend-exposure to COVID-19 cases for a user at a given time is given by:

$$\begin{aligned} FriendExposure_{it} &= \sum_{j=1}^{J} \frac{NumFriends_{ijt}}{\sum_{j=1}^{J} NumFriends_{ijt}} \times COVID19Cases_{jt} \\ &= \sum_{j=1}^{J} FracFriends_{ijt} \times COVID19Cases_{jt} \end{aligned} \tag{1}$$

$NumFriends_{ijt}$ gives the number of friends of person i in county j at time t, and $COVID19Cases_{jt}$ gives the cumulative number of COVID-19 cases reported in county j before time t.[7] In our baseline specifications, we define $FriendExposure_i$ using case counts from March 15, capturing individuals' early perceptions of the pandemic. Table 1 shows that there is substantial variation in this measure, with a mean of 10.35 friend weighted cases and a standard deviation of 19.34. For the first weeks of the outbreak, the correlation of $FriendExposure_{it}$ across time is high, as a similar set of U.S. locations had the highest cumulative case counts. In Section 3.5, we analyze a longer time frame using *changes* in the log of friend-exposure over time, a measure that has a pronounced negative correlation across months, as different parts of the country were most affected by the pandemic at different points in time.

While we want to understand the relationship between friend-exposure to COVID-19 and individuals' mobility choices, it is important to emphasize that friend-exposure in March 2020 is not randomly assigned. Instead, given the geographic concentration of U.S. COVID-19 cases in mid-March, friend-exposure likely correlates with individual characteristics that might also affect behavior during a pandemic. To understand the relationship between friend-exposure to COVID-19 and individual and ZCTA-level characteristics, we regress several such factors on the log of $FriendExposure_{i,March15}$.

[Table 3]

Column 1 of Table 3 shows that, on March 15, users who have college experience, have an iPhone, or live in higher income ZCTAs are more likely to have greater friend-exposure to COVID-19 than others. An R-squared of over 0.38 indicates that these variables explain a substantial share of the variation in friend-exposure. A regression featuring ZCTA fixed effects alone has an R-squared of about 0.67 —

[7] In this section, we primarily use measures that do not normalize cases by the county populations. In the early stages of the pandemic, when cases counts were low, we believe that the raw number of cases was likely a more salient measure of COVID-19 exposure than a normalized measure. For example, the areas with highest case exposures on March 15th were King County and New York City, each widely covered as early pandemic hot spots. By contrast, the areas with highest population-normalized infection rates were Pitkin and Eagle counties in Colorado. The outbreaks in these small counties received relatively little attention. In column 3 of Appendix Table A3 we show that our primary results holds when normalizing by population. In Section 3.5 we use normalized measures of exposure when exploring later stages of the pandemic.

consistent with the fact that a substantial share of most individuals' friends live nearby (see Bailey et al., 2018b) and that the severity of local outbreaks plays a large role in determining distancing behavior. In Column 3, we add fixed effects for national percentiles of friend weighted median household income, population density, and share urban population. More precisely, for each user i we calculate:

$$FriendMetric_i = \sum_{j=1}^{J} FracFriends_{ij} \times Metric_j \quad (2)$$

Here, $Metric_j$ is, for county j, either median household income, population density, or the share of the population living in urban areas.[8] We rank all users based on the resulting three metrics and group them into percentiles. Including these controls raises the R-squared to 0.85, consistent with the fact that in mid-March, most COVID-19 cases in the United States were in urban areas. Column 4 shows that the magnitude of the coefficients on the other observable characteristics falls substantially when including fixed effects for ZCTA and other friend-weighted factors, though younger users, male users, users with a college degree, and users who own an iPhone remain more likely to have high friend-exposure to COVID-19. Because these analyses highlight that certain demographics are correlated with friend-exposure on March 15 in an important way, we will always include a rich set of control variables when assessing potential effects of friend-exposure to COVID-19 on social distancing. More importantly, in Section 3.5 we also use the change in exposure over time—a measure that does not have a consistent relationship with demographic observables—to document the robustness of our conclusions.

3.2 Social Distancing Behavior of Users with High and Low Friend-Exposure

Since much of the variation in friend-exposure to COVID-19 is determined by home ZCTA, we first distinguish between users with high and low levels of friend-exposure relative to others *within* the same ZCTA. Concretely, for every ZCTA k, we calculate the median friend-exposure to COVID-19 as of March 15. We then define $HighExp_i$ for user i as an indicator of whether the user is above or below their ZCTA median. This measure allows us to introduce our baseline results with simple graphical representations.

[Figure 2]

Figures 2a and 2b present time series plots for our two measures of mobility—the probability of staying home and the average number of tiles visited—split by $HighExp_i$. Figure 2a shows that conditional on ZCTA, users with high and low levels of friend-exposure behaved similarly before the pandemic. Through February and early March, the probability of staying at home for both groups was between 17% and 20%, with differences never reaching half of a percentage point. Consistent with this, the bottom rows of Table 2 show that from February 25 to March 2, the probability of staying at home was 18.5% for high friend-exposure users and 18.2% for low friend-exposure users. Starting in mid-March, however, high friend-exposure users became substantially more likely to stay home. By early April, individuals with high friend-exposure have a probability of staying at home of close to 35% compared to less than 32% for users with lower levels of friend-exposure. Through April and May the probability of staying home declines steadily for both groups, but the difference between them remains roughly

[8] The data on median household income and population density come from the 5-year ACS from 2014-2018 and the share of the population living in urban areas comes from the 2010 Census.

the same. Figure 2b tells a similar story: while no differences in levels and in trends can be observed between high and low friend-exposure group before the outbreak of the pandemic in the U.S., after the beginning of the lockdown in mid-March, users with high friend-exposure are substantially more likely to engage in social distancing. As before, these early differences persist through mid-May. That individuals' early perceptions of the crisis could shape their later beliefs is consistent with existing survey evidence showing that those in areas of early COVID-19 spread remained the most worried about the virus in later months (Blumenthal, 2020).

Difference-in-Differences Analysis. To further explore the effects of friend-exposure on social distancing behavior, we estimate the following difference-in-differences specification:

$$Y_{it} = \beta_0 + \beta_{1t} \times HighExp_i \times week_t + \beta_{2t} \times X_i \times week_t + \eta_i + \epsilon_{it} \tag{3}$$

The dependent variable, Y_{it}, is a measure of mobility (either the probability of staying home or the number of tiles visited) for individual i during week t. $HighExp_i$ is an indicator equal to one if user i has friend-exposure greater than their ZCTA median on March 15. The vector $week_t$ includes indicators for each week of the sample. We interact these weekly dummies with each of X_i, our controls for ZCTA, college attainment, ownership of iPhone and tablet, age group, and gender, as well as for percentiles of friend-weighted median household income, population density, and share urban. This allows us to control for time-varying effects of differences in demographics on social distancing behavior. We also include individual-level fixed effects, η_i. We cluster standard errors by ZCTA.

Figures 2c and 2d show estimates for the coefficients of interest, β_{1t}, by week, for both measures of mobility. In Figure 2c, the coefficient estimates prior to mid-March are close to zero, and nearly always statistically insignificant. Then, as lockdowns began, the coefficient estimates jump. During the week of March 16th, moving from below ZCTA-level median friend-exposure to above corresponds to an increase in the probability of staying at home of about 0.7 percentage points. The coefficient estimates continue to rise until late April, reaching levels above 1 percentage point. During the week of April 6, the coefficient estimate of close to 0.012 corresponds to a 3.8% increase in the probability of staying home relative to the average of 32% among users with low friend-exposure. Towards the end of the sample period, the coefficients begin to decline slightly, yet remain close to 0.01 and always highly significant.[9] The coefficient estimates for the number of tiles visited in Figure 2d are consistent with this: movement patterns look similar for high and low friend-exposure users through mid-March, then gradually decrease to levels around -0.2, where they remain through May. Together, Figures 2c and 2d suggest sizeable effects of friend-exposure to COVID-19 on social distancing behavior. Although the two groups' movement look nearly identical prior to the pandemic in the U.S., high friend-exposure

[9]While the qualitative patterns in Figures 2c and 2d are very consistent with Figures 2a and 2b, it is worth noting that the effect size is noticeably smaller. This decrease is due to the rich set of control variables in equation 3—individual demographics and other measures of friend-exposure—and highlights that differences in mobility, even within ZCTA, are related to factors other than friend-exposure. That we observe sizeable and highly significant coefficient estimates with the full set of controls suggests that friend-exposure to COVID-19 does indeed effect social distancing. However, we stress the importance of controlling for other factors (in particular those related to ability to work from home) when exploring this relationship. In Section 3.4 we will also conduct a set of analyses to further test our identification.

users are substantially less mobile after the outbreak begins. This effect is persistent through May.

3.3 Benchmarking the Effects of Friend-Exposure Effects on Social Distancing Behavior

To benchmark the importance of friend-exposure to COVID-19 in determining social distancing behavior relative to other factors, we return to our measure from Table 2: the change in the probability of staying home between the week of February 25-March 2 (prior to the U.S. pandemic) and the week of April 14-20 (during a period of widespread social distancing). We then conduct the following multivariate analysis:

$$\Delta Y_i = \alpha_0 + \alpha_1 \log(FriendExposure_i) + \alpha_2 Z_i + \alpha_3 C_i + \epsilon_i \tag{4}$$

Here, ΔY_i is individual i's change in the probability of staying home. $FriendExposure_i$ is defined as in equation 1 (as of March 15). Z_i is a vector consisting of age dummies, gender, educational attainment, ownership of iPhone and tablet, ZCTA-level income, and local exposure to COVID-19 (county-level COVID-19 cases per resident as of March 15). As before, for each of the last three covariates, we rank areas nationally and assign them to tertiles. The coefficients of interest are the vectors α_1 and α_2. We also include other controls, C_i, which vary across specifications.

[Table 4]

Table 4 presents the results; Appendix Table A2 presents corresponding results using the percentage change in tiles visited as the dependent variable. Consistent with the univariate patterns documented in Table 2, column 1 of Table 4 shows that older users engage in more social distancing: in the presence of other controls, users aged 55 and above increase the probability of staying at home by 1.4 percentage points more than users aged 18-34. Female users and those who attended college also increase their probabilities of staying at home by 4.4 and 2.9 percentage points more than male users and those without college education, respectively. Column 2 shows that the heterogeneities with respect to individual level characteristics are robust to including ZCTA fixed effects.

In columns 3-5 of Table 4, we include coefficient estimates for friend-exposure to COVID-19 cases, as well as fixed effects for friend-weighted characteristics as described in the discussion of equation 2.[10] Column 3 includes ZCTA fixed effects but omits all other individual-level characteristics of columns 1 and 2. Given a standard deviation in $\log(FriendExposure_i)$ of 1.35, the coefficient estimate on friend-exposure of 0.92 indicates that a one standard deviation increase in friend-exposure is associated with an increase in the probability of staying at home of about 1.2 percentage points, an 8.8% increase relative to the sample mean of 13.7%. Adding additional individual-level characteristics to the regression in column 4 slightly decreases the coefficient estimate for α_1 only slightly to 0.85. Comparing the coefficient estimates for friend-exposure to COVID-19 to those for other individual-level characteristics highlights that friend-exposure is an important determinant of social distancing. An increase in friend-exposure by one standard deviation corresponds to an increase in social distancing that is more than two thirds as large as being age 55 or older, and hence belonging to a group that is considered most vulnerable to the health risks of COVID-19. Similarly, the effect of a one standard deviation increase friend-exposure is

[10] To further assess the robustness of our results to including additional friend characteristics, in Appendix Table A3, we also control for the fraction of friends living in China, South Korea, Italy, and Spain, all of which were early hotspots of the COVID-19 pandemic. Reassuringly, coefficient estimates are virtually unchanged when adding these control variables.

roughly half of the effect of having college experience, often cited as an important driver of individuals' ability to engage in social distancing behavior (see Dingel and Neiman, 2020).

In column 5 of Table 4, we include the full interaction of individual-level controls with ZCTA fixed effects. This has no additional impact on the estimated coefficient estimate for α_1.[11] In Figure 3, we present a binned scatter plot corresponding to this specification (Appendix Figure A7 presents the corresponding binned scatter plot for the percentage change in average tiles visited). This figure confirms the linear relationship between the change in mobility and the log of friend-exposure. The relative stability of the α_1 coefficient to the addition of ever-tighter individual-level controls suggest a limited confounding effect of other characteristics that might be correlated with friend-exposure, at least once we control for ZCTA fixed effects.

[Figure 3]

3.4 Ruling Out Alternative Interpretations

The prior analyses in this section suggest the potential presence of important effects of friend-exposure to COVID-19 on social distancing behavior. We find that users with higher levels of friend-exposure are substantially more likely to stay at home than others in subsequent weeks, even after controlling flexibly for a rich set of demographics and friend characteristics. Nevertheless, it is still possible that other unobserved factors drive some of the relationship between social distancing and friend-exposure to COVID-19 cases. In particular, differences in the ability to work from home or differences in preferences are plausible confounding factors. We next present robustness checks to alleviate such concerns.

Friend-Exposure Effects on Weekends and Weekdays. Our first robustness check addresses concerns that our estimates might be picking up differential ability to work from home, which could be correlated with friend-exposure to COVID-19 cases. To rule out such effects, we separately analyze the effect of friend-exposure on mobility on weekdays and weekends. If our previous results were predominantly driven by differential abilities to work from home, we would likely observe different effect sizes on weekdays and weekends. In particular, as most individuals do not work on weekends, we would observe smaller differences in distancing on weekends than on weekdays. However, if differences in social distancing behavior are similar for weekdays and weekends, it will help rule out ability to work from home as a confounding factor. In Figure 4, we re-estimate equation 3 with the probability of staying at home as the outcome variable, separately for weekdays and weekends. Appendix Figure A8 shows the corresponding coefficient estimates with the percentage change in the number of tiles visited as the outcome variable. The differential effect of friend-exposure to COVID-19 on changes in mobility are qualitatively and quantitatively similar for weekends and weekdays: in both cases, effect sizes are between 0.01 and 0.012 at the peak, nearly identical to the baseline estimation presented in Figure 2c. In addition, in columns 6-7 of Table 4, we follow the estimation strategy of Section 3.3, but subset to weekend and weekday behavior, respectively. Again, the coefficients are very similar, suggesting our earlier results are unlikely to be driven by differences in ability to work from home.

[11] Our sample size is about 5% smaller in this regression, due to combinations of ZCTA- and individual-level characteristics for which we have only a single observation.

[Figure 4]

Within-College Comparisons. For our second robustness check, we subset our analysis to users who list a college on their Facebook profile and include college fixed effects interacted with week indicators. Since educational attainment plays a large role in determining employment, users who went to the exact same college (and are *also* of similar age, the same gender, live in the same ZCTA, have same ownership of iPhone and tablet, and have generally quite similar friend characteristics) should have comparable levels of ability to work from home. Figure 4c shows the corresponding coefficient estimates.[12] The results are qualitatively and quantitatively very similar to the baseline findings presented in Figure 2c, despite the substantially smaller sample and narrower comparison of users. Consistent with this, column 8 of Table 4 highlights that the coefficient estimates discussed in depth in Section 3.3 remain almost unchanged—and if anything increase—when focusing on this more restricted sample and adding college fixed effects. Together, these finding provide further evidence against unobservables driving our previous results.

3.5 Dynamics of Friend-Exposure to COVID-19 and Social Distancing Behavior Over Time

Finally, rather than focusing on the effects of friend-exposure to COVID-19 at one specific point in time, we study the effects of *changes* in friend-exposure as the pandemic evolves on *changes* in social distancing. We do so to alleviate concerns that the correlations presented in Table 3 between our March 15th exposure measure and certain individual characteristics—or, more worrisome, correlations with *un*observables—are driving our earlier results. As the pandemic progressed, the geography of outbreaks, and therefore the set of individuals exposed to COVID-19 through their friends, changed substantially. It is thus difficult to argue that, in every month, it would be exactly those individuals with friends in parts of the country with the fastest acceleration of the virus that would have independently reduced their mobility for reasons other than their friend-exposure. To operationalize this research design, for each month, we define:

$$ChangeFriendExposure_{it} = Log(1 + FriendExposure100k_{it}) - Log(1 + FriendExposure100k_{it-1}) \quad (5)$$

with

$$FriendExposure100k_{it} = \sum_{j=1}^{J} FracFriends_{ijt} \times \frac{COVID19Cases_{jt}}{Residents100k_j} \quad (6)$$

In words, for each month we take the log of friend-exposure to COVID-19 (per 100,000 inhabitants) at the end of that month and subtract the same measure from the end of the preceding month.[13]

[12] In Figure 4c the dependent variable is the probability of staying at home. Appendix Figure A8c presents the corresponding estimates for the number of tiles visited as LHS variable.

[13] We normalize cases by population to alleviate concerns that potential effects might be driven by factors that are correlated with population size. See footnote 7 for an explanation of why we do not normalize in our previous measure of exposure. Our earliest measure of exposure uses cases as of March 27th. Our decision to use the difference of logs will weight new cases in new outbreaks more than new cases in old outbreaks—for example, because $log(100) - log(1) > log(199) - log(100)$—which we believe to be consistent with the salient, behavior-altering change we hope to measure.

[Figure 5]

To support the notion that the set of individuals with high levels of $ChangeFriendExposure_{it}$ changes substantially over time, in Figure 5 we present maps showing the ZCTA-level average of this measure by month. Bright green colors indicate large increases in friend-exposure to COVID-19 for a given month, while darker blues indicate smaller changes. The Figure shows that the geography of $ChangeFriendExposure_{it}$ shifts substantially over time. In March, the measure is particularly high in New York, Seattle, Denver and Louisiana. By April, however, the highest average levels are throughout the Midwest; in May, hotspots appear in Minnesota, Iowa, and North Carolina, while in June they are in Texas, Oklahoma, and Arizona. Finally, in July it is in particular places in Texas bordering Mexico that exhibit high average increases in friend-exposure to COVID-19.

[Table 5]

We provide further evidence of the variation in the characteristics of users with high levels of new friend-exposure to COVID-19 over time in Table 5, which regresses $ChangeFriendExposure_{it}$ for each month from March to July on various individual- and region-characteristics. We find that the relationship between friend-exposure to COVID-19 and other factors varies across months. For example, users with a college degree were more exposed to COVID-19 through their friends at the beginning of the pandemic (when the pandemic was primarily an urban phenomenon); in later months, as the pandemic spread across the United States, the relationship reverses. Similar reversals occur for each of the other control variables, most interestingly for friends' incomes and population density.

Together, Figure 5 and Table 5 highlight that the set of individuals with the largest increases in friend-exposure to COVID-19 varies substantially over time. While our previous analyses suggest that very early friend-exposure plays a substantial role in shaping behavior throughout the first months of the pandemic, we next test if the *changes* in friend-exposure lead to *changes* in social distancing behavior. Given that those who experience the largest increase in this exposure look very different from month to month, this result would be unlikely to be driven by omitted variables. We estimate the following equation:

$$\Delta Y_{it} = \sigma_0 + \sigma_1 ChangeFriendExposure_{it} + \sigma_{2t} X_{it} + \sigma_{3t} N_{it} + \epsilon_i \tag{7}$$

ΔY_{it} is the change in the probability of staying at home from the last week of the prior month to the last week of the current month t (e.g., from end of February to end of March). X_{it} is a flexible vector of controls. Figure 6 presents the resulting estimate of σ_1. In Panel (a), we control for individual-level variables as in Equation 4 all interacted with each other, ZCTA, and month. In addition, we control for month interacted with fixed effects of percentiles of friend weighted urbanity, population density and median household income, N_{it}. In Panel (b), we additionally control for user fixed effects. Consistent with the hypothesis that changes in exposure result in changes in social distancing behavior, we observe a strong and positive relationship between friend-exposure to COVID-19 and the probability of staying at home. Depending on the specification, the magnitude of the coefficient (based on a linear regression) is between 0.21 and 0.26, and is in both cases highly significant.

[Figure 6]

We next test the relationship between changes in social distancing behavior and friend-exposure for each month separately. Concretely, for each month, we regress the change in the probability of staying at home on changes in friend-exposure to COVID-19 over that month, as well as over all preceding months:

$$\Delta Y_{it} = \sigma_0 + \sum_{j=1}^{t} \sigma_{1j} ChangeFriendExposure_{ij} + \sigma_2 X_i + \sigma_3 N_i + \epsilon_{it} \tag{8}$$

We conjecture that if friend-exposure to COVID-19 is a key explanatory factor in determining social distancing, changes in social distancing in a given month should respond most to the change in friend-exposure to COVID-19 during that same month. In all our regressions, we control for interactions of individiual-level characteristics and ZCTA fixed effects (X_i) as well as other friend weighted characteristics (N_i).

[Table 6]

Table 6 presents the results of this analysis, which support our interpretation. In March, the increase in friend-exposure to COVID-19 has a highly significant positive effect on the change in the probability to stay at home: a doubling of the increase in friend-exposure leads to a roughly 0.14 percentage point increase in the probability of staying at home.[14] More importantly, in subsequent months, changes in social distancing behavior is driven by changes in the friend-exposure of the corresponding months. Across all months, we find that the most recent changes in the rate of friend-exposure are the most important, though our results in April are not statistically significant. These findings support our hypothesis that friend-exposure to COVID-19 has a sizeable effect on social distancing behavior. Because characteristics of users with high friend-exposure to COVID-19 change substantially over time, the *dynamic* relationship between changes in friend-exposure and changes in social distancing behavior allows us to rule out many unobserved effects that could potentially influence our results. In addition, in Appendix Table A4, we regress changes in mobility only on changes in friend-exposure for the same month (without also including changes in prior months).[15] The results are again consistent with a causal interpretation of changes in friends' exposure to COVID-19 on own social distancing behavior.

4 Heterogeneity of Friend-Exposure Effects

We next explore how the effect of friend-exposure on social distancing behaviors differs with an individual's own characteristics as well as characteristics defining the friendships underlying our measure of friend-exposure.

Heterogeneity of Friend-Exposure Effects — Own Characteristics. To explore heterogeneity in the effect of friend-exposure by an individual's own characteristics, we amend equation 4 to estimate the effect of friend-exposure (on March 15) separately for several sets of mutually exclusive characteristics.

[14]Noticeably, this effect is somewhat smaller than our baseline presented in Table 4 (and for normalized cases in Table A3). Differences are driven by two factors: 1) here we measure exposure on March 27, as opposed to March 15 and 2) here we measure change in mobility as of March 24-30, as opposed to April 14-20. These two reasons also result in noticeable differences in mean and variance of the exposure metrics: in the context of Table 4, log of friend exposure had a mean of 1.45 with a standard deviation of 1.35, compared with 2.80 and 0.85, respectively for the definition used in Table 6.

[15]Formally, we estimate $\Delta Y_{it} = \sigma_0 + \sigma_{1t} ChangeFriendExposure_{it} + \sigma_2 X_i + \sigma_3 N_i + \epsilon_{it}$ and report the σ_{1t} coefficient estimates.

We consider several dimensions, looking separately at age, gender, educational attainment, ZCTA-level income, and county-level COVID-19 exposure. In all of these specifications, we use the set of controls included in column 5 in Table 4.

[Table 7]

Table 7 explores the characteristics that influence the response of "staying at home" to friend-exposure to COVID-19 cases; Appendix Table A5 shows the corresponding estimates for average daily tile movement. The effect of friend-exposure is much stronger for younger users: the effect for those aged 55+ is only about one third the size of the effect among those aged 18-34. In addition, the coefficient estimates for users with a college degree are about three times as large as corresponding estimates for non-college educated users. Coefficient estimates for females are somewhat larger than for males, though the differences are economically small. At the ZCTA-level, the effect of friend-exposure to COVID-19 on own mobility monotonically increases with the average income of an area: while a doubling of friend-exposure leads to an increase in the probability of staying at home of around 1.1 percentage points for users living in high-income ZCTAs, that effect size drops to around a quarter for users living in the lowest-income areas.[16] Interestingly, regardless of these heterogeneities, for all the various groups we consider, we find that friend-exposure leads to noticeable increases in social distancing behavior.

Heterogeneity of Friend-Exposure Effects — Friendship Characteristics. We also study heterogeneity in the effects of friend-exposure to COVID-19 by characteristics of the friendship. Specifically, we investigate the extent to which the effects of friend-exposure vary with strength of the friendship.

We consider a ranking of friendship 'closeness' based on the number of interactions between users on Facebook. Our specification amends equation 4 by replacing replace $FriendExposure_i$ by the four variables $FriendExposure_{1-25,i}$, $FriendExposure_{25-50,i}$, $FriendExposure_{51-75,i}$, and $FriendExposure_{76-100,i}$. For example, $FriendExposure_{1-25,i}$ is defined as friend-exposure to COVID-19 cases among the 25 closest friends of user i. It is important that each group contains the same number of friends, since the same average friend-exposure among a larger group of friends may induce larger effects. We then run specifications that are otherwise identical to the ones underlying column 5 in Table 4. The regression results are presented in the rightmost column of Table 7.

We find that the effects of friend-exposure tend to be strongest for the closest friends, with effect size falling off among more marginal friends. A doubling of friend-exposure among a person's 25 closest friends leads to a 0.14 percentage point increase in the likelihood that the person stays home. A similar increase in friend-exposure among the person's *next* 25 closest friends (those ranked 26-50) leads to a substantially smaller effect, of 0.08 percentage points. The effect size is smaller for the two more distant friend groups, so that the friend-exposure effects associated with the 25 closest friends are about

[16] In Appendix Figure A9 we show the same results in the event-study framework described in Section 3.2, where we estimate separate coefficients for the effects of friend-exposure by week for the various demographic groups. The corresponding results using the average daily tiles visited as an outcome are shown in Appendix Figure A10. The results are very consistent with the ones shown in Table 7.

twice as large as those associated with friends ranked 76-100.[17] The decrease in the effect size of friend-exposure as we move toward more 'distant' friends supports our hypothesis that the observed effects on health behavior are indeed driven by friend-exposure to COVID-19 cases rather than omitted variables and selection effects. Users are more connected to their closer friends and, therefore, more likely to obtain information from them, including about COVID-19. This informational channel is one possible mechanism for our previous results which we discuss in greater detail below.

5 Mechanisms

In our previous analyses, we have identified an effect of friend-exposure to COVID-19 on social distancing behavior, finding that greater friend-exposure to COVID-19 is associated with a substantial increase in a person's likelihood of staying at home. In this section, we explore the mechanisms behind these findings. Several mechanisms seem plausible. Most eminently, users are likely to learn from their friends – that is, they receive information about the threat of COVID-19 and hence adjust their behavior accordingly. Alternatively, users might simply respond to restrictions they are facing: the reduction in mobility as a result of greater friend-expoure to COVID-19 could reflect the fact that users are unlikely to visit friends and relatives and are thus more likely to remain at home. To differentiate between these and other potential mechanisms, we leverage data on public posts and public group membership on Facebook. In addition, we combine several sources of public data to highlight that the observed reduction in mobility as a result of greater friend-exposure to COVID-19 is consistent with individuals *choosing* to reduce physical interactions to minimize the transmission of the virus.

5.1 Posting Behavior

We begin by considering users' public Facebook posts, which can be viewed by any other user on the platform. We use these public posts to construct three different measures. First, we use regular expression searches to measure the percentage of a user's public posts that mention the coronavirus; this measure captures the user's level of general engagement in discussions about the coronavirus. Second, we identify common phrases used to support or oppose social distancing measures, to quantify a user's level of opposition to social distancing measures. Specifically, we create a metric that is defined as the number of posts opposed to social distancing as a fraction of all 'signed' posts (i.e., all posts identified as either supporting or opposing these measures). Third, we measure the sentiment of public posts using a text analysis algorithm, constructing monthly averages of the sentiment of each user's posts. Appendix C provides details on these measurements.

We then amend equation 4 slightly to estimate the effect of friend-exposure to COVID-19, as well as other individual- and ZCTA-level characteristics, on our three public post outcomes:

$$Y_i = \delta_0 + \delta_1 \log(FriendExposure_i) + \delta_2 Z_i + \delta_3 C_i + \epsilon_i \tag{9}$$

Y_i here corresponds to one of the three outcomes described above. $FriendExposure_i$, Z_i and C_i are defined

[17] The focus in this section is on the relative magnitudes of the effects of different equally-sized friends group. The absolute magnitudes cannot be easily compared to those from Table 4, since that looks at the effect of an increase in the average exposure among a much larger group of friends, which is likely to have a substantially larger effect.

as in equation 4. Table 8 presents the coefficients of interest, δ_1 and δ_2. In columns 1 and 2, Y_i is the share of public posts between February and April 2020 that are about the coronavirus. In columns 3 and 4, Y_i is the percentage of signed posts between February and April 2020 that are opposed to social distancing measures. In columns 5 and 6, Y_i is the change in sentiment of a user's posts between February and April. We impose the sample restrictions described in Sections 2 and 2.1 and further require that users have posted publicly at least once in February, March, or April of 2020. Unlike in our examination of mobility, we do not limit to users with location sharing and storage permissions, which increases our sample size substantially.[18] Summary statistics for this sample are shown in Appendix Table A6.

[Table 8]

Column 1 of Table 8 shows that, conditional on posting, users with college experience are close to 0.6 percentage points more likely to post about COVID-19 than users without college experience. This effect is large—a 33% increase from the baseline average of 1.8% of posts related to coronavirus. Individuals in areas with higher median incomes are also more likely to post about the coronavirus.

Friend-exposure to COVID-19 cases also has substantial effects on posting behavior. In column 1, we see that a doubling in friend-exposure corresponds to a increase in the share of posts about the coronavirus of about 0.22 percentage points, a 12% increase relative to the average. This effect remains sizeable when we include fixed effects for ZCTA interacted with individual characteristics in column 2. Figure 7a shows a binned scatter plot that corresponds to our analysis in column 2. The relationship between the percentage of posts about COVID-19 and friend-exposure is strong, with a functional form that is close to linear.

[Figure 7]

This first analysis suggest that users with higher levels of friend-exposure to COVID-19 are generally more likely to talk about the coronavirus; however, it does not allow us to capture the nature of individuals' posts. Specifically, our measure includes both (a) posts supportive of the notion that the virus poses a great threat to public health and endorsing measures to contain the risk, and (b) posts downplaying the threat of the virus or calling for an end to various restrictions. We therefore repeat our analysis in columns 3 and 4 of Table 8, focusing on the percentage of 'signed' posts which either oppose social distancing guidelines and shutdowns, or call for an earlier re-opening of the economy. For this analysis, we concentrate on the subset of users who share a 'signed' post in February, March or April.

The coefficient estimates in Table 8 show that friend-exposure to COVID-19 decreases the likelihood that users oppose social distancing measures in their posts. The effect is large even in our most conservative specification: a doubling in friend-exposure corresponds to a 1.3 percentage point reduction in the share of signed posts opposing distancing. This implies a 4% reduction given a baseline average of 36%. In Figure 7b we present a binned scatter plot corresponding to the specification of column 4.

To further analyze the effects of friend-exposure on user behavior, we use the VADER algorithm described in Hutto and Gilbert (2014) to estimate the average sentiment of individuals' posts over time. We replace Y_i in equation 9 with the change in average post sentiment between February 3-23 and April

[18] Importantly, we still observe an assumed ZCTA of residence (based on IP address, profile information, and other factors) allowing us to include ZCTA-level controls in our regressions.

6-26, and present the results in Columns 5 and 6 of Table 8. Users with higher levels of friend-exposure to COVID-19 cases have significantly larger decreases in post sentiment, even in our strictest specification.

5.2 Group Membership

We next explore the effect of friend-exposure to COVID-19 cases on users' decisions to join various Facebook groups. Facebook users can create and join various forms of groups to chat, meet and otherwise engage with others. For our analysis, we focus on membership in public groups, which any Facebook user can access without additional restrictions. Since no restrictions on posting behavior and/or location settings are necessary for this part of the analysis, we simply focus on active users who meet the non-mobility data requirements described in Section 2. We present summary statistics for this group of users in Appendix Table A7.

To further explore the determinants of health behavior, we focus on groups created between March 1 and June 28, 2020 with names that suggest support for an early reopening (or "liberation") of the economy. Appendix C provides details on how we identify these groups. We then estimate:

$$ReopenGroup_i = \gamma_0 + \gamma_1 \log(FriendExposure_i) + \gamma_2 Z_i + \gamma_3 C_i + \epsilon_i \qquad (10)$$

where $ReopenGroup_i$ is an indicator equal to one if on June 28 user i is a member of at least one group advocating for the lifting of COVID-19 related restrictions. $FriendExposure_i$, Z_i and C_i are defined as in equations 4 and 9 with the exception that, in addition to the control variables used above, we include fixed effects for national percentiles of group membership as of February 2020. That is, based on data from February 2020, we construct national percentiles of the number of groups a user is a member of. We then assign each user to a percentile based on the number of groups they are a member of as of February 2020. We do so in order to correct for potential differences in Facebook usage behavior. We present coefficient estimates for the coefficients of interest (i.e., γ_1 and γ_2) in columns 7 and 8 of Table 8.

Given the small number of sample restrictions, the number of observations included in this part of the analysis greatly increases. About 1.2% of all users are a member of at least one Reopen Group. Column 7 of Table 8 shows that older and male users are substantially more likely to be members of Reopen Groups than younger and female users, respectively. Users without college experience and users living in higher income areas are also more likely to be members of a group supporting reopening. These results are generally consistent with those observed for mobility and posting behavior: younger users, female users, and users with college experience are more likely to engage in social distancing and appear more supportive of restrictions on public life. Returning to our main variable of interest, in the presence of ZCTA-level fixed effects interacted with individual level covariates, we find a significant negative effect of friend-exposure to COVID-19 on membership in a group supporting reopening. In column 8, a doubling in friend-exposure to COVID-19 decreases the probability of being a member of a group that backs reopening by about 0.09 percentage points, or 7.5%.

Taken together, our results suggest that higher friend-exposure to COVID-19 cases increases the likelihood of posting about COVID-19 and increases the share of users who support social distancing guidelines. In addition, higher friend-exposure to COVID-19 appears to lower the likelihood that a user joins groups opposing public health guidelines. Together, these results add an important insight into the mechanisms driving our findings in Section 3: individuals more exposed to COVID-19 are more concerned about COVID-19 and support public health measures to a greater extent. This is consistent with the explanation that greater friend-exposure affects social distancing by shaping individuals' information and attitudes towards the disease. Moreover, the findings speak against alternative explanations based on the notion that individuals with higher friend-exposure might reduce their mobility simply as a response to restrictions they are facing without updating their beliefs and perceptions about the threat associated with COVID-19.[19]

5.3 Disaggregation of Friend-Exposure Effects

To extend our findings, we next use public data on movement and transactions from Safegraph that has been used extensively in contemporaneous research on COVID-19. We combine these data with Social Connectedness Index (SCI) data from Facebook (Bailey et al., 2018b). Our analysis serve two purposes. First, in Appendix Section A, we replicate our findings on the effects of friend-exposure on social distancing at the ZCTA-level. While this analysis does not allow us to effectively control for many individual-level characteristics that are correlated with changes in social distancing behavior and exposure to COVID-19, it has some advantages. Most importantly, the Safegraph mobility data are based on a different and larger set of individuals, thus mitigating concerns that the results discussed in Section 3 are merely an artifact of the somewhat selected sample of Facebook users who have consented to sharing and storing their location information with the platform. We find that our qualitative observation that friend-exposure to COVID-19 increases social distancing replicates fully at the ZCTA-level using alternative measures of mobility. Second, in the following sections, we expand our analysis at the ZCTA-level to explore a number of other observable outcomes that speak to the mechanism behind the observed reduction in mobility. This analysis also allows us to provide some evidence on the economic effects of the mobility reduction. More concretely, we disaggregate our results by type of establishment (or point-of-interest) and merchant type to understand whether individuals with high friend-exposure to COVID-19 not only reduce their mobility in general, but in particular stop visiting "nonessential" establishments where close physical interaction is common.

5.3.1 Mobility Effects By Type of Establishment

Safegraph aggregates cellphone GPS data to measure the number and duration of visits to points of establishments on a daily basis. We use these data through July 28, 2020 to construct a measure of the total number of POI visits by ZCTA, both for all POIs and for POIs of selected industries.[20] With the objective of distinguishing between 'essential' and 'nonessential' places, we focus on the following categories: (i) Arts, Entertainment, and Recreation (NAICS code 71), (ii) Food Services and Drinking Places (NAICS code 722), (iii) Retail Trade Excl. Food and Beverage Stores (NAICS codes 44 and 45, excluding 445), (iv) Food and Beverage Stores (NAICS code 445), (v) Parks (NAICS code 712190); and (vi) Health Care and Social Assistance (NAICS code: 62). We think of (i)-(iii) as less essential places that can be avoided in order to reduce physical interaction. In contrast, (iv)-(vi) are either more essential or entail very limited physical interaction. We aggregate visits to the weekly level.

[19] In addition to the results presented in this Section, in Appendix Tables A8, A9, A10, we study heterogeneities in the observed effects of friend-exposure to COVID-19, finding results largely consistent with those presented in Section 4.

[20] For the sample period, there are on average 27.5 million POI visits each day, distributed over roughly 5.4 million POIs.

To study the effect of friend-exposure to COVID-19 on POI visits, we construct a measure of friend-exposure that is very similar to the one used in Section 3, but differs in that it is a ZCTA-level average rather than being observable at the individual level. In particular, $HighExp_i$ is an indicator equal to one if ZCTA i has friend-exposure to COVID-19 higher than the median for the county it is located in, based on the number of COVID-19 cases as of March 17. We discuss the construction of this ZCTA-level friend-exposure metric in detail in Appendix Section A. Following the difference-in-differences analysis in Section 3, we then estimate:

$$Y_{it} = \beta_0 + \beta_{1t} \times HighExp_i \times week_t + \beta_{2t} \times X_{it} + \epsilon_{it} \tag{11}$$

where Y_{it} corresponds to the log of one plus the number of POI visits (by type of establishment) in a given ZCTA i per week t. We control for county-time fixed effects together with ZCTA fixed effects as well as various ZCTA-level covariates interacted with time fixed effects. These covariates are median household income of the area, as well as the fraction of individuals in in each of the following demographic groups: male, Asian, Black, White, service employee, manager, art or science employee, high-speed internet user, high-school educated, some college completion, college educated. We also control for the fraction of individuals in different age buckets.[21] All these measures are obtained from the most recent 5-year ACS (2014-2018). Following Section 3, we also control for national percentiles of friend weighted median household income, population density and urbanity. We cluster standard errors at the ZCTA-level.

[Figure 8]

Figure 8 shows coefficient estimates for β_{1t}, with each panel corresponding to a different type of POI. For reference, we include results for all POIs aggregated in the gray series. The patterns are consistent with the hypothesis that people in places with high friend-exposure to COVID-19 disproportionately reduce their mobility to avoid unnecessary physical interactions. While differential responses in POI visits are negative for nonessential POIs in Panels (a)-(c), they are close to zero and insignificant for essential POIs in Panels (d)-(f). More concretely, the coefficient estimates for arts, recreation, and entertainment locations (Figure 8a) show that the difference in the change of visits between high and low exposure places can be as large as 0.05 log points (in absolute magnitude). Similar effects can be observed for retail destinations (Figure 8b), and restaurants and bars (Figure 8c). Although coefficient estimates return to zero well before the end of the sample period, they are negative and highly significant for the period from mid-March to mid-April. In contrast, coefficient estimates for visits of food and beverage stores (Figure 8d), health care and social assistance (Figure 8e) and parks (Figure 8f) are insignificant and substantially smaller, suggesting that there is no differential reduction in these types of visits among individuals with differential friend-exposure to COVID-19. Reassuringly, all coefficient estimates in every panel are very close to zero prior to March, indicating no differential behavior before the outbreak of the pandemic. Note that since friend-exposure is defined within counties—and distancing policies were nearly always administered at the federal, state, or county level—differences in business closures across

[21] These buckets are the fraction of individuals 18 or younger, between 18 and 24, between 25-34, between 35-44, between 45-54, between 55-64, between 65-74 and above 75.

places are unlikely to drive our results. This set of results thus provides evidence that high friend-exposure to COVID-19 not only causes individuals to reduce their mobility in general, but causes them to do so in a way that is consistent with the objective of minimizing interactions. Individuals with high friend-exposure *choose* to participate in social distancing. These findings thus once again provide evidence against potential concerns that the observed effects of friend-exposure on social distancing might merely pick up differences in the ability to work from home or simply reflect responses to government restrictions.

5.3.2 Economic Effects By Type of Merchant

We next study the extent to which individuals who are more exposed to COVID-19 through their social network behave differently in terms of their spending. To do so, we use data on transactions and spending behavior provided by Facteus. These data are collected primarily from smaller banks (e.g. N26, Simple), payroll cards, and government cards, and include aggregated (by ZCTA of the card holder) information on the number of cards used, the number of transactions made, and the total amount of all transactions combined for a given merchant. Although the sample is skewed towards younger individuals and lower-income households, and is therefore not perfectly representative, it is very large and includes information on cardholders from over 21,000 ZCTAs across the U.S., with an average of 5.7 million transactions and \$218.7 million in spending per day. Thus, these data allow us to make meaningful statements about the U.S. population.[22]

In Appendix Figure A12, we first plot two spending outcomes over time: the log of one plus the total amount spent in USD and the log of one plus the total number of transactions. At the beginning of the year, the number of transactions and the total amount of money spent were relatively low, but both increase in late January before peaking in mid-February. This peak can be explained by tax rebate season: because our data over-represents low-income households, spending is particularly responsive to the receipt of a large cash rebate. As the pandemic begins in the U.S., spending and transactions decrease sharply, reaching a low at the end of March. Both then slowly increase in the middle of April when many Americans received stimulus checks (or in anticipation thereof). Given these seasonal fluctuations, in order to estimate the effect of friend-exposure to COVID-19 on spending behavior we thus build on equation 11 and estimate a triple-diff in order to smooth out seasonal fluctuations. We do so by using expenditure data from 2020 and 2019.

$$
\begin{aligned}
Y_{it} = \beta_0 &+ \beta_{1t} \times HighExp_i \times week_t \times year2020_t \\
&+ \beta_{2t} \times HighExp_i \times week_t \\
&+ \beta_{3t} \times X_{it} \\
&+ \beta_{4t} \times Cov_i \times week_t \times year2020 + \epsilon_{it}
\end{aligned}
\tag{12}
$$

$HighExp_i$, $week_t$, and ϵ_{it} are defined as before. Y_{it} is defined as the log of one plus the total number of

[22] These statistics are calculated after excluding the top 0.001% of ZCTAs, outliers that exhibit spending behavior unlikely to represent the true spending behavior of individuals living in these areas. Also note that because of data quality issues, we focus on spending data until April 16, 2020.

transactions made by cardholders residing in ZCTA i during week t.[23] $year2020_t$ is an indicator equal to one for the year 2020 and zero other otherwise. X_{it} is a vector of county-time fixed effects as well as ZCTA fixed effects. Cov_i is a vector of covariates as in equation 11. We again cluster standard errors at the ZCTA-level.

[Figure 9]

Figure 9 shows our β_{1t} estimates. In Figure 9a we present results aggregating across all types of merchants. We cannot detect large effects of friend-exposure to COVID-19 on overall transcation behavior. More importantly, these data again allow for a disaggregated analysis of spending behavior. We thus re-estimate equation 12 for different types of merchants to capture the extent to which individuals with greater friend-exposure to COVID-19 change their transaction behavior to reduce social interactions. We focus on two merchants in particular: Starbucks and Amazon. Transactions made at Starbucks are almost exclusively in-person (and discretionary) while transactions made at Amazon are exclusively online. Thus, we would expect individuals seeking to socially distance to reduce their shopping at Starbucks more than their shopping at Amazon. Figure 9b shows patterns consistent with this hypothesis. Before the pandemic, there are no notable differences in spending between groups at either merchants, but after the outbreak places with high friend-exposure are close to 0.1 log points less likely to make transactions at Starbucks than low friend-exposure places. This effect is large and statistically significant. In contrast, the coefficient estimates for Amazon spending are all statistically insignificant and very close to zero. This case study suggests that those with higher friend-exposure to COVID-19 reduced discretionary in-person spending without changing their online spending patterns where no physical interactions are required.

In sum, in this section, we use disaggregated information on POI visits and purchases to show that the overall reductions in mobility are consistent with the objective of social distancing and avoiding nonessential interactions: while the effects of friend-exposure to COVID-19 on nonessential visits are large, negative, and highly significant, effects are near-zero and insignificant for essential places. Similarly, in places with high friend-exposure, we observe a great reduction in transactions made at Starbucks (which involve in-person interactions) and no effect on transactions made on Amazon (which are entirely online).

6 Conclusion

In this paper, we use de-identified data from Facebook to explore how personal connections shape social distancing behavior during the COVID-19 pandemic. We find that U.S. users whose friends lived in areas with worse coronavirus outbreaks on March 15 reduced their mobility in subsequent months more than otherwise similar users without friends in areas affected by COVID-19. As the outbreak progressed in the U.S., users with more friends in emerging hotspots in one month were more likely to reduce their mobility in the same month than others. Because we measure mobility at the individual level, we are able to rule out various alternative explanations that might confound an observed relationship between social networks and distancing behavior at the aggregate level. We are also able to explore how personal

[23] All of the results shown below are robust to using the log of one plus the total amount spent in USD as outcome variable.

connections shape distancing behavior differentially across demographics; for example, the effect of having friends in areas with many COVID-19 cases on distancing behavior is substantially larger for younger users and for users with college experience.

We also use data on Facebook posts and group memberships, as well as public disaggregated mobility and spending data, to illuminate the mechanisms driving these relationships. We find evidence that friend-exposure to COVID-19 raises awareness about the risks of the disease, thereby inducing individuals to participate in mitigating public health behavior. Specifically, users with higher friend-exposure to COVID-19 are more likely to post about the coronavirus and are less likely to oppose distancing in these posts. These users are also less likely to join Facebook groups advocating for a reopening of the economy during the early months of the pandemic. At the zip code level, friend-exposure to COVID-19 results in substantial decreases in visits to restaurants, bars, and places related to the arts, entertainment, and recreation; by contrast, we observe no effects for visits to essential places such as grocery stores, or for places of health care and social assistance.

We conclude that an individual's personal connections play an important role in shaping their health behaviors during a pandemic. In the context of the current outbreak, these results add important insights to a growing literature that explores the factors affecting individuals' social distancing behaviors. More broadly, our work illustrates how data from online social networks can help social scientists overcome important measurement challenges and better understand the determinants of public health outcomes.

7 Primary Tables and Figures

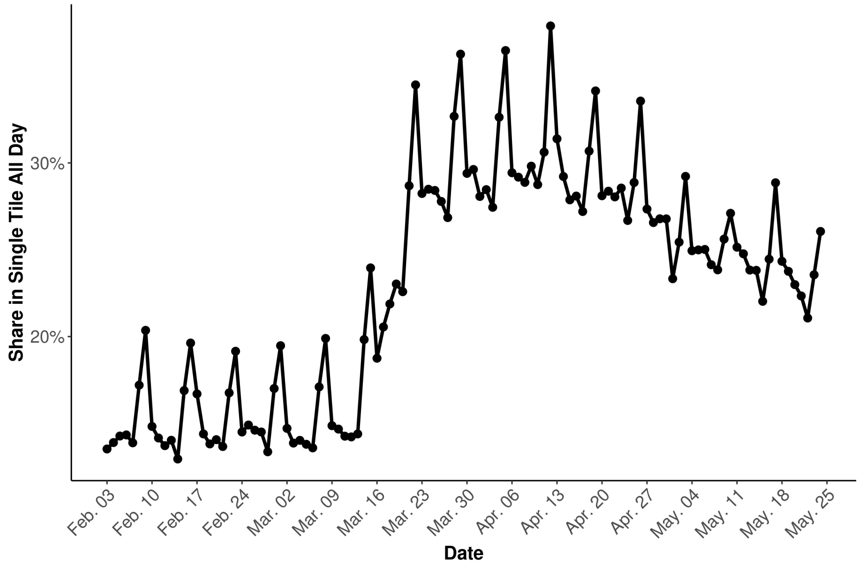
Table 1: Summary Characteristics - Mobility Sample

	Mean	SD	P10	P25	P50	P75	P90
Age	43.58	14.93	26	32	42	54	63
Female	0.53	0.50	0	0	1	1	1
Has College	0.53	0.50	0	0	1	1	1
Has iPhone	0.25	0.43	0	0	0	0	1
Has Tablet	0.53	0.50	0	0	1	1	1
Zip Code Income	$58,792	$21,961	$36,160	$43,648	$54,000	$69,203	$88,096
Number of Friends	532.80	326.61	193	276	441	718	1047
Friend Exposure to Cases	10.35	19.34	0.74	1.77	4.49	11.12	26.31
Staying at home (Feb)							
- All	18.33	29.35	0	0	0	28.57	66.67
- Weekend	19.39	34.44	0	0	0	50.00	100.00
- Weekday	16.83	29.80	0	0	0	20.00	66.67
Bing tiles visited (Feb)							
- All	10.96	9.07	1.57	3.43	9.00	15.86	23.43
- Weekend	10.57	9.79	1.00	3.00	7.50	15.50	24.50
- Weekday	11.34	9.77	1.50	3.40	9.00	16.20	24.60

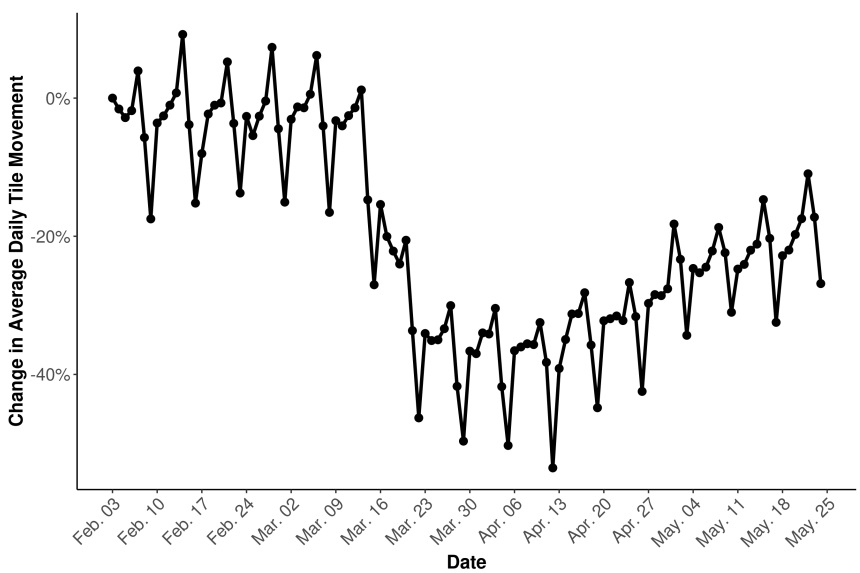
Note: Table presents summary statistics describing individuals analyzed in our mobility sample of users. Individual-level characteristics include age, gender, whether the user has a college listed on Facebook, whether the user primarily accesses Facebook mobile from an iPhone, whether the individual has accessed Facebook from a tablet, number of friends, friend-exposure to COVID cases on March 15th, and patterns of mobility during the week of February 25th to March 2nd. The table also includes information on the users' home ZCTA 2018 median household income. [Return to text]

Figure 1: Mobility Over Time

(a) Probability of Staying at Home



(b) Average Number of Tiles Visited



Note: Figures show average patterns of mobility from February 3rd to May 24th according to two metrics described in Section 2.1. Panel (a) shows the probability of staying at home and panel (b) shows the percent change in average number of tiles visited from February 3rd. [Return to text]

Table 2: Change in Probability Staying at Home

	Stay at Home					
	All		Weekdays		Weekends	
	Level Feb	ΔFeb-Apr	Level Feb	ΔFeb-Apr	Level Feb	ΔFeb-Apr
Overall	18.33	13.68	16.83	13.58	19.39	14.29
By Age Group						
18-34	14.49	13.17	13.23	13.30	14.54	13.12
35-54	16.57	13.22	14.95	13.13	17.98	13.79
55+	25.68	14.99	24.10	14.64	27.04	16.29
By Gender						
Female	20.15	15.68	18.72	15.76	21.19	15.89
Male	16.21	11.33	14.62	11.02	17.26	12.39
By College						
Has College	17.66	15.27	16.11	15.33	18.94	15.48
No College	19.10	11.84	17.66	11.56	19.90	12.89
By Zip Code Income						
Bottom Tertile	19.27	11.54	17.84	11.29	19.96	12.50
Middle Tertile	18.19	12.78	16.69	12.65	19.33	13.43
Top Tertile	17.55	16.69	15.98	16.76	18.88	16.85
By County Total Cases/Population						
Bottom Tertile	18.62	10.86	17.15	10.65	19.75	11.66
Middle Tertile	17.97	15.17	16.56	15.07	18.71	15.84
Top Tertile	18.15	16.75	16.55	16.80	19.30	17.02
By Exposure through Friends						
High Exposure	18.46	14.82	16.97	14.77	19.45	15.34
Low Exposure	18.21	12.55	16.70	12.40	19.33	13.23

Note: Table describes changes in social distancing across different user characteristics. Social distancing is measured as the average probability of staying home. Characteristic splits include age group, gender, whether the user has a college listed on Facebook, the tertile of home ZCTA median household income, the tertile of county-level cases per resident as of March 15th, and whether the log of friend-exposure to COVID cases on March 15th is above (high exposure) or below (low exposure) the user's home ZCTA median. Columns 1, 3, and 5 show the levels for the week of February 25th to March 2nd (prior to the pandemic). Columns 2, 4, 6 show the difference between the week of April 14th to 20th (during the early stages of the pandemic) and this baseline. Columns 1 and 2 include movement on all days; 3 and 4 include weekdays only; and 5 and 6 include weekends only. [Return to text.]

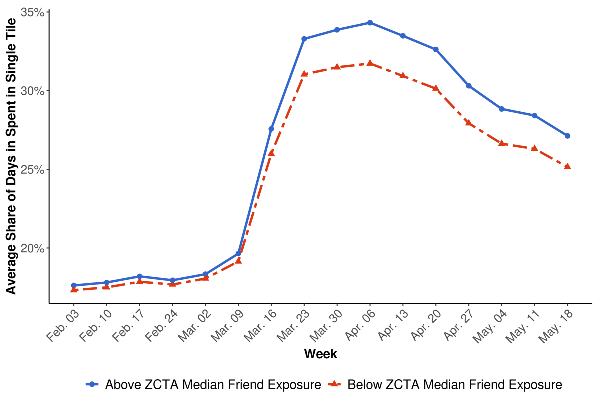
Table 3: Relationship Between Friend-Exposure and Individual Characteristics

	DV: log(Friend Exposure)				
Age Group					
35-54	-0.005***			0.017***	
	(0.002)			(0.001)	
55+	-0.055***			0.022***	
	(0.004)			(0.001)	
Female	-0.100***			-0.015***	
	(0.001)			(0.001)	
Has College	0.185***			0.052***	
	(0.003)			(0.001)	
Has iPhone	0.090***			0.004***	
	(0.002)			(0.001)	
Has Tablet	0.045***			0.014***	
	(0.001)			(0.000)	
Zip Code Income					
Middle Tertile	0.120***				
	(0.019)				
Top Tertile	0.415***				
	(0.019)				
County Cases/Pop					
Middle tertile	1.030***				
	(0.015)				
Top Tertile	1.676***				
	(0.020)				
Zip Code FE		Y	Y	Y	
Other network exposure FE			Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone					Y
R-Squared	0.377	0.671	0.851	0.851	0.873
Sample Mean	1.458	1.458	1.458	1.458	1.487
N	6,803,762	6,803,761	6,803,761	6,803,761	6,400,738

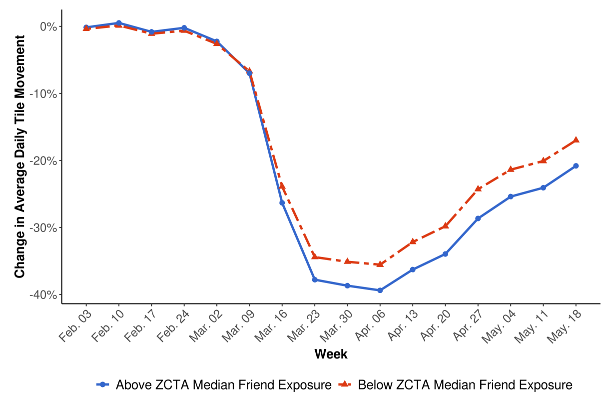
Note: Table shows results from regressing various measures on the log of friend-exposure to COVID cases on March 15th. Each observation is an individual. Column 1 includes controls for age groups, gender, whether the individual has a college listed on Facebook, whether the individual primarily accesses mobile Facebook from an iPhone, whether the individual has accessed Facebook from a tablet, the tertile of home ZCTA median household income, and the tertile of home county cases per resident as of March 15th. Column 2 includes only ZCTA fixed effects. Column 3 adds percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Column 4 adds back the individual-level controls from column 1. Column 5 adds fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

Figure 2: Effects of Friend-Exposure to COVID-19 on Mobility Behavior

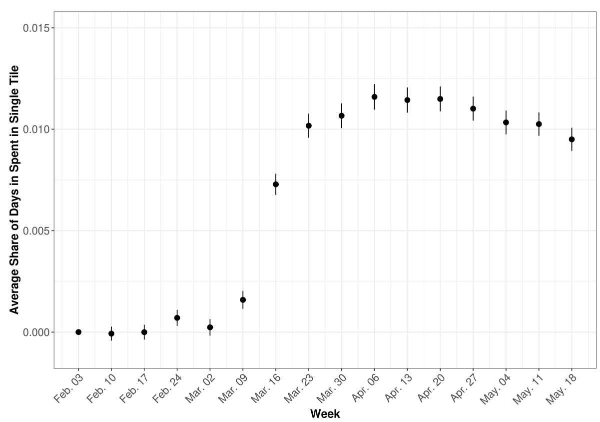
(a) Times Series: Probability of Staying Home



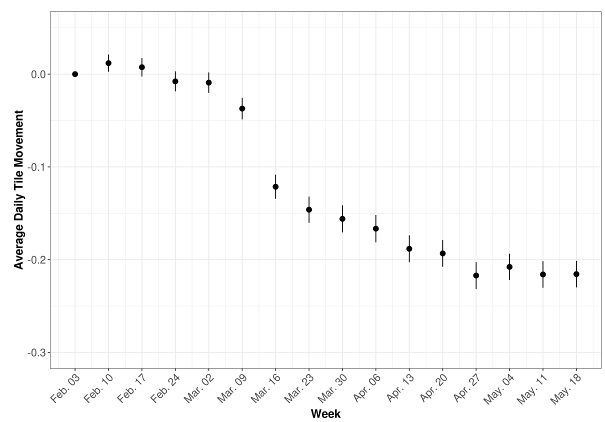
(b) Time Series: Average Tiles Visited



(c) Diff-In-Diff: Probability of Staying Home



(d) Diff-In-Diff: Average Tiles Visited



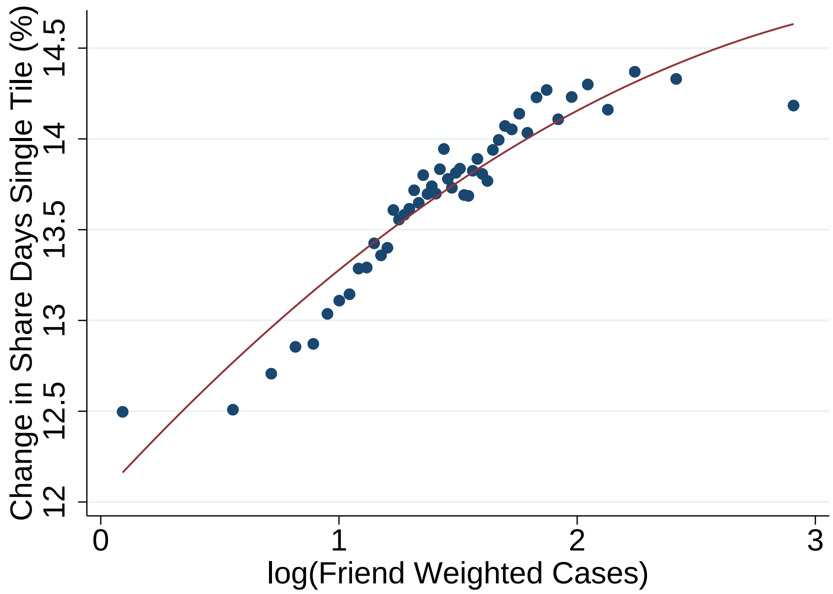
Note: Figures show the relationship between friend-exposure to COVID-19 on March 15th cases and mobility behavior. Panels (a) and (b) show weekly averages of the probability of staying at home and the average number of tiles visited from the week of February 3rd to the week of May 18th, seperately for individuals above their ZCTA median friend-exposure and below their median friend-exposure. Panels (c) and (d) show coefficients estimated using the difference-in-differences setup specified in equation 3 with the outcome variable as the probability of staying at home and the average number of tiles visited, respectively. Both specifications include fixed effects at the individual level as well as the following groups interacted with week: ZCTA, age group; gender; has college listed on Facebook; has iPhone; has tablet; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Standard errors are clustered by ZCTA. [Return to text]

Table 4: Social Distancing by Demographics: Probability of Staying at Home

	DV: Δ Stay at Home (Feb - Apr)							
Age Group								
35-54	-0.394***	-0.360***		-0.347***				
	(0.036)	(0.032)		(0.032)				
55+	1.381***	1.544***		1.589***				
	(0.045)	(0.038)		(0.037)				
Female	4.404***	4.718***		4.852***				
	(0.031)	(0.030)		(0.029)				
Has College	2.876***	2.538***		2.228***				
	(0.029)	(0.026)		(0.025)				
Has iPhone	0.147***	-0.332***		-0.465***				
	(0.035)	(0.032)		(0.032)				
Has Tablet	0.936***	0.900***		0.851***				
	(0.024)	(0.023)		(0.023)				
Zip Code Income								
Middle Tertile	1.001***							
	(0.109)							
Top Tertile	3.671***							
	(0.109)							
County Cases/Pop								
Middle tertile	3.816***							
	(0.089)							
Top Tertile	5.105***							
	(0.120)							
log(Friend Exposure)			0.923***	0.849***	0.878***	0.825***	0.919***	0.961***
			(0.026)	(0.026)	(0.028)	(0.037)	(0.030)	(0.045)
Zip Code FE		Y	Y	Y				
Other Network Exposure FE			Y	Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone					Y	Y	Y	Y
College FE								Y
Sample						Weekend	Weekday	College
R-Squared	0.021	0.041	0.035	0.044	0.175	0.159	0.174	0.193
Sample Mean	13.683	13.683	13.683	13.683	13.800	14.415	13.704	15.852
N	6,804,168	6,804,167	6,803,761	6,803,761	6,400,738	5,808,187	6,309,820	2,616,959

Note: Table shows results from regression 4. Each observation is an individual. The outcome in all columns is the change in probability of staying at home from the week of February 25-March 2, 2020 (prior to the pandemic) to April 14-20, 2020. Column 1 includes controls for age groups, gender, whether the individual has a college listed on Facebook, whether the individual primarily accesses Facebook mobile from an iPhone, whether the individual has accessed Facebook from a tablet, the tertile of home ZCTA median household income, and the tertile of home county cases per resident as of March 15th. Column 2 adds ZCTA fixed effects, but maintains the individual level controls. Column 3 includes only the log of friend-exposure to COVID cases on March 15th; ZCTA fixed effects; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Column 4 adds back the individual-level controls from column 1. Column 5 adds fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. In Column 6 the outcome is measured using weekend movement and in column 7 using weekday movement. Column 8 limits to individuals that attended a college, limiting to colleges with more than 100 individuals, and adds a fixed effect for each individual college. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

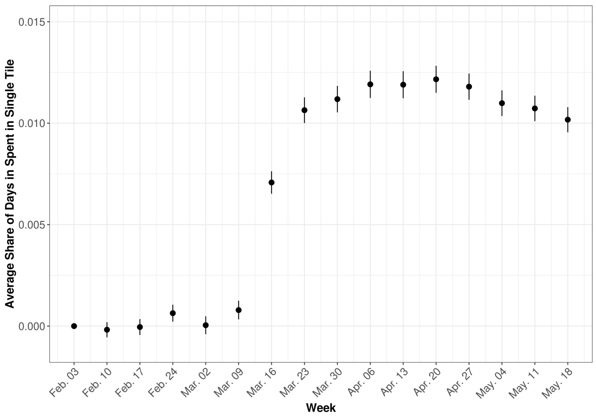
Figure 3: Probability of Staying at Home vs. Friend-Exposure to COVID-19



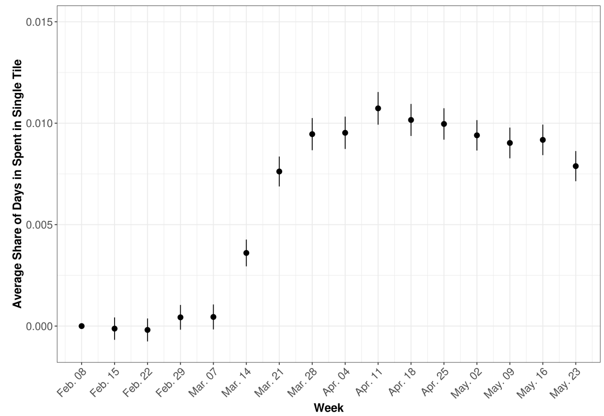
Note: Figure shows a binned scatter plot of the log of friend weighted friend-exposure to COVID-19 on March 15th and the change in probability of staying at home from the week of February 25-March 2, 2020 (prior to the pandemic) to April 14-20, 2020. The plot controls for fixed effects constructed from interacting ZCTA, age group, gender, has a college listed on Facebook, has iPhone, and has tablet. It also controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. [Return to text]

Figure 4: Robustness: Effects of Friend-Exposure to COVID-19 on Probability of Staying Home

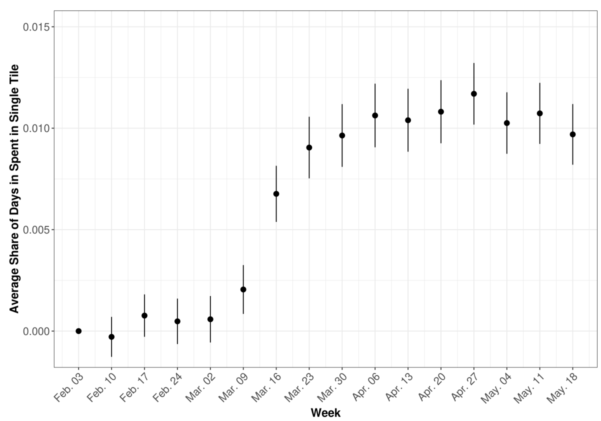
(a) Weekdays



(b) Weekends

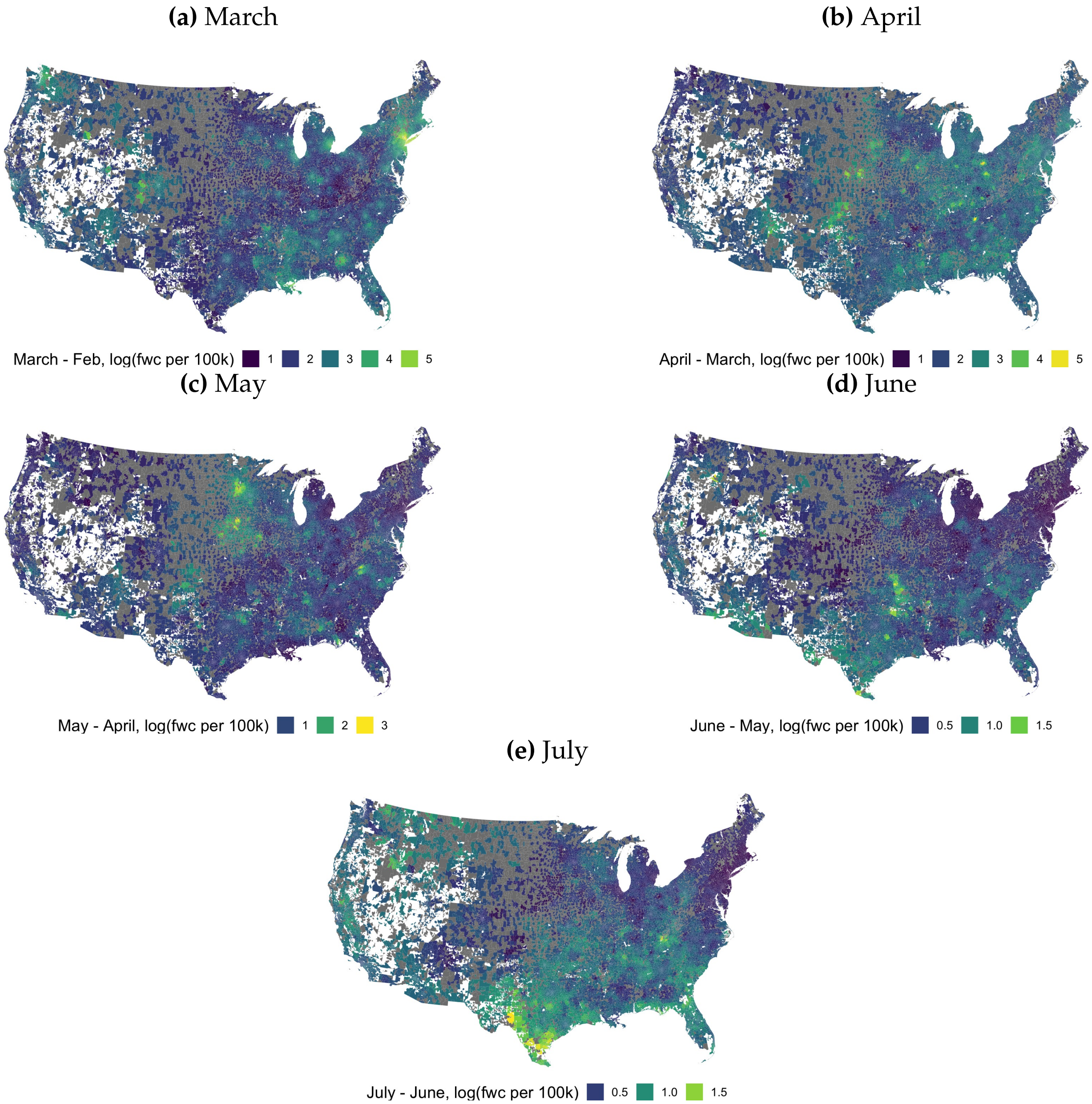


(c) Controlling for Exact College



Note: Figures show coefficients estimated using the difference-in-differences setup specified in equation 3 with the outcome variable as the probability of staying at home. The outcome is measured on weekdays in panel (a) and weekends in panel (b). Panel (c) limits to individuals that attended college, limiting to colleges with more than 100 individuals, and adds a fixed effect for each individual college interacted with week. All specifications include fixed effects at the individual level as well as the following groups interacted with week: ZCTA; age group; gender; has a college listed on Facebook; has iPhone; has tablet; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Standard errors are clustered by ZCTA. [Return to text]

Figure 5: Local Variation in Friend-Exposure to COVID-19 by Month



Note: Figures show average changes in log friend-exposure to COVID-19 cases per 100k residents (as described in equation 6) by ZCTA for the continental U.S. Exposure is measured on the last Friday of each month: panel (a) shows that change from February to March, panel (b) shows the change from March to April, panel (c) shows the change from April to May, panel (d) shows the change from May to June and panel (e) shows the change from June to July. The sample of users is restricted to those for whom location can be observed at the end of each of the two relevant months. Darker blue indicates a smaller increase and brighter green and yellow indicate a larger increase. Grey areas are excluded because of small populations, white areas are not assigned ZCTAs by the Census Bureau. [Return to text]

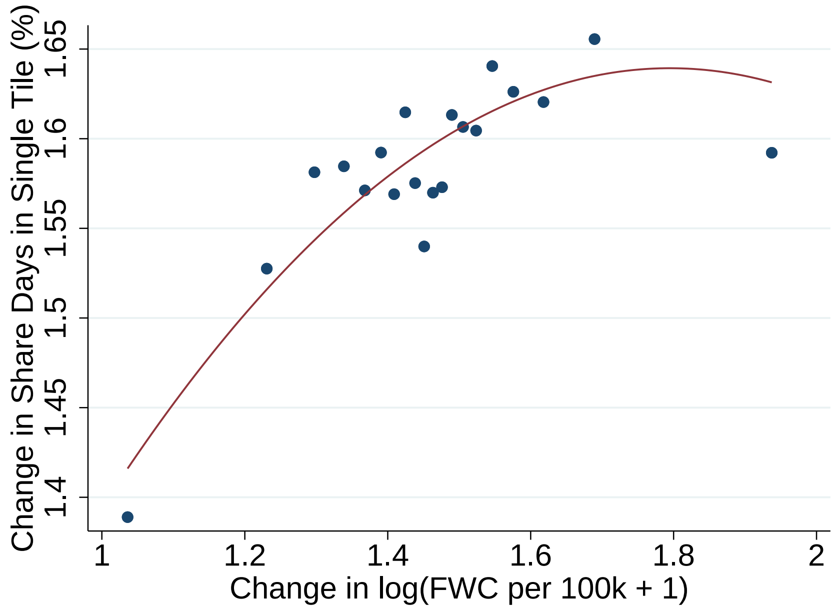
Table 5: Determinants of Δ Friend-Exposure to COVID-19 by Month

	Monthly Change in log(Friend Exposure + 1)									
	March	April	May	June	July	March	April	May	June	July
Age Group										
35-54	0.040***	0.014***	-0.013***	-0.008***	-0.001**	0.015***	0.005***	-0.004***	-0.003***	0.001***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
55+	0.076***	0.015***	-0.026***	-0.018***	-0.004***	0.024***	0.007***	-0.006***	-0.006***	0.001
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.021***	0.006***	0.003***	0.006***	0.004***	-0.004***	0.002***	0.001***	0.000**	-0.000***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Has College	0.039***	-0.031***	-0.007***	-0.008***	-0.004***	0.003***	-0.013***	-0.003***	0.002***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Has iPhone	0.011***	0.005***	-0.007***	0.008***	0.013***	0.002***	-0.002***	-0.000	0.002***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Has Tablet	0.005***	-0.009***	-0.008***	-0.001***	0.002***	0.003***	-0.002***	-0.002***	-0.000***	0.001***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Network-Exposure Median HH Income ($k)	0.015***	-0.004***	0.001***	-0.009***	-0.013***					
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)					
Network-Exposure Population Density	349.495***	-34.302***	-65.142***	-71.383***	-88.601***					
(residents/meter^2)	(5.622)	(1.527)	(1.280)	(1.582)	(1.764)					
Network-Exposure Fraction of Pop. Urban	1.112***	-0.076**	-0.263***	0.319***	0.456***					
	(0.035)	(0.019)	(0.016)	(0.014)	(0.016)					
Zip Code Income										
Middle Tertile	-0.034***	-0.017**	0.007	-0.011**	-0.004					
	(0.011)	(0.008)	(0.006)	(0.005)	(0.006)					
Top Tertile	0.002	-0.026***	-0.008	-0.006	0.005					
	(0.011)	(0.008)	(0.005)	(0.005)	(0.006)					
Zip Code FE						Y	Y	Y	Y	Y
Other Network Exposure FE						Y	Y	Y	Y	Y
R-Squared	0.560	0.044	0.117	0.215	0.281	0.877	0.680	0.728	0.781	0.822
Sample Mean	2.800	2.303	0.810	0.476	0.615	2.800	2.303	0.810	0.476	0.615
N	7,090,255	6,981,142	6,571,618	6,251,614	5,859,728	7,090,254	6,981,141	6,571,617	6,251,614	5,859,728

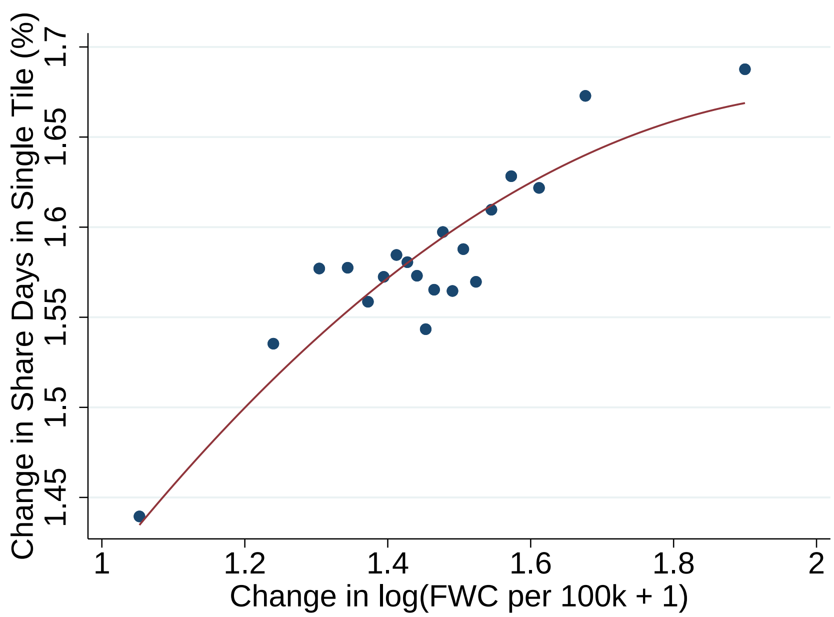
Note: Table shows results from regressing various measures on the change in log of friend-exposure to COVID cases per 100k residents (as described in 6) between the last Fridays of each month (e.g. Feb to March in column 1). Columns 1-5 include age groups; gender; whether the individual has a college listed on Facebook; whether the individual primarily accesses mobile Facebook from an iPhone; whether the individual has accessed Facebook from a tablet; friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas; and the tertile of ZCTA-level median household income. Columns 6-10 control for ZCTA fixed effects and percentiles of the friend weighted exposure metrics. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

Figure 6: Δ Probability of Staying at Home vs. Δ Friend-Exposure to COVID-19

(a) Without User Fixed-Effects



(b) With User Fixed-Effects



Note: Figure shows a binned scatter of the change in log friend-exposure to COVID-19 cases per 100k residents (as described in equation 6) and the change in the probability of staying home. Each observation is a unique individual and month for the months of March, April, May, June and July. Change in exposure is measured as of the last Friday of each month. Change in movement patterns is measured using the Tuesday to Monday week that includes each of these Fridays. Panel (a) includes controls for fixed effects constructed from interacting month, ZCTA, age group, gender, college background, and iPhone and tablet ownership. It also controls for month interacted with percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Panel (b) includes the same controls and also adds user fixed effects. [Return to text]

Table 6: Effects of Friend-Exposure by Month: Δ Probability of Staying at Home

	Monthly Change in Δ Stay at Home						
	All months	All months	March	April	May	June	July
Δ log(Friend Exposure + 1), All Months	0.206***	0.261***					
	(0.029)	(0.032)					
Δ log(Friend Exposure + 1), March			0.207***	0.006	-0.076**	0.097	0.037
			(0.046)	(0.040)	(0.048)	(0.054)	(0.064)
Δ log(Friend Exposure + 1), April				0.035	0.096	0.329***	0.069**
				(0.052)	(0.056)	(0.061)	(0.071)
Δ log(Friend Exposure + 1), May					0.379***	0.044	-0.057
					(0.082)	(0.078)	(0.094)
Δ log(Friend Exposure + 1), June						0.854***	-0.329*
						(0.114)	(0.127)
Δ log(Friend Exposure + 1), July							0.323**
							(0.138)
User FE		Y					
Other Network Exposure FE	Y		Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y	Y	Y	Y
R-Squared	0.211	0.287	0.174	0.141	0.150	0.146	0.145
Sample Mean	1.611	1.456	14.214	-0.923	-5.989	-1.068	0.679
N	30,742,008	29,777,929	6,688,448	6,579,359	6,169,176	5,848,722	5,456,303

Note: Table shows results from regression 8. Each observation is an individual. The outcome variable is the change in the probability of staying home between the final weeks of a given month and the previous months' final week: February 25-March 2 for February; March 24-March 30 for March; April 21-April 27 for April; May 26-June 1; June 23-June 29; July 21-July 28. The sample of users is restricted to those for whom location can be observed at the end of each of the two relevant months. In the first two columns, we pool the changes over all months from March to July. In the next columns, we consider changes by month. In all columns we control for interactions of ZCTA fixed effects, age groups, gender, whether the individual has a college listed on Facebook, whether the individual primarily accesses mobile Facebook from an iPhone, and whether the individual has accessed Facebook from a tablet. Columns 1 and 3-7 also include fixed effects for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. In column 2, we include user fixed effects. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

Table 7: Heterogeneity of Friend-Exposure Effects: Probability of Staying at Home

	%Δ Stay at Home					
log(Friend Exposure) x I(Age < 35)	1.241***					
	(0.042)					
log(Friend Exposure) x I(Age 35-55)	0.960***					
	(0.033)					
log(Friend Exposure) x I(Age > 55)	0.412***					
	(0.038)					
log(Friend Exposure) x Female		0.949***				
		(0.032)				
log(Friend Exposure) x Male		0.796***				
		(0.033)				
log(Friend Exposure) x College			1.321***			
			(0.034)			
log(Friend Exposure) x No College			0.443***			
			(0.031)			
log(Friend Exposure) x Zip Income First Tertile				0.386***		
				(0.037)		
log(Friend Exposure) x Zip Income Second Tertile				0.794***		
				(0.036)		
log(Friend Exposure) x Zip Income Third Tertile				1.608***		
				(0.045)		
log(Friend Exposure) x County Cases First Tertile					0.676***	
					(0.030)	
log(Friend Exposure) x County Cases Second Tertile					1.384***	
					(0.058)	
log(Friend Exposure) x County Cases Third Tertile					1.245***	
					(0.055)	
log(Friend Exposure - Rank 1 - 25)						0.204***
						(0.017)
log(Friend Exposure - Rank 26 - 50)						0.112***
						(0.017)
log(Friend Exposure - Rank 51 - 75)						0.082***
						(0.017)
log(Friend Exposure - Rank 76 - 100)						0.098***
						(0.017)
Other Network Exposure FE	Y	Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y	Y	Y
R-Squared	0.175	0.175	0.175	0.175	0.175	0.177
Sample Mean	13.800	13.800	13.800	13.800	13.800	14.488
F Test (Rank 1-25 = Rank 76-100)						17.328***
N	6,400,738	6,400,738	6,400,738	6,400,738	6,400,738	5,684,469

Note: Table shows results from regressions of friend-exposure to COVID-19 on March 15th, interacted with individual characteristics, on the percentage change in the probability of staying at home. Friend-exposure is interacted with age groups in rows 1-3; gender in rows 4-5; whether the individual has a college listed in Facebook in rows 6-7; ZCTA median household income in rows 8-10; county-level cases of COVID-19 in rows 11-13; and friend rank (i.e. a measure for how close friends are) in rows 14-16. All columns include controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. All columns include fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

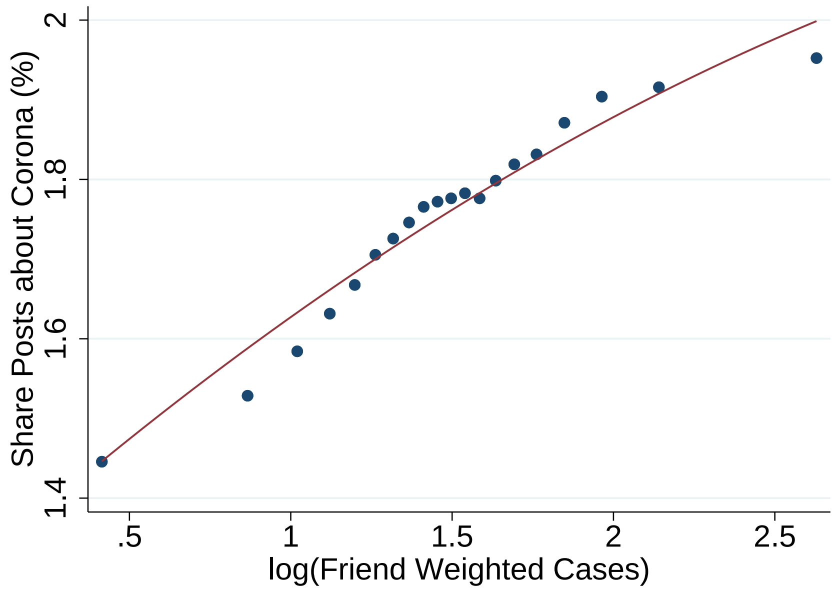
Table 8: Posting Behavior and Group Membership

	DV: Share Posts about Corona (Feb - Apr)		DV: Share "Signed Posts" Opposed to Distancing (Feb - Apr)		Δ Sentiment (Feb - Apr) All Posts		DV: Member "Reopen Group" by June 28, 2020	
log(Friend Exposure)	0.324***	0.249***	-1.659***	-1.929***	-0.109***	-0.094***	-0.003	-0.129***
	(0.006)	(0.006)	(0.107)	(0.245)	(0.016)	(0.025)	(0.018)	(0.007)
Age Group								
35-54	0.579***		-2.196***		-0.480***		0.767***	
	(0.005)		(0.168)		(0.026)		(0.011)	
55+	0.351***		4.667***		-0.031		0.851***	
	(0.005)		(0.194)		(0.030)		(0.012)	
Female	-0.266***		-17.713***		0.942***		-0.582***	
	(0.003)		(0.142)		(0.024)		(0.010)	
Has College	0.637***		-2.392***		-0.283***		-0.188***	
	(0.004)		(0.141)		(0.023)		(0.006)	
Has iPhone	0.137***		-7.215***		-0.150***		0.019***	
	(0.003)		(0.135)		(0.023)		(0.006)	
Has Tablet	0.028***		-1.997***		0.039*		-0.048***	
	(0.003)		(0.125)		(0.023)		(0.003)	
Zip Code Income								
Middle Tertile	0.069***		-0.886***		-0.075*		0.211***	
	(0.013)		(0.229)		(0.031)		(0.041)	
Top Tertile	0.269***		-1.946***		-0.121***		0.379***	
	(0.016)		(0.250)		(0.035)		(0.044)	
County Cases/Pop								
Middle tertile	-0.064***		1.458***		0.027		0.219***	
	(0.014)		(0.256)		(0.037)		(0.049)	
Top Tertile	-0.097***		1.049***		0.034		0.204***	
	(0.013)		(0.240)		(0.036)		(0.048)	
Percentiles of Total Number of Groups (Feb 2020)							Y	Y
Other Network Exposure FE	Y	Y	Y	Y	Y	Y	Y	Y
Zip Code FE								
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone		Y		Y		Y		Y
Sample	People With Any Posts February - April		People With "Signed Posts" February - April		People With Posts between Feb 3 and May 3		People With Group Memberships	
R-Squared	0.013	0.060	0.087	0.445	0.000	0.118	0.013	0.074
Sample Mean	1.750	1.755	39.806	35.979	-1.817	-1.823	1.217	1.216
N	34,828,054	34,528,373	546,499	277,776	11,209,068	10,777,790	119,384,394	119,145,833

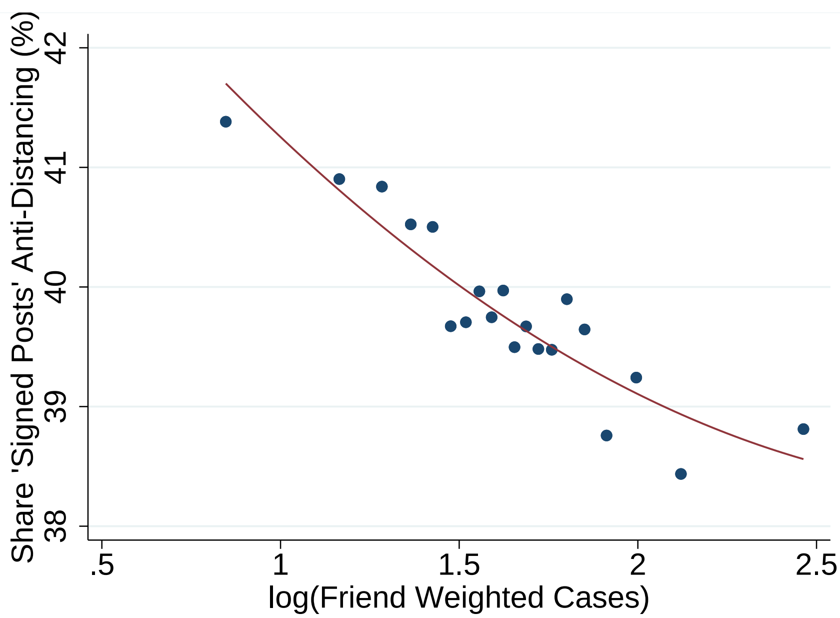
Note: Table shows results from regression 9 and 10. Each observation is an individual. The outcome in columns 1-2 is the percentage of individual posts that are about COVID-19; in columns 3-4 it is the percentage of pro- or anti-distancing posts that are anti-distancing; in columns 5-6 it is the change in the average sentiment of the posts from February 3rd through 23rd to April 6th through 26th; in columns 7-8 it is whether the individual was a member of a 'Reopen' Facebook group as of June 28th. For ease of interpretation and because of small magnitudes, we rescale coefficients and standard errors by 100, so that they correspond to percentages. Post classification is based on the regex in Appendix C.1. Sentiment is measured on a scale from -100 to 100 using the VADER algorithm described in Hutto and Gilbert (2014). Group classification is determined by the regular expression described in Appendix C.2. Columns 1, 3, 5 and 7 include controls for the log of friend-exposure to COVID-19 on March 15th; age groups; gender; whether the individual has a college listed on Facebook; whether the individual primarily accesses Facebook mobile from an iPhone; whether the individual has accessed Facebook from a tablet; the tertile of home ZCTA median household income; the tertile of home county cases per resident as of March 15th; and percentiles of friend-exposures (as described in equation 2) for median household income, population density, and the share of the population living in urban areas. Columns 2, 4, 6, 8 add fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. For the group based analysis in columns 7 and 8 we also include fixed effects for the percentile of the number of groups an individual was in as of February 2020. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

Figure 7: Posting Behavior vs. Friend-Exposure to COVID-19

(a) Percentage of Posts About COVID-19



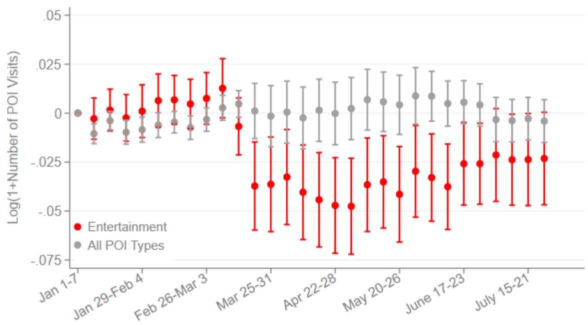
(b) Percentage of Signed Posts Opposing Distancing



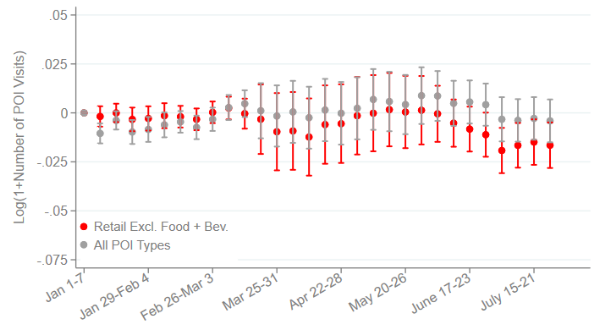
Note: Figures show binned scatter plots of the log of friend-exposure to COVID-19 on March 15th and Facebook post based measures. The outcome variable in panel (a) is the percentage of individual posts that are about COVID-19 and in panel (b) it is the percentage of pro- or anti-lockdown posts that are anti-distancing. Post classification is based on the regex in Appendix C.1. The plots control for fixed effects constructed from interacting ZCTA, age group, gender, has college, has iPhone, and has tablet. They also control for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. [Return to text]

Figure 8: Coefficient Estimates for Different Types of POI Places

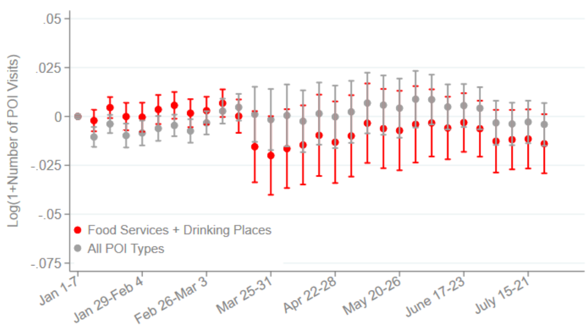
(a) Arts, Entertainment & Recreation



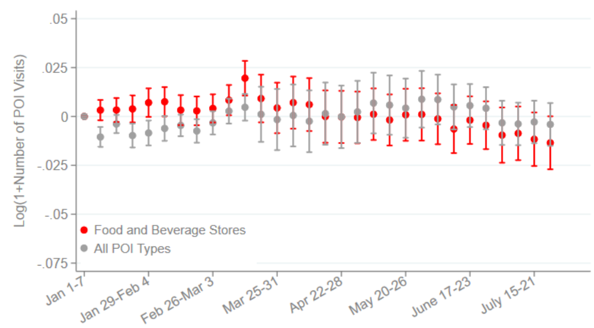
(b) Retail Trade, excl. Food & Beverage Stores



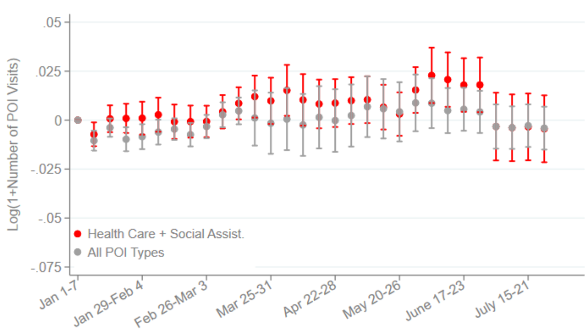
(c) Food Services & Drinking Places



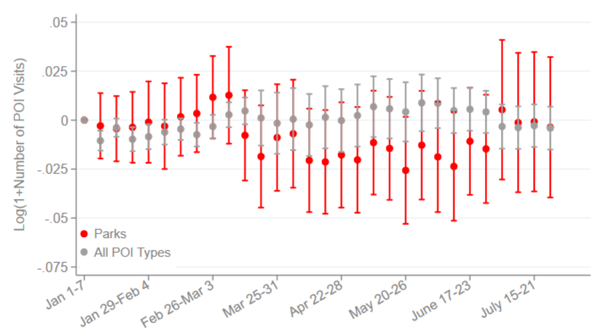
(d) Food & Beverage Stores



(e) Health Care & Social Assistance



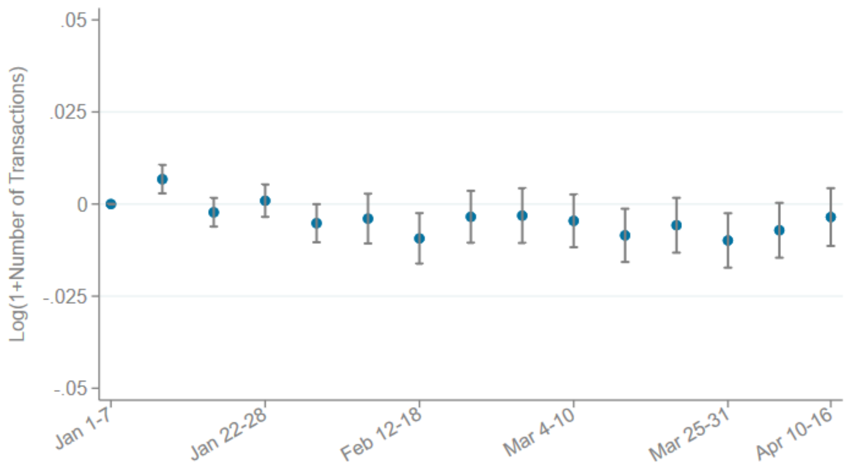
(f) Parks



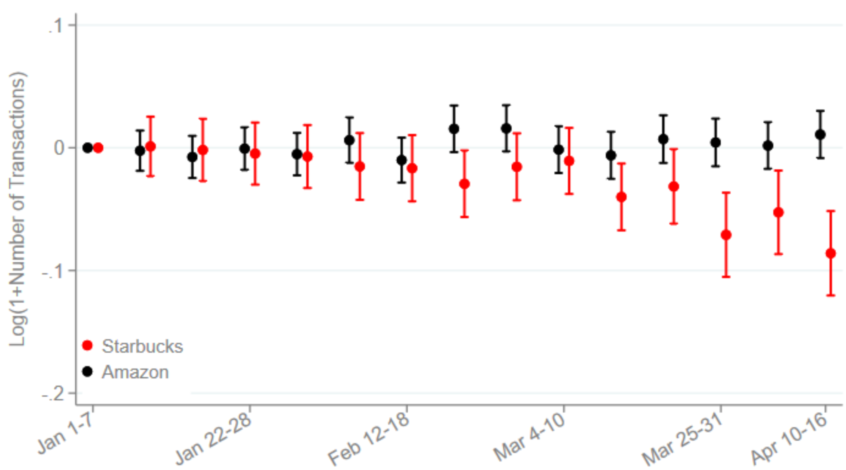
Note: Figures show coefficient estimates based on equation 11 for various types of POIs. For reference, we include estimates aggregating across all types of POIs in gray in all panels. We control for ZCTA fixed effects, county fixed effects interacted with week indicators as well as a rich set of covariates interacted with week indicators. These covariates are the fraction of people being male, the fraction of asian/black/white people, median household income, the fraction of individuals working in service occupations, the fraction of individuals working in production or transportation, the fraction of individuals working in management, arts or science, the fraction of individuals with a high school degree, some college education and a college degree as well as the fraction of households with high speed internet. We also include various age-related controls, i.e. the fraction of individuals 18 or younger, between 18 and 24, between 25-34, between 35-44, between 45-54, between 55-64, between 65-74 and above 75. All these control variables are obtained from the most recent 5-year ACS (2014-2018). In addition, we also control for ventiles of friend-exposure to other characteristics, namely income, population density (both from 2014-2018 ACS) and urbanity (from 2010 Census), again interacted with week indicators. Standard errors are clustered at the ZCTA level. [Return to text]

Figure 9: Coefficient Estimates for β_{1t} from equation 12

(a) Log Number of Transactions Made



(b) Log Number of Transactions at Amazon and Starbucks



Note: Figures present regression estimates of β_{1t} based on equation 12. In Panel (a) we present coefficient estimates for the log of one plus the number of transactions made. In addition to controlling for ZCTA fixed effects as well as county fixed effects interacted with week indicators, we also control for a rich set of covariates interacted with week indicators. These covariates are the fraction of people being male, the fraction of asian/black/white people, median household income, the fraction of individuals working in service occupations, the fraction of individuals working in production or transportation, the fraction of individuals working in management, arts or science, the fraction of individuals with a high school degree, some college education and a college degree as well as the fraction of households with high speed internet. We also include various age-related controls, i.e. the fraction of individuals 18 or younger, between 18 and 24, between 25-34, between 35-44, between 45-54, between 55-64, between 65-74 and above 75. All these control variables are obtained from the most recent 5-year ACS (2014-2018). We also control for ventiles of friend-exposure to other characteristics, namely income, population density (both from 2014-2018 ACS) and urbanity (from 2010 Census), again interacted with week indicators. In panel (b), we instead present coefficient estimates where the dependent variable is the log of one plus the number of transactions made at Starbucks and at Amazon. The black series corresponds to transactions made at Amazon while the red series corresponds to transactions made at Starbucks. The set of control variables included is identical to panel (a). In all regressions standard errors are clustered at the ZCTA level. [Return to text]

References

Georgij Alekseev, Safaa Amer, Manasa Gopal, Theresa Kuchler, JW Schneider, Johannes Stroebel, and Nils Wernerfelt. The effects of covid-19 on u.s. small businesses. Technical report, 2020.

Hunt Allcott, Levi Boxell, Jacob Conway, Billy Ferguson, Matthew Gentzkow, and Benny Goldman. Economic and health impacts of social distancing policies during the coronavirus pandemic. *Available at SSRN 3610422*, 2020a.

Hunt Allcott, Levi Boxell, Jacob C Conway, Matthew Gentzkow, Michael Thaler, and David Y Yang. Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. Working Paper 26946, National Bureau of Economic Research, April 2020b. URL http://www.nber.org/papers/w26946.

Michael Bailey, Rachel Cao, Theresa Kuchler, and Johannes Stroebel. The economic effects of social networks: Evidence from the housing market. *Journal of Political Economy*, 126(6):2224–2276, 2018a.

Michael Bailey, Rachel Cao, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. Social connectedness: Measurements, determinants, and effects. *Journal of Economic Perspectives*, 32(3):259–80, 2018b.

Michael Bailey, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel. House price beliefs and mortgage leverage choice. *The Review of Economic Studies*, 86(6):2403–2452, 2019a.

Michael Bailey, Patrick Farrell, Theresa Kuchler, and Johannes Stroebel. Social connectedness in urban areas. Working Paper 26029, National Bureau of Economic Research, 2019b.

Michael Bailey, Drew M Johnston, Theresa Kuchler, Johannes Stroebel, and Arlene Wong. Peer effects in product adoption. Working Paper 25843, National Bureau of Economic Research, 2019c.

Michael Bailey, Abhinav Gupta, Sebastian Hillenbrand, Theresa Kuchler, Robert Richmond, and Johannes Stroebel. International trade and social connectedness. Working paper, 2020a.

Michael Bailey, Theresa Kuchler, Dominic Russel, Bogdan State, and Johannes Stroebel. Social connectedness in europe. Working paper, 2020b.

Olivier Bargain and Ulugbek Aminjonov. Trust and compliance to public health policies in times of covid-19. 2020.

John M Barrios and Yael Hochberg. Risk perception through the lens of politics in the time of the covid-19 pandemic. Working Paper 27008, National Bureau of Economic Research, April 2020. URL http://www.nber.org/papers/w27008.

John M Barrios, Efraim Benmelech, Yael V Hochberg, Paola Sapienza, and Luigi Zingales. Civic capital and social distancing during the covid-19 pandemic. Working Paper 27320, National Bureau of Economic Research, June 2020. URL http://www.nber.org/papers/w27320.

Mark Blumenthal. The spread of coronavirus to red america is not shifting opinions, 2020. URL https://today.yougov.com/topics/politics/articles-reports/2020/06/09/spread-coronavirus-red-america-analysis.

Daniela Del Boca, Noemi Oggero, Paola Profeta, and Mariacristina Rossi. Women's work, housework and childcare, before and during covid-19. 2020.

Abel Brodeur, David M Gray, Anik Islam, Suraiya Jabeen Bhuiyan, et al. A literature review of the economics of covid-19. Technical report, Institute of Labor Economics (IZA), 2020a.

Abel Brodeur, Idaliya Grigoryeva, and Lamis Kattan. Stay-at-home orders, social distancing and trust. 2020b.

Adam Brzezinski, Valentin Kecht, David Van Dijcke, and Austin L Wright. Belief in science influences physical distancing in response to covid-19 lockdown policies. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2020-56), 2020.

Leonardo Bursztyn, Aakaash Rao, Christopher Roth, and David Yanagizawa-Drott. Misinformation during a pandemic. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2020-44), 2020.

Ben Charoenwong, Alan Kwan, and Vesa Pursiainen. Social connections to covid-19-affected areas increase compliance with mobility restrictions. *Available at SSRN*, 2020.

Nicholas A. Christakis and James H. Fowler. The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4):370–379, 2007. doi: 10.1056/NEJMsa066082. URL https://doi.org/10.1056/NEJMsa066082. PMID: 17652652.

Nicholas A. Christakis and James H. Fowler. The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, 358(21):2249–2258, 2008. doi: 10.1056/NEJMsa0706154. URL https://doi.org/10.1056/NEJMsa0706154. PMID: 18499567.

Joshua Coven and Arpit Gupta. Disparities in mobility responses to covid-19. 2020.

Jonathan I Dingel and Brent Neiman. How many jobs can be done at home? Working Paper 26948, National Bureau of Economic Research, April 2020. URL http://www.nber.org/papers/w26948.

Maeve Duggan, Shannon Greenwood, and Andrew Perrin. Social media update 2016. Report, Pew Research Center, 2016. URL https://www.pewresearch.org/internet/2016/11/11/social-media-update-2016/.

Facebook. Facebook form 10-k, 2019 annual report, 2020. URL http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/45290cc0-656d-4a88-a2f3-147c8de86506.pdf.

Paola Giuliano and Imran Rasul. Civic capital and social distancing: Evidence from italians' response to covid-19, 2020a. URL https://voxeu.org/article/civic-capital-and-social-distancing.

Paola Giuliano and Imran Rasul. Compliance with social distancing during the covid-19 crisis, 2020b. URL https://voxeu.org/article/compliance-social-distancing-during-covid-19-crisis.

David Holtz, Michael Zhao, Seth G. Benzell, Cathy Y. Cao, Mohammad Amin Rahimian, Jeremy Yang, Jennifer Allen, Avinash Collis, Alex Moehring, Tara Sowrirajan, Dipayan Ghosh, Yunhao Zhang, Paramveer S. Dhillon, Christos Nicolaides, Dean Eckles, and Sinan Aral. Interdependence and the cost of uncoordinated responses to covid-19. *Proceedings of the National Academy of Sciences*, 117(33): 19837–19843, 2020. ISSN 0027-8424. doi: 10.1073/pnas.2009522117. URL https://www.pnas.org/content/117/33/19837.

Clayton J Hutto and Eric Gilbert. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*, 2014.

Theresa Kuchler, Lin Peng, Johannes Stroebel, Yan Li, and Dexin Zhou. Working Paper 27299, National Bureau of Economic Research, 2020a.

Theresa Kuchler, Dominic Russel, and Johannes Stroebel. The geographic spread of covid-19 correlates with structure of social networks as measured by facebook. Working Paper 26990, National Bureau of Economic Research, 2020b.

Paige Maas, Shankar Iyer, Andreas Gros, Wonhee Park, Laura McGorman, Chaya Nayak, and P Alex Dow. Facebook disaster maps: Aggregate insights for crisis response and recovery. *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, 2019. URL https://research.fb.com/wp-content/uploads/2019/04/iscram19_camera_ready.pdf.

Christos Makridis and Tao Wang. Learning from friends in a pandemic: Social networks and the macroeconomic response of consumption. *Available at SSRN 3601500*, 2020. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3601500.

Oliver Rehbein, Simon Rother, et al. Distance in bank lending: The role of social networks. Technical report, University of Bonn and University of Mannheim, Germany, 2020.

Ryoko Sato and Yoshito Takasaki. Peer effects on vaccination behavior: Experimental evidence from rural nigeria. *Economic Development and Cultural Change*, 68(1):93–129, 2019. doi: 10.1086/700570. URL https://doi.org/10.1086/700570.

Joe Schwartz. Bing maps tile system, 2018. URL https://docs.microsoft.com/en-us/bingmaps/articles/bing-maps-tile-system.

Andrey Simonov, Szymon K Sacher, Jean-Pierre H Dubé, and Shirsho Biswas. The persuasive effect of fox news: Non-compliance with social distancing during the covid-19 pandemic. Working Paper 27237, National Bureau of Economic Research, May 2020. URL http://www.nber.org/papers/w27237.

Yuan Tian, Maria Esther Caballero, and Brian K. Kovak. Social learning along international migrant networks. Working paper, July 2020. URL http://www.contrib.andrew.cmu.edu/~bkovak/tian_caballero_kovak.pdf.

Riley Wilson. The impact of social networks on eitc claiming behavior. 2019.

Appendices

A Replication of Friend-Exposure Effects Using Public Data

We use public data sources on movement and social connections to replicate and extend our previous findings. By doing so, we assess the robustness of our key results using other, well-studied data sources. This helps us benchmark the magnitude of our estimated effects against a broader literature exploring the factors shaping social distancing behavior. As is done in Section 5.3, we can also use these data sources to provide evidence regarding the mechanisms of the observed effects of friend-exposure to COVID-19.

A.1 Data

A.1.1 Safegraph

In light of the outbreak of the COVID-19 pandemic, Safegraph Inc. has released several data products that allow for a detailed understanding of consumer spending and of mobility patterns across time and space. These data are available to researchers on request. We use three different data products from Safegraph: (a) social distancing data, (b) point of interest (POI) visit data and (c) transaction data provided by Facteus. All three of these data sources have been used by other contemporaneous research on the COVID-19 pandemic. For the purpose of this Section, we focus on social distancing data. In Section 5, we exploit data sources (b) and (c) to shed light on the mechanisms of friend-exposure effects on social distancing behavior.

The Safegraph Social Distancing data contains location data obtained from a number of smartphone applications. Safegraph uses each user's location history to impute their Census block group of residence, and provides aggregated data for each block group from January 1, 2020. We use data through July 28. In particular, we use three metrics: the number of devices that are assigned to a given Census block group on a given day, the number of devices that do not leave their home location during a given day[24] as well as the average distance traveled.[25] The average number of devices observed on a given day in our sample period is about 19 million. Using these metrics, we calculate (a) the fraction of devices that remain at home over the course of a day and (b) the average distance traveled in kilometers. We believe that these two ZCTA-level measures of social distancing correspond nicely to the measures employed in the individual level analysis, i.e. the probability of staying at home and average daily tile movement. As before, for ease of computation and to smooth out fluctuations, we construct weekly averages.

A.1.2 Friend-Exposure to COVID-19

To construct a measure of friend-exposure to COVID-19 we combine data from Facebook on social connectedness and data from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. The Social Connectedness Index (SCI) (Bailey et al., 2018b) is a scaled metric of relative

[24] In this case, home location corresponds to the geohash-7 in which home is located. A geohash-7 is a region about 500 feet on each side.

[25] We construct the average distance traveled based on the number of devices per bin of travel distance. Where possible, we use the mean of highest and lowest value of the bin. For the open ended top bin ($> 50km$) we assign a value of $75km$.

connectedness of different ZCTAs across the U.S., defined as:

$$SCI_{ij} = \frac{FBConnections_{ij}}{FBUsers_i \times FBUsers_j} \tag{13}$$

$FBConnections_{ij}$ is the scaled number of connections between ZCTA i and ZCTA j, and $FBUsers_i$ and $FBUsers_j$ are the respective numbers of users for ZCTA i and j. To create our measure of friend-exposure, we begin by calculating per-user connections between ZCTA i and county k:

$$PerUserConnect_{ik} = \sum_{j \in k} SCI_{ij} * Pop_j \tag{14}$$

Pop_j is the population of ZCTA j that is in county k. Note that in the absence of public data on user counts, we use population counts rather than user counts. In constructing this measure we have two objectives. First, since the data on COVID-19 cases is only available at the county level this measure moves us from zips to counties. Second, this measure of per-user connections helps to construct friend-exposure that is independent of the number of users or friends on Facebook (which might systematically differ with the way Facebook is used across regions). Next, for each ZCTA i we calculate the fraction of per-user connections from county k relative to all counties:

$$FracConnect_{ik} = \frac{PerUserConnect_{ik}}{\sum_{k \in K} PerUserConnect_{ik}} \tag{15}$$

We can loosely think of this measure as the fraction of all friends a user in ZCTA i has in county k. As a final step, we multiply this metric with the number of COVID-19 cases in county k and sum over all counties in order to create our measure of friend-exposure to COVID-19. Since the number of cases varies over time, this metric is also time-variant.

$$FriendExpCOVID_{it} = \sum_{k \in K} FracConnect_{ik} \times Cases_{kt} \tag{16}$$

Here, t denotes the time-dimension which in our case is weeks. Together, these simplifications allow us to generate a measure of friend-exposure similar to the one used for our prior individual-level analyses.

A.2 Replication of Individual Level Results

To validate the findings presented in Section 3, we now estimate the effect of having high exposure at the zip level on social distancing behavior at the zip level. More concretely, we estimate:

$$Y_{it} = \beta_0 + \beta_{1t} \times HighExp_i \times week_t + \beta_{2t} \times X_{it} + \epsilon_{it} \tag{17}$$

Y_{it} is our measure of social distancing for ZCTA i during week t t constructed from Safegraph data, i.e. either (a) the average fraction of devices at home full-time for a given ZCTA or (b) the percentage change in the average distance traveled relative to the month of January 2020.[26] $HighExp_i$ is an indica-

[26]More precisely, based on our measure of average distance traveled for ZCTA i during week t, i.e. $AvgDist_{it}$, we calculate $\%\Delta Dist_{it} = \frac{AvgDist_{it} - AvgDist_{iJan20}}{AvgDist_{iJan20}} * 100$.

tor equal to one if ZCTA i has friend-exposure to COVID-19 higher than the median for the county it is located in, based on the number of COVID-19 cases as of March 17. We include a rich set of controls: in addition to including county-time fixed effects together with ZCTA fixed effects in our regressions, we control for various zip-level covariates interacted with time fixed effects. These are the median household income of the area, as well as the fraction of individuals in in each of the following demographic groups: male, Asian, black, white, service employee, manager, art or science employee, high-speed internet user, high-school educated, some college completion, college educated. We also control for the fraction of individuals in several age buckets.[27] All these control variables are obtained from the most recent 5-year ACS (2014-2018). Finally, as described in depth in Section 3, we control for national ventiles of friend-exposure to other factors, i.e. median household income, population density and urbanity.[28] In Table A11, we show the differences between high and low friend-exposure places with respect to these characteristics.

[Table A11]

While high and low friend-exposure places appear balanced with respect to many key demographic and socio-economic characteristics, a few differences are noticeable. In particular, high exposure places are slightly more racially diverse, have a somewhat lower median household income, and include individuals more likely to have a college degree. High exposure places also have larger populations, are more densely populated, and have more POIs. While none of these differences is very large, they might affect the the average ability or willingness of residents to engage in social distancing in a way that is independent of friend-exposure. We therefore control for all the above-mentioned set of covariates and allow for the value of these controls to vary over time. Together, these control variables help to alleviate concerns that any observed effects are merely driven by differences in demographic, socio-economic or other work-related variables that are correlated with social distancing behavior. Figure A11 depicts the corresponding β_{1t} estimates from Equation 17. These coefficients capture the effect of having a level of (ZCTA-level) friend-exposure to COVID-19 that is above the county mean. Standard errors are clustered at the ZCTA-level.

[Figure A11]

Figure A11 shows changes in mobility as a result of friend-exposure to COVID-19 that are qualitatively very consistent with the results presented in Section 3. As is in apparent both in Figure A11a and in Figure A11b, in January and February—before the outbreak of the pandemic in the U.S.—changes in mobility between high and low-exposure places are always very close to zero. Beginning in the week of March 4, these coefficients begin to shift, indicating that groups with more friend-exposure have begun to stay home more and travel less. For the fraction of devices at home, coefficients continue to rise, reaching levels of around 0.025 in late-March and early-April. Thereafter, coefficients slowly return to values closer to zero, yet they remain statistically significant for several more weeks and up until some-

[27] These are the fraction of individuals 18 or younger, between 18 and 24, between 25-34, between 35-44, between 45-54, between 55-64, between 65-74 and above 75.

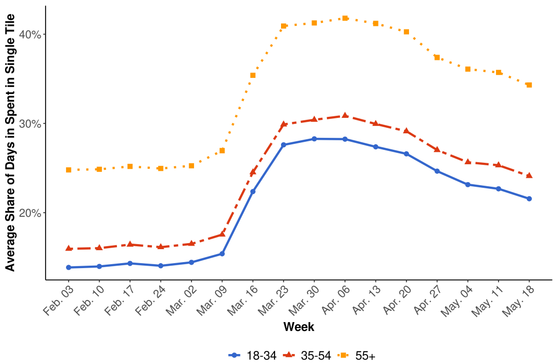
[28] These friend-exposure variables are constructed as $FriendExpMetric_i = \sum_{k \in K} FracConnect_{ik} \times Metric_k$ where $Metric_k$ is one of population density, median household income (both from ACS 2014-2018) and the fraction of the population residing in urban settings (from 2010 Census).

time in mid-May. In line with these patterns, for the percentage change in the average distance traveled, coefficients continue to fall during late March and stay low, i.e. around -1.5, for much of April before they gradually return to zero, although never quite reaching zero. Together, these estimates highlight that as the COVID-19 pandemic hits the U.S., places with greater friend-exposure to COVID-19 reduce their mobility more than places with lower friend-exposure. These effects are persistent over time and cannot entirely be explained by our measures of differential ability and/or willingness to engage in social distancing. In spite of the different data source, the different level of analysis and the different sample, these results are thus consistent with the evidence presented in Section 3: friend-exposure to COVID-19 matters when trying to explain differences in social distancing behavior across individuals and across places.

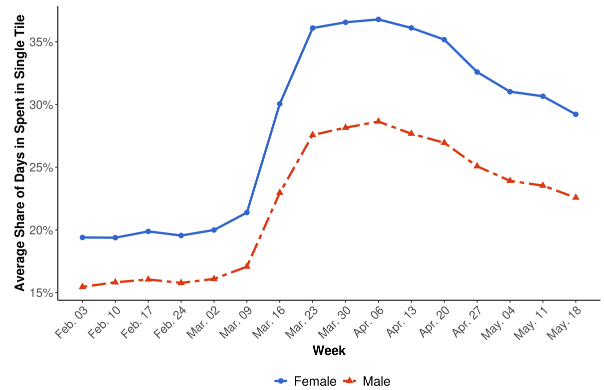
B Additional Tables and Figures

Figure A1: Heterogeneity in Probability of Staying at Home

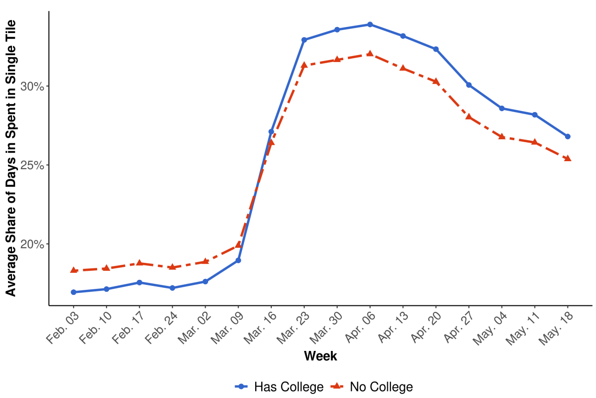
(a) Age



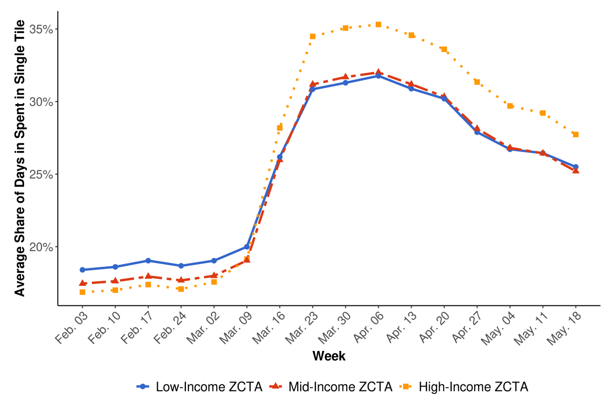
(b) Gender



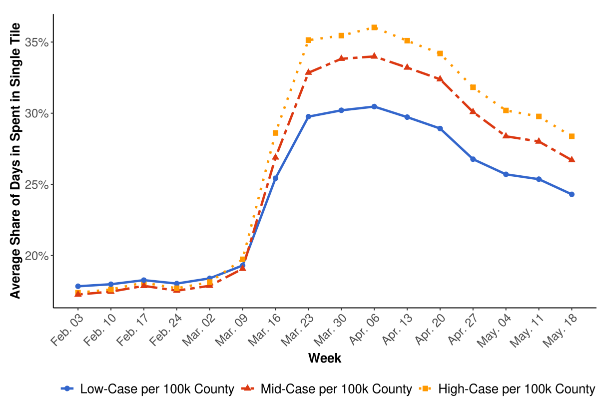
(c) Educational Attainment



(d) Income



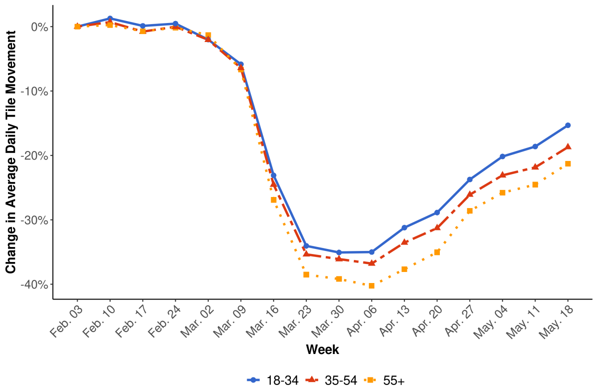
(e) Local Exposure to COVID-19



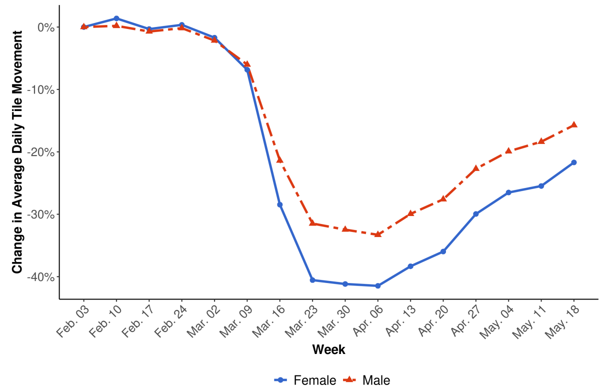
Note: Figures show weekly averages of the probability of staying at home from the week of February 3rd to the week of May 18th across certain characteristics. Panel (a) shows age; panel (b) shows gender; panel (c) shows whether the user has a college listed on Facebook; panel (d) shows the tertile of home ZCTA median household income; and panel (e) shows the tertile of county-level cases per resident as of March 15th.

Figure A2: Heterogeneity in Change in Average Tiles Visited

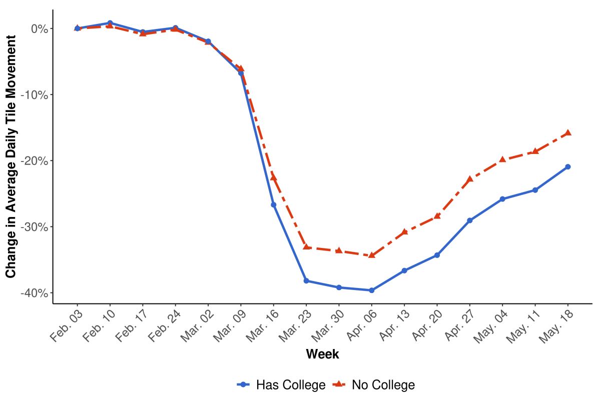
(a) Age



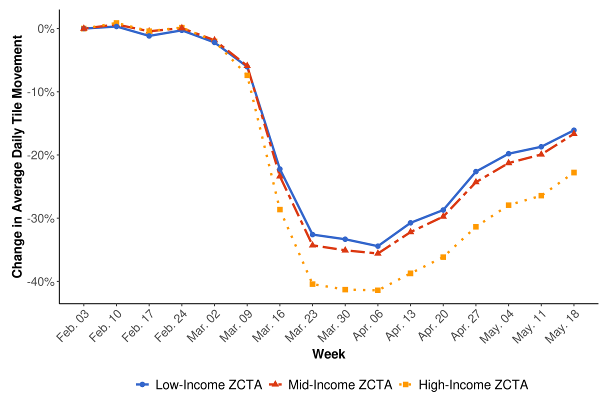
(b) Gender



(c) Educational Attainment



(d) Income



(e) Local Exposure to COVID-19

Note: Figures show the percent change in the weekly average of daily tiles visited from the week of February 3rd to the week of May 18th across certain characteristics. Panel (a) shows age; panel (b) shows gender; panel (c) shows whether the individual has college information in Facebook; panel (d) shows the tercile of ZCTA-level median household income; and panel (e) shows the tercile of county-level cases per resident as of March 15th.

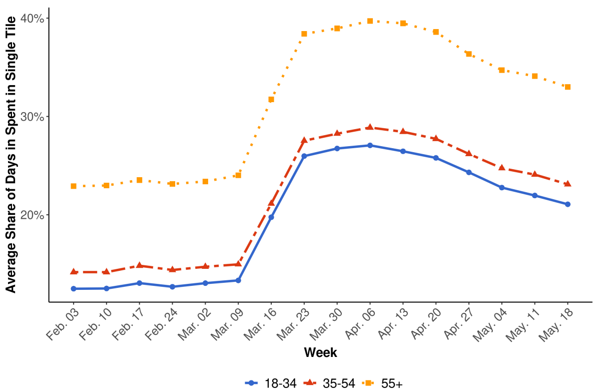
Table A1: Change in Average Tiles Visited

	Bing Tile Visited					
	All		Weekdays		Weekends	
	Level Feb	ΔFeb-Apr	Level Feb	ΔFeb-Apr	Level Feb	ΔFeb-Apr
Overall	10.957	-3.590	11.339	-3.632	10.570	-3.714
By Age Group						
18-34	11.590	-3.593	11.883	-3.587	11.555	-3.843
35-54	11.507	-3.753	11.952	-3.818	10.975	-3.834
55+	9.287	-3.307	9.656	-3.358	8.804	-3.381
By Gender						
Female	9.729	-3.641	9.937	-3.697	9.694	-3.711
Male	12.398	-3.530	12.985	-3.555	11.602	-3.717
By College						
Has College	11.041	-3.945	11.395	-4.012	10.714	-4.014
No College	10.862	-3.179	11.275	-3.193	10.405	-3.362
By Zip Code Income						
Bottom Tertile	10.735	-3.146	11.110	-3.147	10.392	-3.372
Middle Tertile	10.899	-3.367	11.265	-3.386	10.530	-3.525
Top Tertile	11.238	-4.247	11.642	-4.353	10.787	-4.228
By County Total Cases/Population						
Bottom Tertile	10.670	-2.916	11.006	-2.883	10.358	-3.186
Middle Tertile	11.317	-4.066	11.713	-4.129	10.939	-4.174
Top Tertile	11.140	-4.246	11.579	-4.382	10.643	-4.174
By Exposure through Friends						
High Exposure	10.959	-3.849	11.333	-3.900	10.599	-3.968
Low Exposure	10.956	-3.331	11.345	-3.365	10.542	-3.460

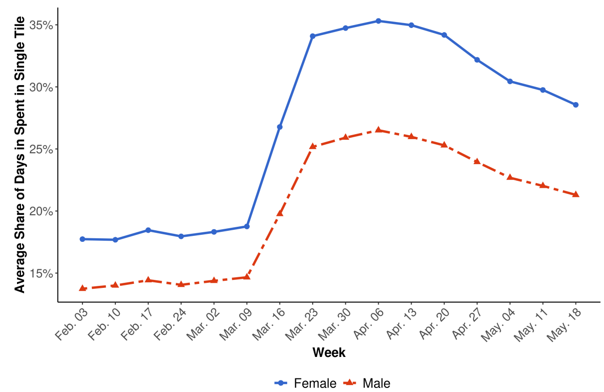
Note: Table describes changes in social distancing across different user characteristics. Social distancing is measured as the average number of daily Bing tiles visited. Characteristic splits include age group, gender, whether the individual has college information in Facebook, the tercile of ZCTA-level median household income, the tercile of county-level cases per resident as of March 15th, and whether the log of friend-exposure to COVID cases on March 15th is above (high exposure) or below (low exposure) the users' home ZCTA median. Columns 1, 3, and 5 show the levels for the week of February 25th to March 2nd (prior to the pandemic). Columns 2, 4, 6 show the difference between the week of April 14th to 20th (during the early stages of the pandemic) and this baseline. Columns 1 and 2 include all days; 3 and 4 include weekdays only; and 5 and 6 include weekends only.

Figure A3: Weekdays, Heterogeneity in the Probability of Staying at Home

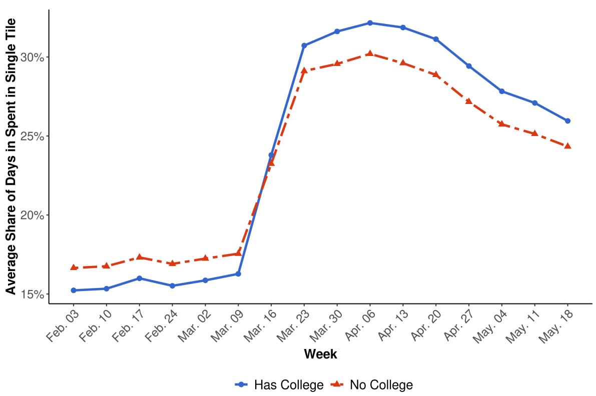
(a) Age



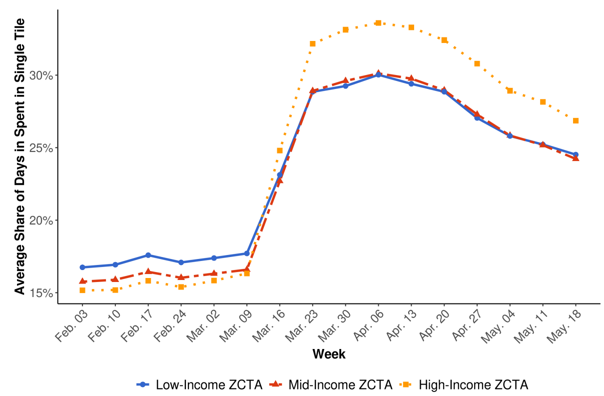
(b) Gender



(c) Educational Attainment



(d) Income

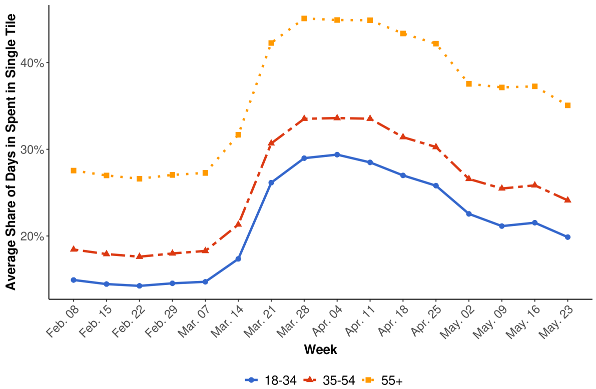


(e) Local Exposure to COVID-19

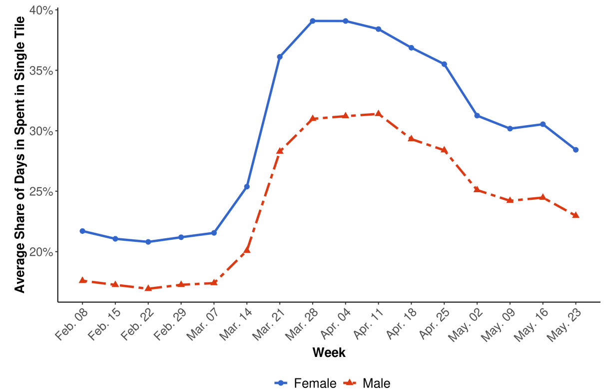
Note: Figures show weekly averages, *of weekdays*, of the probability of staying at home from the week of February 3rd to the week of May 18th across certain characteristics. Panel (a) shows age; panel (b) shows gender; panel (c) shows whether the individual has college information in Facebook; panel (d) shows the tercile of ZCTA-level median household income; and panel (e) shows the tercile of county-level cases per resident as of March 15th.

Figure A4: Weekends, Heterogeneity in the Probability of Staying at Home

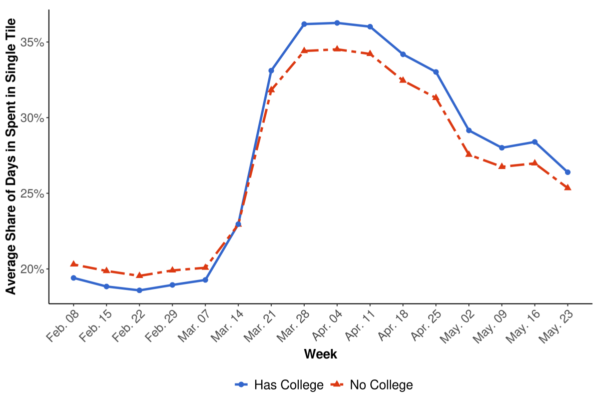
(a) Age



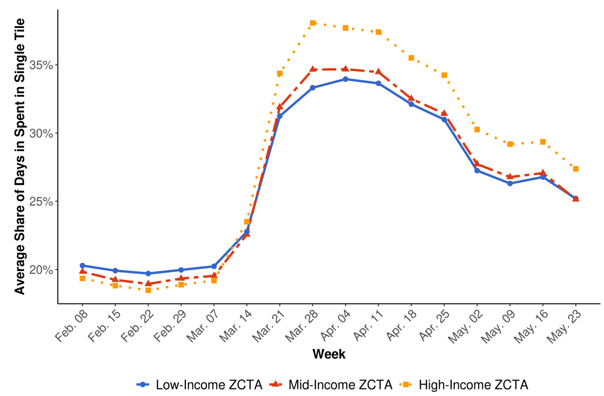
(b) Gender



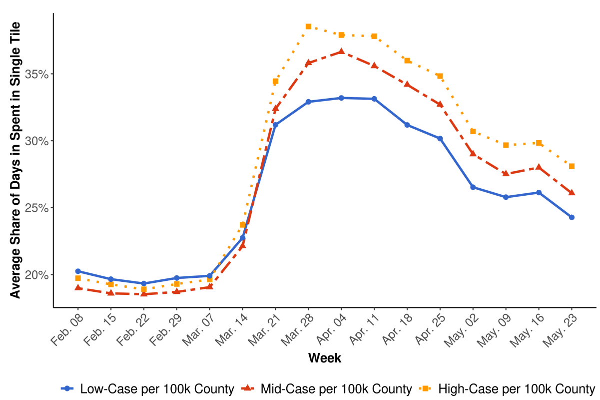
(c) Educational Attainment



(d) Income



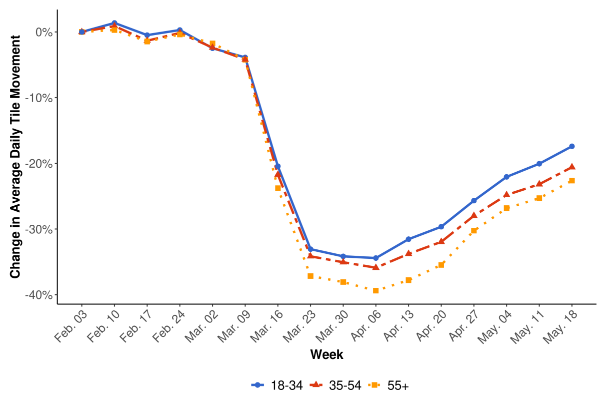
(e) Local Exposure to COVID-19



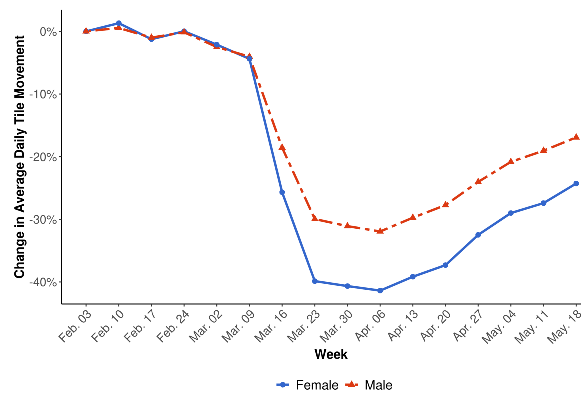
Note: Figures show weekly averages, *of weekend days*, of the probability of staying at home from the weekend of February 8th to the weekend of May 23rd across certain characteristics. Panel (a) shows age; panel (b) shows gender; panel (c) shows whether the individual has college information in Facebook; panel (d) shows the tercile of ZCTA-level median household income; and panel (e) shows the tercile of county-level cases per resident as of March 15th.

Figure A5: Weekdays, Heterogeneity in the Average Number of Tiles Visited

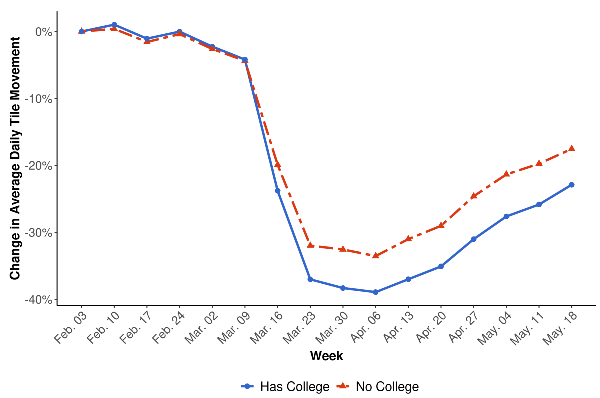
(a) Age



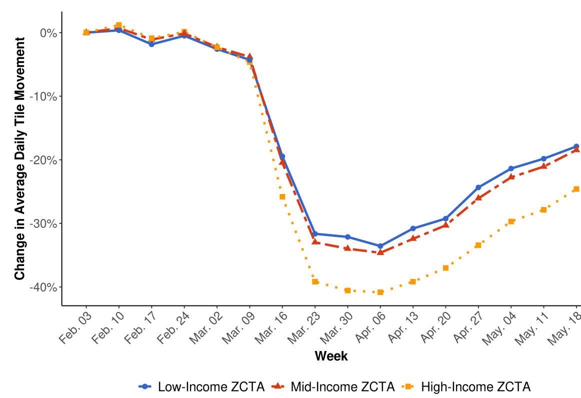
(b) Gender



(c) Educational Attainment



(d) Income

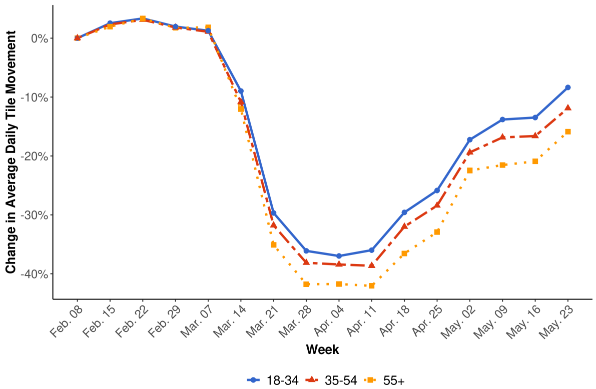


(e) Local Exposure to COVID-19

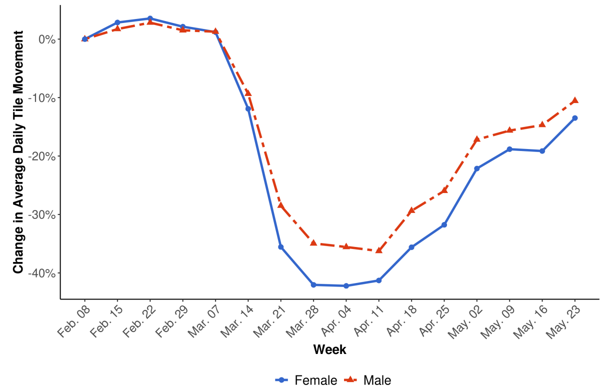
Note: Figures show the percent change in the weekly average, *of weekdays*, of daily tiles visited from the week of February 3rd to the week of May 18th across certain characteristics. Panel (a) shows age; panel (b) shows gender; panel (c) shows whether the individual has college information in Facebook; panel (d) shows the tercile of ZCTA-level median household income; and panel (e) shows the tercile of county-level cases per resident as of March 15th.

Figure A6: Weekends, Heterogeneity in the Average Number of Tiles Visited

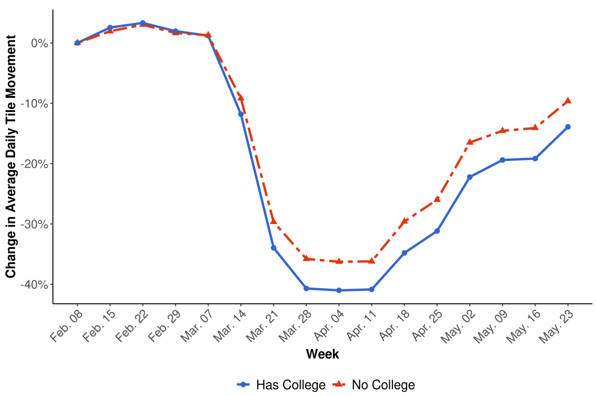
(a) Age



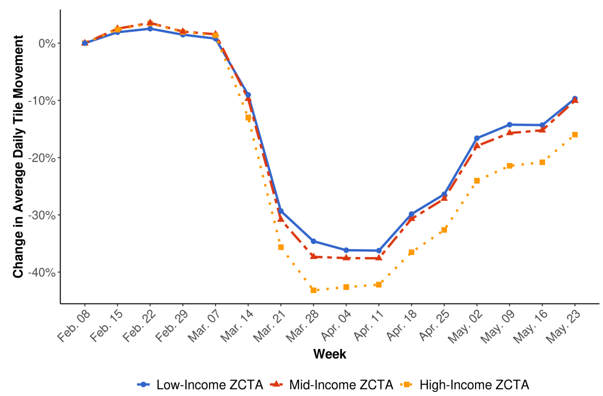
(b) Gender



(c) Educational Attainment



(d) Income



(e) Local Exposure to COVID-19

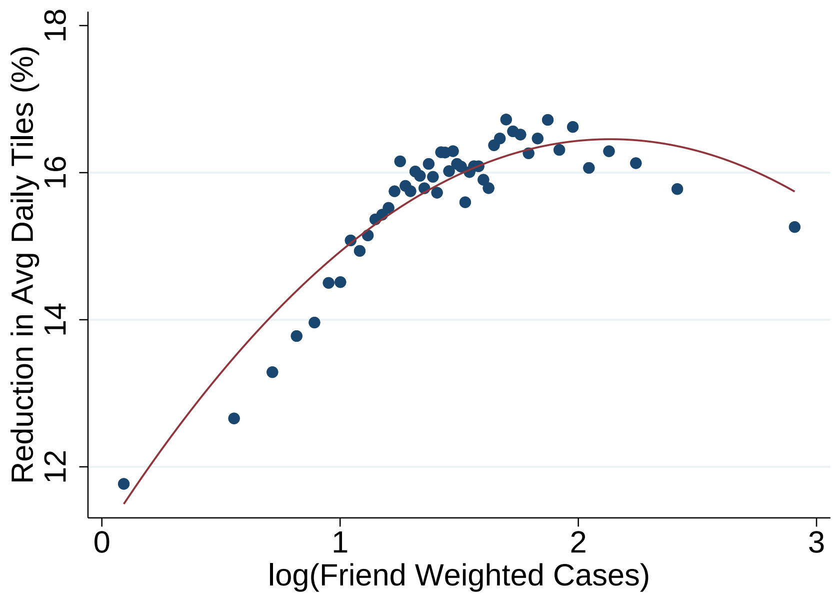
Note: Figures show the percent change in the weekly average, *of weekend days*, of daily tiles visited from the weekend of February 8th to the weekend of May 23rd across certain characteristics. Panel (a) shows age; panel (b) shows gender; panel (c) shows whether the individual has college information in Facebook; panel (d) shows the tercile of ZCTA-level median household income; and panel (e) shows the tercile of county-level cases per resident as of March 15th.

Table A2: Social Distancing by Demographics: Percent Reduction in Number of Tiles Visited

	DV: % Reduction - Bing Tiles Visited (Feb - Apr)							
Age Group								
35-54	1.073***	0.986***		1.012***				
	(0.104)	(0.101)		(0.101)				
55+	3.534***	3.702***		3.842***				
	(0.119)	(0.112)		(0.112)				
Female	9.577***	10.036***		10.285***				
	(0.084)	(0.082)		(0.082)				
Has College	7.347***	6.825***		6.233***				
	(0.085)	(0.081)		(0.081)				
Has iPhone	5.847***	4.934***		4.635***				
	(0.099)	(0.098)		(0.098)				
Has Tablet	0.141*	0.041		-0.057				
	(0.079)	(0.078)		(0.078)				
Zip Code Income								
Middle Tertile	3.467***							
	(0.229)							
Top Tertile	9.432***							
	(0.226)							
County Cases/Pop								
Middle tertile	8.387***							
	(0.204)							
Top Tertile	9.892***							
	(0.227)							
log(Friend Exposure)			1.802***	1.585***	1.514***	1.455***	1.481***	1.473***
			(0.083)	(0.083)	(0.092)	(0.155)	(0.103)	(0.144)
Zip Code FE		Y	Y	Y				
Other Network Exposure FE			Y	Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone					Y	Y	Y	Y
College FE								Y
Sample						Weekend	Weekday	College
R-Squared	0.009	0.018	0.015	0.020	0.154	0.155	0.156	0.172
Sample Mean	15.640	15.640	15.641	15.641	15.801	-1.943	12.668	20.942
N	6,804,168	6,804,167	6,803,761	6,803,761	6,400,738	5,808,187	6,309,820	2,616,959

Note: Table shows results from regression 4. Each observation is an individual. The outcome in all columns is the percent reduction in average number of Bing tiles visited from the week of February 25th to March 2nd (prior to the pandemic) to April 14th to 20th. Column 1 includes controls for age groups, gender, whether the individual has college information in Facebook, whether the individual primarily accesses mobile Facebook from an iPhone, whether the individual has accessed Facebook from a tablet, the tercile of ZCTA-level median household income, and the tercile of county-level cases per resident as of March 15th. Column 2 adds ZCTA fixed effects, but maintains the individual level controls. Column 3 includes only the log of friend-exposure to COVID cases on March 15th; ZCTA fixed effects; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Column 4 adds back the individual-level controls from column 1. Column 5 adds fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. In Column 6 the outcome is weekend movement and in column 7 the outcome is weekday movement. Column 8 limits to individuals that attended a college, limiting to colleges with more than 100 individuals, and adds a fixed effect for each individual college. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Figure A7: Percent Reduction in Average Number of Tiles Visited vs. Friend-Exposure



Note: Figure shows a binned scatter plot of the log of friend weighted friend-exposure to COVID on March 15th and the percent reduction in average number of tiles visited from the week of February 25th to March 2nd (prior to the pandemic) to April 14th to 20th. The plot controls for fixed effects constructed from interacting ZCTA, age group, gender, has college information in Facebook, has iPhone, and has tablet. It also controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas.

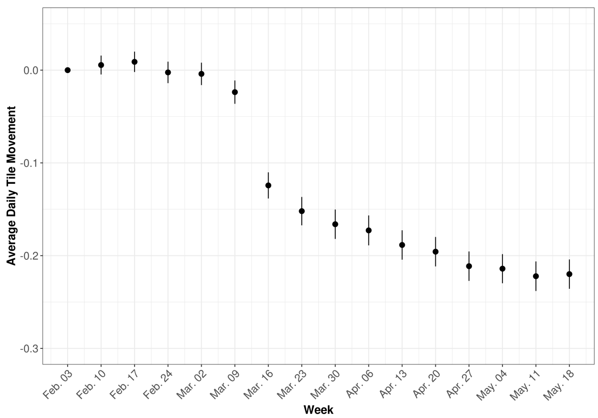
Table A3: Social Distancing and Other Exposure

	DV: Δ Stay at Home (Feb - Apr)							
log(Friend Exposure)	0.878*** (0.028)	0.521*** (0.043)		0.872*** (0.028)	0.875*** (0.028)	0.876*** (0.028)	0.872*** (0.028)	0.861*** (0.028)
log(Friend Exposure, Cases per 100k)			0.778*** (0.029)					
Share Friends China				1.116*** (0.090)				1.075*** (0.089)
Share Friends South Korea					0.215*** (0.022)			0.207*** (0.021)
Share Friends Italy						0.068*** (0.014)		0.053*** (0.014)
Share Friends Spain							0.209*** (0.022)	0.200*** (0.022)
Sample		Friends >100mi						
Other Network Exposure FE	Y	Y	Y	Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y	Y	Y	Y	Y
R-Squared	0.175	0.229	0.175	0.175	0.175	0.175	0.175	0.175
Sample Mean	13.800	14.876	13.800	13.800	13.800	13.800	13.800	13.800
N	6,400,738	2,479,352	6,400,738	6,400,738	6,400,738	6,400,738	6,400,738	6,400,738

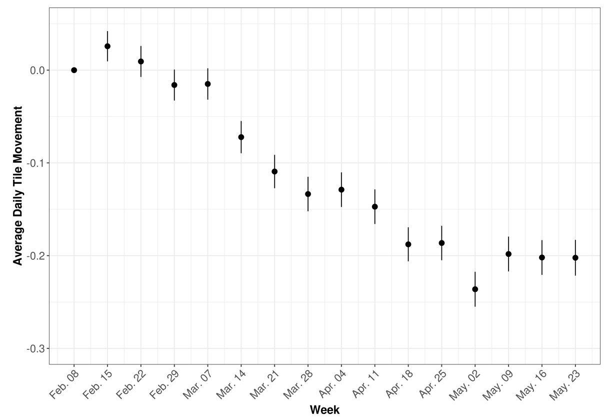
Note: Table shows results from regression 4, using alternative measures of friend-exposure to COVID-19. Each observation is an individual. The outcome in all columns is the percent reduction in average number of Bing tiles visited from the week of February 25th to March 2nd (prior to the pandemic) to April 14th to 20th. Column 1 is the same specification as column 5 of Table 4. Column 2 limits exposure to only friendships with individuals in counties more than 100 miles away. The sample size falls as we restrict to individuals with more than 100 such friends (as described in Section 2.1, we use a similar friend count including *all* friends in our primary sample). Column 3 uses cases per 100k residents (instead of cases) to calculate friend-exposure. Columns 4, 5, 6, and 7 add controls for the share of friends individuals have in China, South Korea, Italy, and Spain respectively. Column 8 adds all four of these country controls at once. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Figure A8: Robustness: Effects of Network-Exposure to COVID-19 on Average Daily Tiles Visited

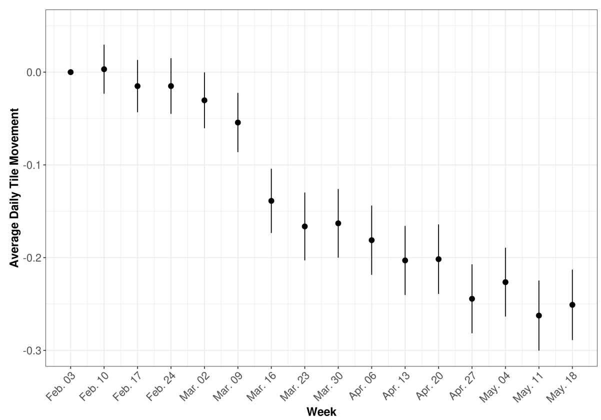
(a) Weekdays



(b) Weekends



(c) Controlling for Exact College



Note: Figures show coefficients estimated using the difference-in-differences setup specified in equation 3 with the outcome variable as the average number of Bing tiles visited. The outcome is measured on weekdays in panel (a) and weekends in panel (b). Panel (c) limits to individuals that attended college, limiting to colleges with more than 100 individuals, and adds a fixed effect for each individual college interacted with week. All specifications include fixed effects at the individual level as well as the following groups interacted with week: ZCTA; age group; gender; has college information in Facebook; has iPhone; has tablet; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Standard errors are clustered by ZCTA.

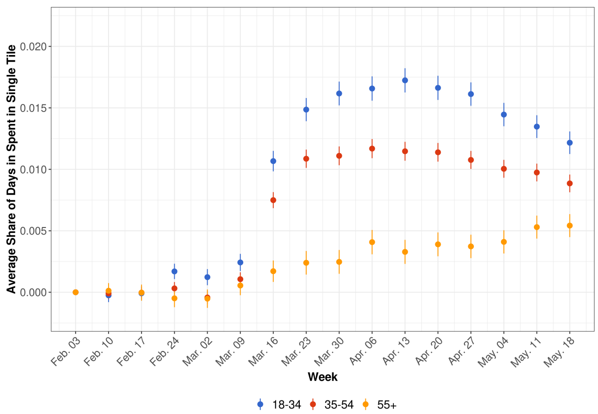
Table A4: Effects of Friend-Exposure by Month: Δ Probability of Staying at Home

	Monthly Change in Δ Stay at Home				
	March	April	May	June	July
Δ log(Friend Exposure + 1), March	0.207***				
	(0.046)				
Δ log(Friend Exposure + 1), April		0.032			
		(0.048)			
Δ log(Friend Exposure + 1), May			0.460***		
			(0.073)		
Δ log(Friend Exposure + 1), June				0.577***	
				(0.089)	
Δ log(Friend Exposure + 1), July					0.076
					(0.089)
Other Network Exposure FE	Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y	Y
R-Squared	0.174	0.141	0.150	0.146	0.145
Sample Mean	14.214	-0.923	-5.989	-1.068	0.679
N	6,688,448	6,579,359	6,169,176	5,848,722	5,456,303

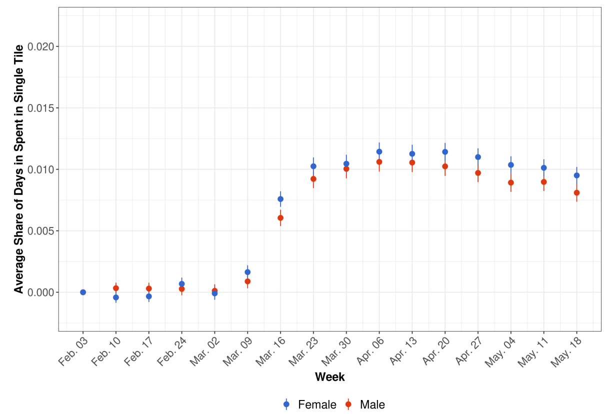
Note: Table shows results from regression 8. Each observation is an individual. The outcome variable is the change in the probability of staying home between the final weeks of a given month and the previous months' final week: February 25-March 2 for February; March 24-March 30 for March; April 21-April 27 for April; May 26-June 1; June 23-June 29; July 21-July 28. We consider changes by month. In all columns we control for interactions of age groups, gender, whether the individual has a college listed on Facebook, whether the individual primarily accesses mobile Facebook from an iPhone, and whether the individual has accessed Facebook from a tablet. We also control for fixed effects for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. In column 2, we include user fixed effects. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

Figure A9: Heterogeneity of Friend Effect: Probability of Staying at Home

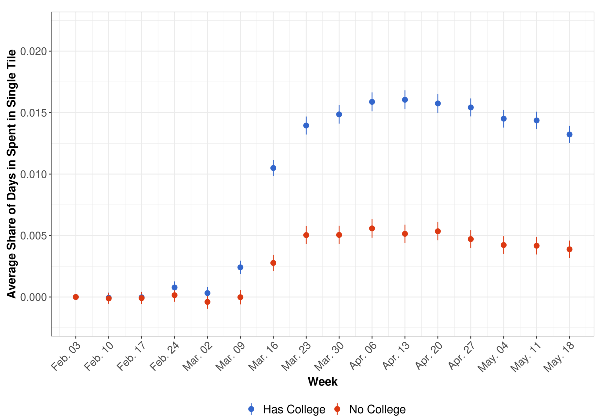
(a) Age



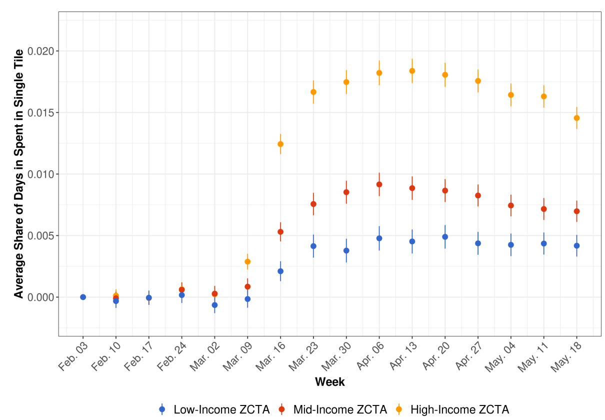
(b) Gender



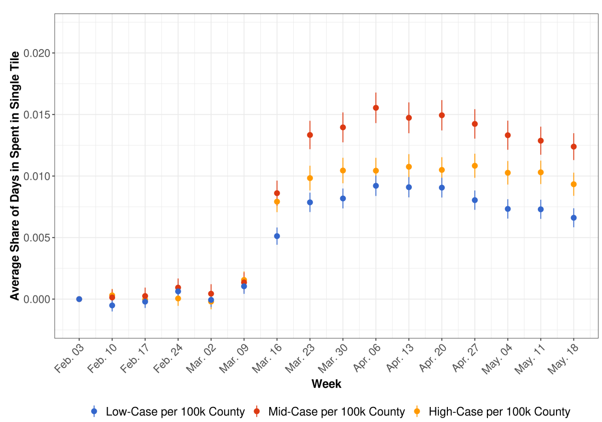
(c) Educational Attainment



(d) Income



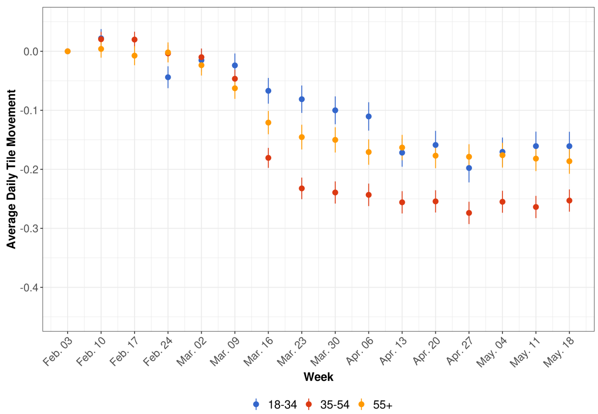
(e) Local Exposure to COVID-19



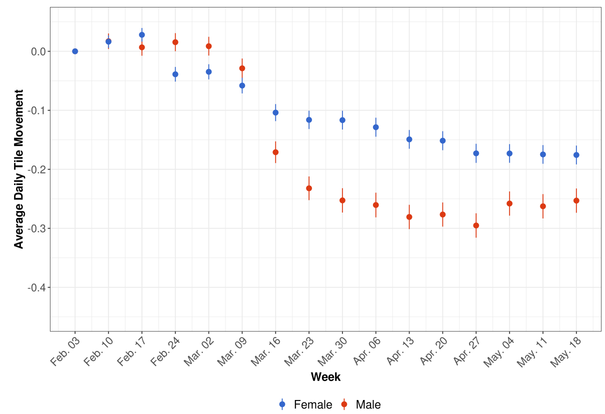
Note: Figures show coefficients estimated using the difference-in-differences setup described in Section 4 with the outcome variable as the probability of staying at home. The heterogeneities interacted with exposure are: age in panel (a), gender in panel (b), whether the individual has a college listed on Facebook in panel (c); the tertile of home ZCTA median household income in panel (d); and the tertile of home county cases per resident as of March 15th in panel (e). All specifications include fixed effects at the individual level as well as the following groups interacted with week: ZCTA; age group; gender; has college; has iPhone; has tablet; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Standard errors are clustered by ZCTA.

Figure A10: Heterogeneity of Friend Effect: Average Daily Tiles Visited

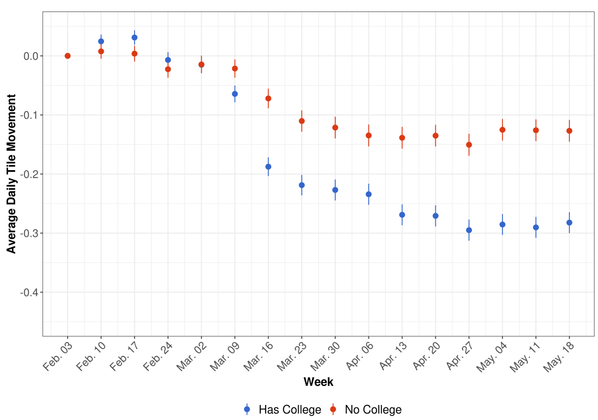
(a) Age



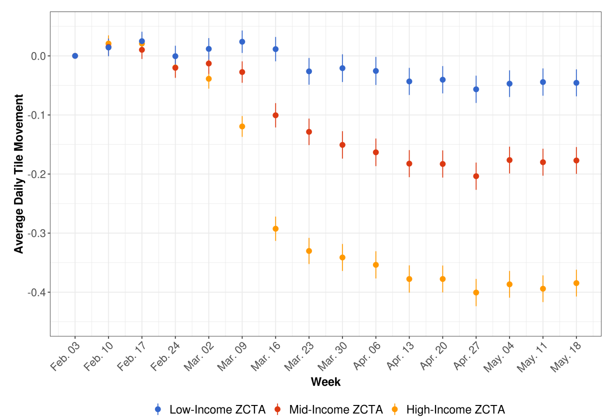
(b) Gender



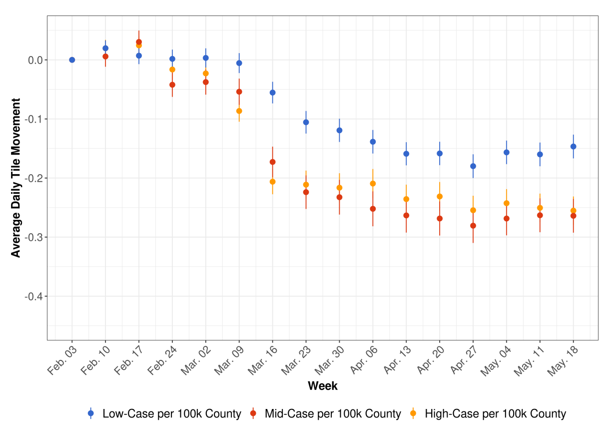
(c) Educational Attainment



(d) Income



(e) Local Exposure to COVID-19



Note: Figures show coefficients estimated using the difference-in-differences described in Section 4 with the outcome variable as the average daily tiles visited. The heterogeneities interacted with exposure are: age in panel (a), gender in panel (b), whether the individual has a college listed on Facebook in panel (c); the tertile of home ZCTA median household income in panel (d); and the tertile of home county cases per resident as of March 15th in panel (e). All specifications include fixed effects at the individual level as well as the following groups interacted with week: ZCTA; age group; gender; has college; has iPhone; has tablet; and percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. Standard errors are clustered by ZCTA.

Table A5: Heterogeneity of Friend-Exposure Effects: Average Daily Tiles Visited

	%Δ Bing Tiles Visited					
log(Friend Exposure) x I(Age < 35)	1.942***					
	(0.146)					
log(Friend Exposure) x I(Age 35-55)	1.860***					
	(0.114)					
log(Friend Exposure) x I(Age > 55)	0.535***					
	(0.123)					
log(Friend Exposure) x Female		1.125***				
		(0.100)				
log(Friend Exposure) x Male		1.971***				
		(0.123)				
log(Friend Exposure) x College			2.030***			
			(0.107)			
log(Friend Exposure) x No College			1.006***			
			(0.114)			
log(Friend Exposure) x Zip Income First Tertile				0.576***		
				(0.136)		
log(Friend Exposure) x Zip Income Second Tertile				1.289***		
				(0.122)		
log(Friend Exposure) x Zip Income Third Tertile				2.990***		
				(0.135)		
log(Friend Exposure) x County Cases First Tertile					0.926***	
					(0.104)	
log(Friend Exposure) x County Cases Second Tertile					2.429***	
					(0.183)	
log(Friend Exposure) x County Cases Third Tertile					3.087***	
					(0.168)	
log(Friend Exposure - Rank 1 - 25)						0.463***
						(0.058)
log(Friend Exposure - Rank 26 - 50)						0.097
						(0.060)
log(Friend Exposure - Rank 51 - 75)						-0.062
						(0.059)
log(Friend Exposure - Rank 76 - 100)						0.139**
						(0.059)
Other Network Exposure FE	Y	Y	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y	Y	Y
R-Squared	0.154	0.154	0.154	0.154	0.154	0.156
Sample Mean	15.801	15.801	15.801	15.801	15.801	17.436
F Test (Rank 1-25 = Rank 76-100)						13.393***
N	6,400,738	6,400,738	6,400,738	6,400,738	6,400,738	5,684,469

Note: Table shows results from regressions of friend-exposure to COVID-19 on March 15th, interacted with individual characteristics, on the percentage change in average tile movement. Each observation is an individual. Friend-exposure is interacted with age groups in rows 1-3; gender in rows 4-5; whether the individual has a college listed in Facebook in rows 6-7; zip-level median household income in rows 8-10; county-level cases of COVID-19 in rows 11-13; and friend rank (i.e. a measure for how close friends are) in rows 14-16. All columns include controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. All columns include fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01). [Return to text]

Table A6: Summary Characteristics - Posting Behavior Sample

	Mean	SD	P10	P25	P50	P75	P90
Age	42.40	15.96	24	29	40	53	64
Female	0.58	0.49	0	0	1	1	1
Has College	0.60	0.49	0	0	1	1	1
Zip Code Income	$61,284	$23,993	$36,729	$44,902	$55,662	$72,704	$94,000
Has iPhone	0.59	0.49	0	0	1	1	1
Has Tablet	0.47	0.50	0	0	0	1	1
Number of Friends	564.85	341.16	196	289	477	776	1103
Friend Exposure to Cases	10.31	19.68	0.78	1.84	4.55	10.83	25.16
Number of Posts Feb	16.12	64.85	0	0	1	8	34
Average Sentiment (Feb)	31.89	35.26	-3.41	3.50	29.91	58.00	83.00
Number of Posts April	20.83	74.95	0	0	2	13	47
Average Sentiment (April)	29.94	34.21	-4.75	3.86	27.80	53.84	79.47
Number Posts about Corona	0.724	4.687	0	0	0	0	2
Average Sentiment Corona Posts	21.46	52.79	-52.75	-10.13	21.09	66.71	93.37
Number Posts Support Lockdown	0.013	0.238	0	0	0	0	0
Number Posts Oppose Lockdown	0.008	0.118	0	0	0	0	0

Note: Table presents summary statistics describing users in our sample underlying the analysis of public posts. Individual-level characteristics include age, gender, whether the user has a college listed on Facebook, whether the user primarily accesses Facebook mobile from an iPhone, whether the individual has accessed Facebook from a tablet, number of friends, friend-exposure to COVID cases on March 15th, and patterns of mobility during the week of February 25th to March 2nd. The table also includes information on the users' home ZCTA 2018 median household income.

Table A7: Summary Characteristics - Group Membership Sample

	Mean	SD	P10	P25	P50	P75	P90
Age	41.97	16.01	24	29	39	53	64
Female	0.57	0.50	0	0	1	1	1
Has College	0.59	0.49	0	0	1	1	1
Zip Code Income	$63,798	$26,081	$36,954	$45,848	$57,600	$76,544	$99,328
Has iPhone	0.61	0.49	0	0	1	1	1
Has Tablet	0.43	0.50	0	0	0	1	1
Number of Friends	502.52	319.56	177	252	410	676	1003
Friend Exposure to Cases	12.42	22.17	0.91	2.23	5.64	13.77	31.75
Number Groups (Feb)	33.03	57.89	3	8	18	38	73
Has Any Groups (Feb)	0.98	0.13	1	1	1	1	1
Number Anti-Lockdown Groups (April)	0.014	0.133	0	0	0	0	0
Has Anti-Lockdown Group (April)	0.012	0.110	0	0	0	0	0

Note: Table presents summary statistics describing users in our sample underlying the analysis of group memberships. Individual-level characteristics include age, gender, whether the user has a college listed on Facebook, whether the user primarily accesses Facebook mobile from an iPhone, whether the individual has accessed Facebook from a tablet, number of friends, friend-exposure to COVID cases on March 15th, and patterns of mobility during the week of February 25th to March 2nd. The table also includes information on the users' home ZCTA 2018 median household income.

Table A8: Heterogeneity of Friend-Exposure Effects - Own Age / Gender / College

	% Posts about Corona	% Corona-Posts Opp. Distancing	Δ Sentiment All Posts	Member "Reopen Group"
log(Friend Exposure) x I(Age < 35)	0.209***	-1.650***	-0.075**	-0.034***
	(0.007)	(0.416)	(0.033)	(0.006)
log(Friend Exposure) x I(Age 35-55)	0.307***	-2.185***	-0.081**	-0.210***
	(0.007)	(0.287)	(0.033)	(0.009)
log(Friend Exposure) x I(Age > 55)	0.213***	-1.572***	-0.143***	-0.127***
	(0.006)	(0.384)	(0.039)	(0.007)
Other Network Exposure FE	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y
R-Squared	0.060	0.445	0.118	0.074
Sample Mean	1.755	35.979	-1.823	1.216
N	34,528,373	277,776	10,777,790	119,145,833
	% Posts about Corona	% Corona-Posts Opp. Distancing	Δ Sentiment All Posts	Member "Reopen Group"
log(Friend Exposure) x Female	0.197***	-1.536***	-0.174***	-0.060***
	(0.006)	(0.262)	(0.028)	(0.006)
log(Friend Exposure) x Male	0.319***	-3.074***	0.034	-0.216***
	(0.007)	(0.388)	(0.034)	(0.008)
Other Network Exposure FE	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y
R-Squared	0.060	0.445	0.118	0.074
Sample Mean	1.755	35.979	-1.823	1.216
N	34,528,373	277,776	10,777,790	119,145,833
	% Posts about Corona	% Corona-Posts Opp. Distancing	Δ Sentiment All Posts	Member "Reopen Group"
log(Friend Exposure) x College	0.352***	-2.281***	-0.122***	-0.171***
	(0.007)	(0.258)	(0.030)	(0.007)
log(Friend Exposure) x No College	0.124***	-0.838**	-0.058*	-0.082***
	(0.005)	(0.399)	(0.031)	(0.000)
Other Network Exposure FE	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y
R-Squared	0.060	0.445	0.118	0.074
Sample Mean	1.755	35.979	-1.823	1.216
N	34,528,373	277,776	10,777,790	119,145,833

Note: Table shows results from regressions of friend-exposure to COVID-19 on March 15th, interacted with individual characteristics, on a number of outcomes. Each observation is an individual. Friend-exposure is interacted with age groups in rows 1-3; gender in rows 4-5; and whether the individual has a college listed in Facebook in rows 6-7. The outcomes in columns 1-2 are the change in probability of staying at home and the percent reduction in the average number of tiles visited, respectively, from the week of February 25 - March 2 (prior to the pandemic) to April 14 - 20. The outcome in column 3 is the percentage of individual posts that are about COVID-19. In column 4 it is the percentage of pro- or anti-distancing posts that are anti-distancing. In column 5 it is the change in the average sentiment of the posts from February 3 - 23 to April 6 - 26. In column 6 it is whether the individual, as of June 28, was a member of a 'Reopen' Facebook group. Post and group classifications are defined in Appendix C. All columns include controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. All columns include fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table A9: Heterogeneity of Friend-Exposure Effects - Own Income / Local Cases

	% Posts about Corona	% Corona-Posts Opp. Distancing	Δ Sentiment All Posts	Member "Reopen Group"
log(Friend Exposure) x Zip Income First Tertile	0.163***	-2.155***	-0.034	-0.080***
	(0.007)	(0.377)	(0.033)	(0.011)
log(Friend Exposure) x Zip Income Second Tertile	0.216***	-1.792***	-0.101***	-0.136***
	(0.007)	(0.335)	(0.034)	(0.012)
log(Friend Exposure) x Zip Income Third Tertile	0.404***	-1.884***	-0.172***	-0.185***
	(0.010)	(0.338)	(0.040)	(0.014)
Other Network Exposure FE	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y
R-Squared	0.060	0.445	0.118	0.074
Sample Mean	1.755	35.979	-1.823	1.216
N	34,528,373	277,776	10,777,790	119,145,833

	% Posts about Corona	% Corona-Posts Opp. Distancing	Δ Sentiment All Posts	Member "Reopen Group"
log(Friend Exposure) x County Cases First Tertile	0.190***	-1.904***	-0.086***	-0.065***
	(0.006)	(0.294)	(0.028)	(0.012)
log(Friend Exposure) x County Cases Second Tertile	0.392***	-2.084***	-0.047	-0.183***
	(0.013)	(0.422)	(0.050)	(0.012)
log(Friend Exposure) x County Cases Third Tertile	0.356***	-1.855***	-0.168***	-0.123***
	(0.012)	(0.399)	(0.046)	(0.011)
Other Network Exposure FE	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y
R-Squared	0.060	0.445	0.118	0.074
Sample Mean	1.755	35.979	-1.823	1.216
N	34,528,373	277,776	10,777,790	119,145,833

Note: Table shows results from regressions of friend-exposure to COVID-19 on March 15th interacted with various ZCTA-level characteristics on a number of outcomes. Each observation is an individual. Friend-exposure is interacted with tertiles of ZCTA median household income in rows 1-3; and tertiles of county cases per resident as of March 15th in rows 4-6. The outcomes in columns 1-2 are the change in probability of staying at home and the percent reduction in the average number of tiles visited, respectively, from the week of February 25 - March 2 (prior to the pandemic) to April 14 - 20. The outcome in column 3 is the percentage of individual posts that are about COVID-19. In column 4 it is the percentage of pro- or anti-distancing posts that are anti-distancing. In column 5 it is the change in the average sentiment of the posts from February 3 - 23 to April 6 - 26. In column 6 it is whether the individual, as of June 28, was a member of a 'Reopen' Facebook group. Post and group classifications are defined in Appendix C. All columns include controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. All columns include fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table A10: Heterogeneity of Friend-Exposure Effects - Friend Characteristics

	Share Posts about Corona	Share "Signed Posts" Opposed to Distancing (Feb - Apr)	Δ Sentiment (Feb - Apr) All Posts	Member "Reopen Group" by May 24, 2020
log(Friend Exposure - Rank 1 - 25)	0.061***	-0.360***	-0.032**	-0.053***
	(0.002)	(0.149)	(0.016)	(0.002)
log(Friend Exposure - Rank 26 - 50)	0.046***	-0.299*	0.013	-0.036***
	(0.002)	(0.160)	(0.016)	(0.002)
log(Friend Exposure - Rank 51 - 75)	0.033***	-0.433**	0.008	-0.053***
	(0.002)	(0.158)	(0.017)	(0.002)
log(Friend Exposure - Rank 76 - 100)	0.022***	-0.016	-0.037**	-0.051***
	(0.002)	(0.159)	(0.017)	(0.002)
Percentiles of Total Number of Groups (Feb 2020)				Y
Other Network Exposure FE	Y	Y	Y	Y
Zip Code x Age Group x Gender x Has College x Has Tablet x Has iPhone	Y	Y	Y	Y
R-Squared	0.060	0.446	0.122	0.074
Sample Mean	1.869	35.319	-1.869	0.012
F Test (Rank 1-25 = Rank 76-100)	184.345***	2.180	0.045	1.352
N	30,814,578	255,095	9,482,790	108,911,020

Note: Table shows results from regressions of friend-exposure to COVID-19 on March 15th, calculated using limited friend sets, on a number of outcomes. Each observation is an individual. Friend-exposure is calculated using only subsets friends based on the strength of friendship connections. The outcomes in columns 1 and 2 are the change in probability of staying at home and the percent reduction in the average number of tiles visited, respectively, from the week of February 25th to March 2nd (prior to the pandemic) to April 14th to 20th. The outcome in column 3 is the percentage of individual posts that are about COVID-19. In column 4 it is the percentage of pro- or anti-lockdown posts that are anti-distancing. In column 5 it is the change in the average sentiment of the posts from February 3rd through 23rd to April 6th through 26th. In column 6 it is whether the individual, as of June 28th, was a member of a 'Reopen' Facebook group. Post and group classifications are defined in Appendix C. All columns include controls for percentiles of friend-exposures (as described in equation 2) for median household income, population density and the share of the population living in urban areas. All columns also include fixed effects for every group constructed from interacting ZCTA, age group, gender, has college, has tablet, and has iPhone. Standard errors are clustered by ZCTA. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

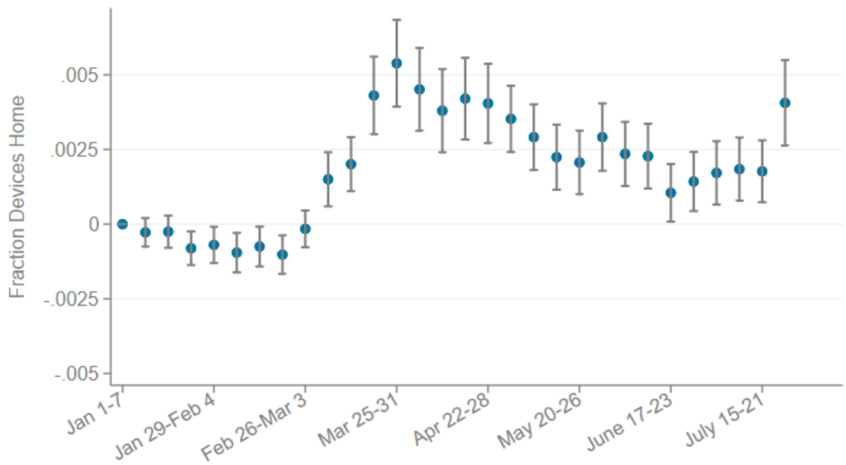
Table A11: Summary Statistics of ZCTAs with High and Low Friend-Exposure to COVID-19

	Low Friend-Exposure		High Friend-Exposure	
	Mean	SD	Mean	SD
Fraction Male	0.49	0.03	0.49	0.03
Fraction White	0.74	0.23	0.72	0.21
Fraction Black	0.12	0.18	0.13	0.18
Fraction Asian	0.05	0.08	0.06	0.09
Median HH Inc.	$65426.94	$24643.08	$64707.44	$28886.90
Management, Business, Science, Arts	0.17	0.08	0.18	0.08
Service Occupations	0.08	0.02	0.08	0.03
Production + Transportation	0.07	0.03	0.06	0.03
Fraction Age <18	0.23	0.05	0.22	0.05
Fraction Age 18-24	0.09	0.03	0.10	0.07
Fraction Age 25-34	0.14	0.04	0.14	0.05
Fraction Age 35-44	0.13	0.02	0.13	0.02
Fraction Age 45-54	0.14	0.02	0.13	0.02
Fraction Age 55-64	0.13	0.03	0.13	0.03
Fraction Age 65-74	0.09	0.03	0.09	0.03
Fraction Age >= 75	0.06	0.03	0.07	0.03
Fraction High School / GED	0.20	0.07	0.17	0.07
Fraction Some College	0.20	0.05	0.19	0.05
Fraction College Degree	0.19	0.11	0.23	0.13
Population Density	1606.18	4122.32	1531.22	3490.32
Fraction High-Speed Internet	0.80	0.11	0.80	0.11
Population	30175.11	21494.15	32923.65	20222.50
Mean Number of POIs	435.91	372.62	538.79	376.59
Number of ZCTAs	14079		11880	

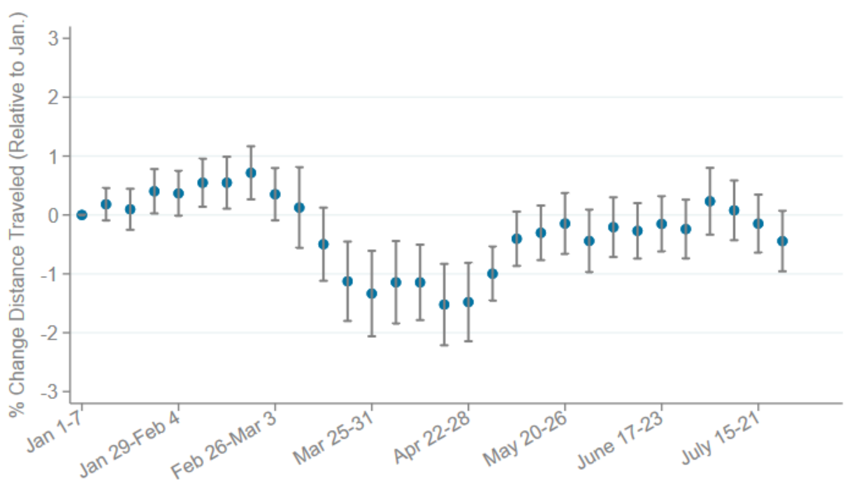
Note: Table presents ZCTA-level summary statistics for the sample used in Section 5.3. Definitions of high- and low-exposure areas are based on friend-exposure to COVID-19 as defined in equation 16. High-exposure ZCTAs are ZCTAs with friend-exposure to COVID-19 above the median for corresponding county. Similarly, low-exposure ZCTAs are places with friend-exposure below that median. Medians are defined based on the number of COVID-19 cases as of March 17. Data on covariates is obtained from the 2014-2018 ACS data. [Return to text]

Figure A11: Coefficient Estimates for β_{1t} Equation 17

(a) Fraction of Devices Home



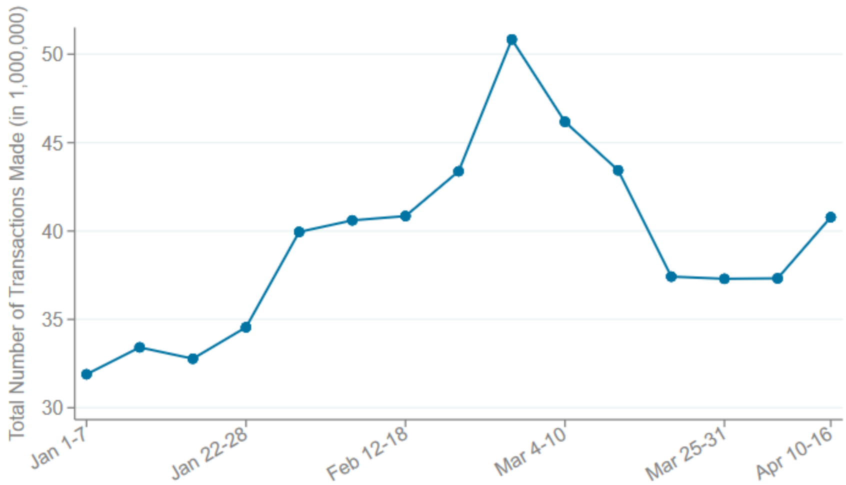
(b) Percentage Change Avg. Distance Traveled



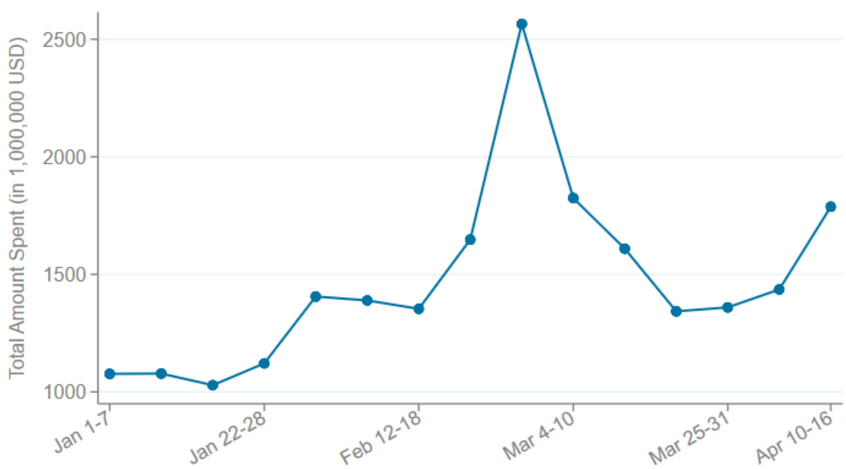
Note: Figures show coefficient estimates based on equation 17. In Panel (a), the dependent variable is the fraction of devices at home, while in Panel (b), the dependent variable is the percentage change in average distance traveled relative to the month of January 2020. The unit of observation is ZCTA by week. Regressions include a rich set of controls: in addition to ZCTA fixed effects and county fixed effects interacted with week indicators, we additionally control for a rich set of covariates interacted with week indicators. These covariates are the fraction of people being male, the fraction of asian/black/white people, median household income, the fraction of individuals working in service occupations, the fraction of individuals working in production or transportation, the fraction of individuals working in management, arts or science, the fraction of individuals with a high school degree, some college education and a college degree as well as the fraction of households with high speed internet. We also include various age-related controls, i.e. the fraction of individuals 18 or younger, between 18 and 24, between 25-34, between 35-44, between 45-54, between 55-64, between 65-74 and above 75. All these control variables are obtained from the most recent 5-year ACS (2014-2018). In addition, we also control for ventiles of friend-exposure to other characteristics, namely income, population density (both from 2014-2018 ACS) and urbanity (from 2010 Census), again interacted with week indicators. Standard errors are clustered at the ZCTA-level. [Return to text]

Figure A12: Total Number of Transactions by Week

(a) Total Number of Transactions Made



(b) Total Amount Spent



Note: Figures show variation in spending behavior over time using data from Facteus. In Panel (a) we present a time series plot of the total number of transactions made by week in millions. In Panel (b) we show a time series plot of the total amount spent in USD in millions.

C Logic for Post and Group Classifications

To classify posts and groups in certain analyses, we use regular expression searches. Posts or groups are flagged if they match one more of the regular expressions described.

C.1 Post Classification

We classify public Facebook posts made between February 3rd and May 3rd according to the regular expressions in Table B1. Posts that match any of "neutral-lockdwon", "pro-lockdown", or "anti-lockdown" are classified as COVID posts.

Table B1: Posts Regular Expression Classification

	Neutral Lockdown	
%corona%	%covid%	%pandemic%
%sars%	%#socialdistancing%	%lockdown%
%stay at home%		
	Pro Lockdown	
%#staysafe%	%#stayhome%	%#bendthecurve%
%bend the curve%	%#flattenthecurve%	%flatten the curve%
%#crushthecurve%	%crush the curve%	%#safeathome%
	Anti Lockdown	
%#liberate%	%#endtheshutdown%	%#endthelockdown%
%#reopen%	%#openamerica%	%#stoptheshutdown%
%#stopthelockdown%	%against%quarantine%	%end the lockdown%
%end the shutdown%	%open now%	%hysteria%
%open the states%	%openthestates%	%lockdown%dictator%
%lockdown%oppress%	%lockdown%tyranny%	%lockdown%liberty%
%lockdown%freedom%	%shutdown%dictator%	%shutdown%oppress%
%shutdown%tyranny%	%shutdown%liberty%	%shutdown%freedom%
%dictator%lockdown%	%oppress%lockdown%	%tyranny%lockdown%
%liberty%lockdown%	%freedom%lockdown%	%dictator%shutdown%
%oppress%shutdown%	%tyranny%shutdown%	%liberty%shutdown%
%freedom%shutdown%		

Note: Table presents the regular expressions used to flag posts about COVID. % is a wildcard capturing any number of characters (including 0).

C.2 'Reopen Group' Classification

We classify public Facebook groups as a 'Reopen Group' if it was created between March 1st and June 28th, 2020 and has a (case-insensitive) name that matches one of the following regular expressions:[29] "%reopen%", "%liberate%", "%end%shutdown%", "%end%lockdown%", "%against%quarantine%."

[29]% is a wildcard capturing any number of characters (including 0).