The Minimal Effects of Public Health Campaigns on Travel During the COVID-19 Pandemic

August 12, 2024

Adam Zelizer[a]
Ethan Bueno de Mesquita[b]
Mehdi Shadmehr[c]
Benjamin Shaver[d]

Abstract

We explore the effectiveness of a public health campaign via a large-scale field experiment during the COVID-19 pandemic. In 2020, approximately 125,000 households were sent mail from health experts encouraging social distancing. Targeted geographies received varying dosages of treatments, and some targeted households were informed that neighbors also received the messages, to highlight the strategic considerations underlying social distancing. We find minimal effects of the campaign on holiday travel and on other social distancing behavior. We consider our results in light of a similar intervention and conclude that travel was unlikely to be influenced by messages from public health experts.

Keywords:
COVID-19; Social distancing; Field Experiment; Mass Media; Social Norms

[a] Harris School of Public Policy, The University of Chicago. 1307 East 60th Street. Chicago, IL 60637. zelizer@uchicago.edu

[b] Harris School of Public Policy, The University of Chicago. 1307 East 60th Street. Chicago, IL 60637. bdm@uchicago.edu

[b] University of North Carolina at Chapel Hill. Abernethy Hall, 209. Chapel Hill, NC 27599. mshadmehr@unc.edu

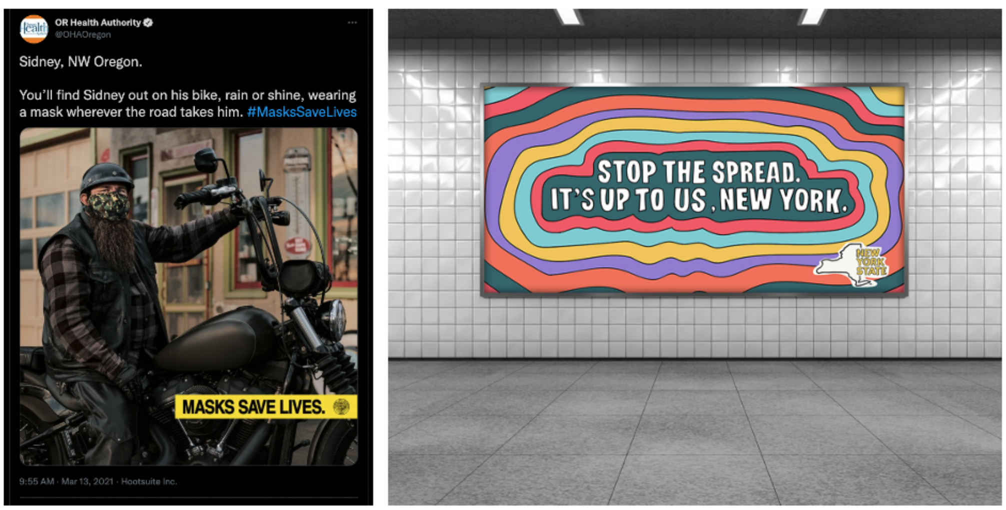
[d] Harris School of Public Policy, The University of Chicago. 1307 East 60th Street. Chicago, IL 60637. blshaver@uchicago.edu

This material is based upon work supported by the National Science Foundation under Grant No. 2035682. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. This work also benefitted from support from the Becker Friedman Institute. This RCT was registered as AEARCTR-0012148 (Zelizer et al., 2023).

Declaration of Interest: Bueno de Mesquita has an ongoing consulting agreement with Facebook on matters unrelated to the research in this paper.

During the COVID-19 pandemic, governments and public health officials relied on mass-media public health campaigns to mitigate the spread of the virus. During the first six months of the pandemic, the National Association of Broadcasters estimated that it provided over $150 million of airtime on radio and television to coverage of the pandemic.[1] As of summer 2022, the Department of Health and Human Services aimed to reach 90% of the American public at least four times per year, and on average 10 times, with content about vaccine safety and efficacy.[2] City and state governments, healthcare providers, educational institutions, and corporations encouraged social distancing, masking, and vaccinations via mass media. Figure 1 shows examples of two such campaigns from Oregon and New York City, on Twitter and in the subways, respectively.

Figure 1: Public Health Messages during Covid-19 Pandemic



A tweet from the Oregon Health Authority (left) and a mural in the New York City subway (right).

Surprisingly, these public health campaigns were carried out despite a limited track record of success. Past observational studies evaluating public health campaigns found mixed evidence for campaigns' effectiveness (Wakefield et al. 2006; Wakefield, Loken, and Hornik 2010; Longshore, Ghosh-Dastidar, and Ellickson 2006; Hornik et al. 2008; Richardson et al. 2011;

[1] `https://web.archive.org/web/20220706192202/https://www.nab.org/coronavirus/`

[2] `https://web.archive.org/web/20220706192339/https://wecandothis.hhs.gov/resource/we-can-do-campaign-background`

Sly, Heald, and Ray 2001). The few field experiments evaluating public health campaigns are even less promising. Experimental evaluations of campaigns targeting smoking (Bauman et al. 1991), teen pregnancy (Green, Zelizer, and Lin 2021), and contraceptive use (Byker, Myers, and Graff 2019) find no evidence that media campaigns change behavior.

Results from recent experiments on the effectiveness of COVID-related health messaging have departed from this trend. In an experiment covering millions of Facebook users across 13 states, Breza et al. (2021b) randomized messages from local healthcare providers on the risks of holiday travel. They report substantial effects on some measures of travel and on COVID infections after the holiday. Larsen et al. (2023) evaluates a different intervention and outcome — the experiment randomized YouTube ads featuring Donald Trump speaking positively about vaccines and tracked vaccination rates — and also finds large effects of the media message.

We add to this emerging experimental literature on the effectiveness of public health media campaigns during the COVID-19 pandemic (Lin and Nan 2022). In partnership with the University of Chicago Medicine, we mailed 125,000 postcards to Cook County residences with guidance from Dr. Emily Landon, the lead epidemiologist for the University on COVID-19, warning against congregating over the holidays. We examine whether these messages reduced holiday travel measured using cell phone tracking data and a follow-up survey.

In addition to evaluating the effectiveness of public health messages, our study also examines the strategic considerations behind social distancing behavior. Like voting, protest, and many political and economic activities, preventing the spread of COVID-19 invokes strategic considerations that complicate individual behavior (Bursztyn, González, and Yanagizawa-Drott 2020; Cantoni et al. 2019; Bueno de Mesquita and Shadmehr 2023). Individuals may free-ride on the pro-health behaviors of others, as others' masking or social distancing reduces the risks of oneself catching COVID, in which case social distancing would be marked by strategic substitutes. Alternatively, people may feel social pressure to comply with norms displayed by others, in which case social distancing would be marked by behavioral complements. As a result, individual behavior depends on beliefs about others' compliance and on social norms regarding compliance (Barrios et al. 2021; Campos-Mercade et al. 2021; Bazzi, Fiszbein, and Gebresilasse 2021; Durante, Guiso, and Gulino 2021).

To evaluate whether social distancing behavior is marked by behavioral complements or substitutes, we randomize the percentage of treated households in neighborhoods assigned to treatment and randomize information about treatment intensity by printing the percentage of the neighborhood treated directly on the mailers. Our focus on the strategic considerations underlying public health behavior is most similar to the survey experiments conducted by Moehring et al. (2023), which finds that informing respondents of fellow citizens' vaccination intentions increases respondents willingness to be vaccinated across 23 countries, and Allen IV et al. (2021), which shows that informing respondents in Mozambique of community support for social distancing could cause either free-riding or conformity depending on COVID-19 case loads in the community.

Across outcomes, we find no evidence that the campaign influenced social distancing behavior. Aggregating the treatment conditions, we estimate precise null effects on both behavioral and survey measures of social distancing. Despite the null effects on behavior, survey respondents did recall the mailers a month after they received them. Recall rates were substantially higher among the publicity treatment conditions, so invoking social considerations did draw attention to the message but did not change behavior.

Study Context and Design

Our experiment was fielded in mid-December 2020. Despite public health officials and policymakers warning of the risks of travel, the Thanksgiving holiday was the busiest travel period since the beginning of the pandemic.[3] Due to a surge of cases after Thanksgiving, public health officials redoubled their efforts prior to the Christmas holiday. Some communications strategies were blunt — Mississippi State Medical Association President Mark Horne said

[3]https://www.cnn.com/travel/article/thanksgiving-travel-volume-2020-pandemic/index.html

"We don't really want to see Mamaw at Thanksgiving and bury her by Christmas" (Wan and Shammas 2020) — while others provided more nuanced, expert guidance about which activities were riskier than others (Wust 2020).

Treatment

Our experiment featured mailers with warnings about holiday travel from the chief infectious disease epidemiologist at The University of Chicago Medicine, Dr. Emily Landon.[4] One side of the half-sheet mailer (pictured in Figure 2) quotes Dr. Landon saying, "Your grandpa doesn't need to go to a bar to get COVID-19 from a bar. He could get it from your cousin who went to a bar last week and is now unknowingly spreading the virus." The other side of the mailer includes gentle pro-social messaging calling on recipients to pull together as a community to keep each other safe by social distancing (Figure 3). Together, expert guidance combined with stark warnings about the costs of socializing are representative of many of the types of health campaigns fielded at the time. All treated individuals received this social distancing guidance from a UChicago Medicine expert.

Our mailings diverged from more standard public health campaigns in an important respect. First, we randomized treatment dosages within treated geographies. In neighborhoods assigned to treatment, we assigned either half or all households to the treatment. Randomizing the dosage allows us to investigate whether there are increasing or decreasing returns. If social distancing behavior was characterized by complementarities, then treatment effects should be increasing in the number of treated individuals; if substitutes, decreasing.

Second, we assigned some treated individuals to a publicity treatment: they were informed of the share of neighbors sent the message (see Figures A1 and A2 in the Appendix). One stated that "half of your neighbors are also receiving this message," while the other said "all of your neighbors are also receiving this message." Both messages were truthful.

[4]Dr. Landon also advised Illinois Governor J.B. Pritzker and the University of Chicago on state and local responses to the COVID-19 pandemic and frequently appeared at news conferences and in the media to discuss the pandemic.

Randomizing the publicity treatment was meant to change recipients' beliefs about the treatment status and social distancing behavior of their neighbors. If randomizing the dosage of treatments exerted too small effects to reveal increasing or decreasing returns to social distancing, then directly telling recipients how many of their neighbors were treated might change their beliefs about others' behavior and thus their own.

Interacting these two treatments on dosage and publicity yields four treatment arms in addition to control. Individuals were either assigned to 50% dosage with no publicity; 50% dosage with publicity; 100% dosage with no publicity; or 100% dosage with publicity.

The COVID-19 pandemic was an appropriate setting for estimating behavioral complementarities or substitutes for two reasons. First, social distancing, masking, and vaccinations all feature strategic considerations. Second, even with disparities in risk, resources, and exposure to the virus, the COVID-19 pandemic had a uniquely broad impact. Nearly everyone's daily life was changed in some way. As a result, we were able to field our experiment with the entire population of the second largest county in the US as our sample.

Figure 2: Mailer Common Content: Message from Emily Landon, MD

Figure 3: Mailer Common Content: Pro-Social Messaging

Study Population and Treatment Assignment

Mailers were sent to residents of Cook County, Illinois. We obtained contact information for the 4.4 million adults residing in Cook County from a private data vendor. Nearly all residents had valid addresses, while slightly less than half had phone numbers. To avoid sending multiple mailers to the same household, we randomly selected one adult per household to receive the mailer and be called for the survey.

The unit of treatment assignment was the Census Block Group (CBG). Cook County has approximately 4,000 CBGs, which range in size from one household to nearly 3,800. To improve statistical power and optimize our budget, we dropped particularly small and large CBGs from the study. Our population consists of the 3,600 CBGs with between 98 and 799 households. This sample covers nearly 3.6 million people. 100 CBGs were randomly assigned to each of the four treatment arms. We clustered treatment assignment at this level because one of our outcome measures, cell phone tracking data, is aggregated at the CBG-level.

Table 1 describes key demographic characteristics of CBGs assigned to control or to any treatment condition.[5] Treatment areas have slightly higher incomes and slightly lower populations of African Americans and Hispanics than control areas, but all differences are small relative to sampling variability.

Table 1: **Balance of Control and Treated Census Block Groups on Key Covariates**

	Control	Any Treatment	Difference
% Female	53.1	53.1	0.1
($\widehat{\text{SE}}$)	(0.1)	(0.2)	(0.2)
Mean Age (in Years)	49.7	50.0	0.3
	(0.1)	(0.2)	(0.2)
Mean Income ($ in '000s)	71.4	75.4	4.0
	(1.1)	(3.1)	(3.3)
% African American	29.6	27.9	−1.7
	(0.6)	(1.8)	(1.9)
% Hispanic	22.5	20.7	−1.9
	(0.4)	(1.1)	(1.2)
N CBGs	3200	400	
N Individuals (in '000s)	3,177.5	391.9	

Outcome Measures

We use both behavioral and survey-based measures of social distancing. The first set of outcomes use cell phone tracking data. We acquired data from a commercial vendor that tracks the movement of individual cell phones. To maintain privacy and anonymity, the vendor aggregates movements at the level of the Census Block Group.[6] The vendor provides daily mobility data so we can precisely track movements for the months before, during, and after the holidays.

No one mobility metric perfectly captures social distancing behavior, so we use several. First, the number of devices in the CBG shows whether there is net inflow or outflow of individuals, for example if someone travelled away from home for longer than one day. The

[5] Table B1 in the Appendix displays demographics across each of the four treatment arms.

[6] We randomized at the level of the CBG to match the availability of this cell phone data. To further ensure individual privacy, the vendor drops CBGs with few phones.

second measure is mean distance travelled,[7] which is one outcome used in a similar study (Breza et al. 2021b). Third and fourth are the percentage of devices that stay at home all day and the amount of time spent outside the home. Complete stay-at-home behavior is a pure, if extreme, type of social distancing, while the amount of time spent out of the house shows whether, and for how long, individuals attended a holiday party, for example. Neither shows whether individuals invited others to their home.

To supplement these data and ask tailored questions about social distancing, we conducted a phone survey. We contacted treated and untreated individuals in January 2021, a month after mailers arrived but sufficiently long after the New Year holiday for any holiday-related travel to end. We called nearly 200,000 households, all the treated households plus nearly 130,000 untreated households, via an automated phone survey.[8]

The survey asked about a range of activities related to social distancing. We asked whether the individual shopped for holiday gifts in-store or online; visited a bar or restaurant; visited the house of a friend or extended family member; wore a mask when out in public; congregated in a group of 10 people or more for the holidays; or saw their neighbors congregate to celebrate the holidays. We concluded each survey by asking whether the responded recalled seeing a mailer from UChicago Medicine highlighting the risks of travelling over the holidays.

Table 2 displays the demographic characteristics of the entire survey sample and those who responded to the survey. Compared to the entire study population described in Table 1, the survey population is substantially older, higher income, and less Hispanic likely due to the profile of individuals who still have landlines. Even among those surveyed, though, respondents are older, higher income, and less likely to be Hispanic. The differences between respondents and the sample result from low response rates and self-selection into the survey.

[7] The vendor filters out any distances travelled larger than 1.5 times the interquartile range for that device, which is undesirable for our purposes. We replicate analyses using the median distance travelled among all trips during a given period, which does not explicitly filter trips.

[8] There were nearly 1.1 million control households; surveying them all would have gained little power at substantial expense. We randomly selected a percentage for the survey.

Table 2: **Demographics of Survey Respondents**

	Survey Sample	Respondents	Difference
% Female	54.2	54.1	-0.1
Mean Age (in Years)	57.1	62.1	5.0
Mean Income ($ in '000s)	85.0	92.5	7.5
% African American	28.6	29.7	1.1
% Hispanic	15.2	6.2	-9.0
N Individuals	195,514	3,393	

Only 1.7% of calls resulted in a completed survey. While low, this rate is characteristic of automated phone surveys, and, importantly, response rates do not vary significantly across treatment conditions (Table 3).

Table 3: **Survey Response Rates (in pp) by Treatment Condition**

		Treatment of Census Block Group					
		Control	Any Treatment	50%-Private	100%-Private	50%-Public	100%-Public
Individual Treated?	No	1.77	1.69	1.71	–	1.66	–
	(N)	(102,110)	(22,849)	(11,550)	–	(11,299)	–
	Yes	–	1.70	1.75	1.78	1.53	1.68
	(N)	–	(70,555)	(11,462)	(23,771)	(11,540)	(23,782)

Methods

To estimate the effects of our mailers on cell phone-based mobility data, we use a difference-in-differences specification. Because our research design is an experiment in which treatment is randomly assigned, we use the difference-in-differences to improve precision, not because of concerns about identification. Our specification is the following:

$$Y_{it} = \gamma_i + \lambda_t + \delta(D_i * Post_t) + \epsilon_{it} \tag{1}$$

with fixed effects γ_i and λ_t for CBGs and days, respectively; D_i indicating treated CBGs and $Post_t$ the holiday period (i.e. the post-treatment period); and δ the average treatment effect. We include geographic and time fixed effects to improve power. We equally weight each CBG and do not reweight by population.[9] Thus the estimand is the average treatment effect across CBGs.[10]

Mailers began to arrive at treated households on December 11, 2020.[11] Our analysis uses data from October 1 – December 10 for the pre-period and December 24 – January 1 for the post-period. Our specification does assume that units cannot anticipate their treatment statuses prior to the arrival of the mailers, which seems reasonable in this case. We invoke no persistence assumptions.[12]

Table 4 displays the estimated effects of any treatment, and of each of the four treatment arms, on social distancing via the cell phone-based mobility data. We find null effects of assigning a CBG to any treatment on all four outcome measures. The number of devices per CBG is estimated to decline by 0.8 ($\widehat{\text{SE}} = 0.8$), or only about 1.5% from baseline levels, due to treatment, which is consistent with sampling variability.[13] Treatment is estimated to slightly increase the proportion of devices that stay-at-home ($\widehat{\text{ATE}} = 0.4$ percentage points, $\widehat{\text{SE}} = 0.3$) but to also increase distance travelled ($\widehat{\text{ATE}} = 0.1$ kilometers, $\widehat{\text{SE}} = 1.0$) and the average amount of time spent away from home ($\widehat{\text{ATE}} = 1.0$ minutes, $\widehat{\text{SE}} = 3.9$). These estimates are precise nulls: we can confidently reject that this treatment caused 1 out of 100 devices to stay home during the holidays or to spend 10 minutes less outside the home.

Table 4 shows that some of the four individual treatments do exhibit large, statistically-

[9]Because treatment was implemented at the same time for all treated units, the two-way specification does not cause the pathological reweighting of units identified in the recent diff-in-diff literature.

[10]Weighting CBGs by population would change the estimand to the ATE among individuals. We see no pressing reason to prefer one estimand to the other in this study, and weighting the CBGs would both decrease precision and potentially bias analyses due to population differences across CBGs.

[11]We sent mailers from each treatment condition to confederates located across Cook County. December 11 was the date the first of these confederates reported seeing the mailer.

[12]One could examine social distancing in the period after the holidays to estimate treatment persistence, but since we find null effects during the holidays, we would not expect, nor trust, non-null persistence effects.

[13]Another reason to estimate the effect of treatment on devices is that if the treatment did cause individuals to stay (or leave) their home during the holidays, the remaining outcome measures would be subject to selection bias. That we find no evidence of large effects on the number of devices assuages that concern.

Table 4: **Estimated Effects of Treatment on Social Distancing: Cell Phone Mobility**

	Any Treatment	50%-Private	100%-Private	50%-Public	100%-Public
Number of Devices (Control Mean: 54.9)					
$\widehat{\text{ATE}}$	-0.8	-0.2	-1.0	0.9	-2.9
$(\widehat{\text{SE}})$	(0.8)	(1.4)	(0.7)	(1.6)	(1.7)
Average Distance Travelled (9.4 km)					
	0.1	-3.0	2.4	2.5	-1.5
	(1.0)	(1.1)	(2.7)	(2.1)	(1.4)
Completely Stay-at-Home Devices (32.9%)					
	0.4	0.3	0.8	0.4	0.2
	(0.3)	(0.6)	(0.6)	(0.6)	(0.6)
Average Time Out-Of-Home (224 minutes)					
	1.0	2.4	-5.5	4.2	3.1
	(3.9)	(8.5)	(6.3)	(7.8)	(7.4)

More conservative of two-way clustered and CBG-clustered standard errors and p-values displayed.

significant effects on outcomes. In particular, the 50% dosage, private treatment is estimated to substantially reduce average distance travelled by 3 kilometers (1.1). With four treatment arms and four outcomes, we interpret this as a chance result that likely arose due to sampling variability and multiple testing. Empirically, some of the other treatment conditions exhibit estimated effects on this same outcome nearly as large but in the opposite direction. Theoretically, there is little reason to expect the 50% private dosage treatment to exert larger effects than the 100% dosage conditions or the 50% public condition. If there were both a direct effect such that treatment increased social distancing but strategic substitution such that compliance decreased with the percentage of treated individuals, we might expect that phenomenon to arise even more strongly in the public conditions.

We now turn to analysis of the follow-up survey. Treatment effects are estimated via the simple difference-in-means. Standard errors and p-values are clustered by CBG.[14] Since only some households in the 50% dosage treatment arms were treated, we analyze untreated

[14]For this analysis, the average treatment effect is across individuals, not CBGs. With 3,393 respondents across 1,656 CBGs, the most populated CBG has 16 respondents.

Table 5: **Estimated Effects of Treatment on Social Distancing: Survey**

	Any Treatment	50%-Private	100%-Private	50%-Public	100%-Public
Holiday Shopping Online (vs. In-Person) (Control Mean: 65.0%)					
$\widehat{\text{ATE}}$	1.8	2.6	2.5	2.7	1.9
$(\widehat{\text{SE}})$	(1.9)	(3.5)	(2.6)	(4.4)	(3.1)
Visited Bar or Restaurant (15.5%)					
	0.0	-2.8	0.5	-3.2	2.4
	(1.3)	(2.4)	(1.9)	(2.6)	(2.1)
Visited House of Friend or Extended Family Member (31.5%)					
	-0.1	-0.2	-0.3	0.8	-1.6
	(1.6)	(3.1)	(2.6)	(4.0)	(2.6)
Wear a Mask in Public (94.3%)					
	-0.7	-2.9	0.1	-0.8	-0.5
	(0.8)	(1.9)	(1.2)	(1.8)	(1.2)
Celebrate Holidays in Group (7.0%)					
	-0.8	-3.3	-2.2	1.9	0.1
	(0.9)	(1.4)	(1.3)	(2.2)	(1.5)
See Neighbors Gathering for Holidays (9.1%)					
	2.6	1.8	1.3	0.3	5.1
	(1.2)	(2.3)	(1.7)	(2.6)	(2.0)
Recall Mailer (13.8%)					
	2.4	0.1	0.4	5.9	4.2
	(1.3)	(2.7)	(2.1)	(2.9)	(2.7)

Standard errors and associated p-values clustered at CBG-level. Control observations only include respondents in control CBGs; untreated respondents in treated CBGs are analyzed separately due to potential spillover.

households in treated CBGs separately and do not include them in the control group.[15]

Table 5 displays estimated average treatment effects on the survey. Again, we observe small, precise, null treatment effect estimates. Treatment is estimated not to have substantially changed respondents' visits to bars, restaurants, friends' houses, or holiday gatherings. The largest estimated effect on behavior is via shopping: treatment is estimated to have increased online shopping relative to in-person by 1.8 percentage points ($\widehat{\text{SE}} = 1.9$).

One reason for these null effects is that respondents to our survey were natural compliers with public health directives: only 7% admitted to celebrating the holidays with a group of 10 people or more. Against this baseline, treatment effects of 0.8 percentage points are substantively large, but still consistent with sampling variability.

The one survey outcome that demonstrates meaningful treatment effects is not about respondents' own behavior, but instead their neighbors'. Treated households were 2.6 percentage points (1.2) more likely to report that their neighbors congregated over the holidays in groups of 10 or more people. Against a baseline rate of 9%, this is a substantial increase. This result may again be spurious, driven by multiple comparisons, and that is our preferred interpretation. Yet there are reasons to expect it is more meaningful than the significant estimate from the cell phone analysis. All four treatment conditions exhibit positive treatment effects on neighbors' congregation, and the largest effect is among the 100% dosage, public treatment condition. If further work verifies this effect is real, it would suggest that treatments encouraging pro-social behavior might cause recipients to monitor or report others' behavior more severely.

Taken together, the behavioral and survey outcomes indicate that messages from a widely-respected, local public health authority had no effect on social distancing over the 2020 holidays. In the next section, we consider why that may have been the case.

[15]Untreated households in treated CBGs were potentially exposed to spillover. We display estimated spillover effects in Table D1 in the Appendix.

Placing Our Results in the Literature

Together, our analyses of social distancing behavior during the 2020 holiday season suggests that messages from a high-profile medical expert exerted no influence on recipients. How do we make sense of our results, particularly with respect to the results of two recent field experiments, including one that examines warnings from health experts against congregating over the 2020 holidays and cell phone-based mobility data?

One explanation for null effects would be if respondents simply did not see or read the mailer. We estimate that treated respondents were 2.4 percentage points more likely to recall seeing the UChicago Medicine mailer than control respondents ($\hat{p} = 0.09$ two-sided). Recall effects were highest at 4 – 5 percentage points in the two public treatments — which included the somewhat unusual, and potentially jarring, notice that the mailers were being sent to 50% or 100% of one's neighbors — but nearly 0 for the two private treatments. Nevertheless, recall rates of 2 – 5 percentage points are not high by absolute standards.

A recall effect of 2 – 5 percentage points is, however, consistent with previous studies of media effects. For comparison, Broockman and Green (2014) estimate recall effects of about 4 percentage points from two experiments with randomly-assigned Facebook ads, and the follow-up surveys in those experiments occurred only one to seven days post-treatment. Because our messages were sent prior to Christmas and our survey fielded after New Year's, a month elapsed before we surveyed voters, which would substantially attenuate recall effects. Green, Zelizer, and Kirby (2018) estimate recall of 12 – 24 percentage points from mailers, but the treatments were much more noticeable, flashy, 4 – 8 page newspapers reporting on local political scandals, and surveys were again fielded within a week or two of treatment. Recall rates of our mailer treatment do not necessarily suggest that recipients tuned out the message; indeed, recall rates were similar to ads delivered via Facebook.

Another explanation for the apparent disagreement is that there is actually no conflict at all—both studies are consistent with small effects of public health messages on travel. The paper that is most similar to ours, Breza et al. (2021b), examines the effects of messages from health experts on mobility during the 2020 holiday season; the key difference is that the treatment is delivered via video, through Facebook, for a much larger study population over 13 states. The paper reports that "average distance traveled in high-intensity counties decreased by -0.993 percentage points (95% confidence interval (CI): -1.616, -0.371; P = 0.002) for the 3 days before each holiday compared to low-intensity counties" (Breza et al. 2021b, p. 1622). Although we estimate a small increase in travel due to messaging, our 95% confidence interval includes the point estimate from Breza et al. (2021b).

Considering the totality of the evidence, we conclude that it is likely that the effects of mass media are toward the smaller end of those estimated in Breza et al. (2021b). We reach this conclusion both because of the new evidence provided by our study and back-of-the-envelope calculations that suggest the point estimates of the treatment effects in Breza et al. (2021b) require quite large changes to individual behavior in response to treatment given low treatment receipt rates. Although Breza et al. (2021b) estimate only a 1 percentage point effect on mobility, this effect implies large behavior changes given three details about the experiment: 1) the study compares mobility in high-intensity counties, where 75% of zip codes were treated, to low-intensity counties, where 25% of zip codes were treated, so the difference in treatment probabilities across the two conditions is 50 percentage points. 2) Only a percentage of Facebook users who were shown the ad actually watched it. In the NBER working paper, Breza et al. (2021a, p. 6) report that 12–13% of Facebook users watched the ads for at least 3 seconds, and 1–2% for at least 15 seconds. If these watch metrics are accurate, a realistic best case scenario is that about 10% of targeted users watched enough of the ad to know what it was about and get a meaningful treatment. 3) Only 35–66% of targeted users were actually reached with the ad, presumably because some users did not log into or spend sufficient time on Facebook during the period when it was pushed to them (Breza et al. 2021b, p. 1,623).

Putting together these three numbers — the 50% difference in treatment assignment probabilities across geographies, the c. 50% of targeted users who were shown the ad, and the c. 10% of users shown the ad who watched a meaningful part of it — shows the difference in the percentage of Facebook users who saw the ad across the two treatment conditions was about 2.5 percentage points. But treating these additional 2.5% of users led to a 1 percentage point decrease in distance travelled among the entire sample, which is a large change in travel behavior.[16]

We conclude that the effect in Breza et al. (2021b) is likely an overestimate due to sampling variability, and that the true effect of messages from the experts was likely substantially smaller. We point this out not to criticize Breza et al. (2021b); the paper is a model of open science. Procedures and outcomes were pre-registered, including this analysis of mobility, and the results in the paper are based on the pre-registered specification. Yet, the extra information revealed during the course of their experiment regarding treatment administration rates — information that the authors themselves disclosed in their papers — combined with the results of our own study lead us to believe that messages from public health experts likely exerted much smaller, if any, influence on public travel over the 2020 holidays.

Conclusion

We evaluate the effectiveness of a public health campaign conducted in Cook County, IL during the COVID-19 pandemic in winter 2020. Messages from the chief epidemiologist for the University of Chicago Medicine were sent to randomly selected Census Block Groups across the county encouraging residents not to travel or congregate during the holidays. Analysis of cell phone-based tracking data as well as a follow-up survey show that the messages had minimal impact on social distancing behavior of any type.

Because the messages we sent out were not influential, we were able to learn little about the strategic aspect of social distancing behavior. We cannot say whether social distancing

[16] One potential explanation for this large effect is that each user who was reached by the video was shown it 2.6–3.5 times (Breza et al. 2021b, p. 1,623). Not only does this dosage make the treatment more effective, but it might also resolve the low watch rates: perhaps a viewer shown the ad three times watches it once, then quickly skips it the next two times. Nevertheless, even tripling the watch rate would still require up to 15% or more of treated users to modify their behavior, and that seems a heavy lift.

is marked by behavioral complements or substitutes because we saw no meaningful changes in behavior at all. The study does provide two pieces of evidence that merit further study. The first concerns recall of messages and publicity. Informing recipients of the share of their neighbors who had received the same message substantially increased recall of the messages. Second, these messages that prescribed social norms and exerted social pressure may have caused recipients to monitor or report on their neighbors' behavior. Establishing new norms around social distancing may make compliers more comfortable judging the activities of others.

Our main result adds to a growing literature in which randomized controlled trials have been used to evaluate the effectiveness of public health campaigns. Across a range of topics — from anti-drugs and anti-smoking to teen pregnancy prevention and social distancing during COVID-19 — such RCTs have found minimal effects of mass media campaigns. These effects are at the margin, so they do not indicate that all public health campaigns are ineffective, but they do suggest that the marginal dollar spent on mass media will not be particularly effective. We hope that future government- or nonprofit-funded mass media campaigns will continue to evaluate their effectiveness so that funding will be efficiently distributed.

References

Allen IV, James, Arlete Mahumane, James Riddell IV, Tanya Rosenblat, Dean Yang, and Hang Yu. 2021. Correcting perceived social distancing norms to combat COVID-19. Technical report National Bureau of Economic Research. http://www.nber.org/papers/w28651.

Barrios, John M, Efraim Benmelech, Yael V Hochberg, Paola Sapienza, and Luigi Zingales. 2021. "Civic capital and social distancing during the Covid-19 pandemic." *Journal of public economics* 193: 104310.

Bauman, Karl E, John LaPrelle, Jane D Brown, Gary G Koch, and Connie A Padgett. 1991. "The influence of three mass media campaigns on variables related to adolescent cigarette smoking: results of a field experiment." *American Journal of Public Health* 81 (5): 597–604.

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse. 2021. ""Rugged individualism" and collective (in) action during the COVID-19 pandemic." *Journal of Public Economics* 195: 104357.

Breza, Emily, Fatima Cody Stanford, Marcella Alsan, Burak Alsan, Abhijit Banerjee, Arun G Chandrasekhar, Sarah Eichmeyer, Traci Glushko, Paul Goldsmith-Pinkham, Kelly Holland, Emily Hoppe, Mohit Karnani, Sarah Liegl, Tristan Loisel, Lucy Ogbu-Nwobodo, Benjamin A. Olken, Carlos Torres, Pierre-Luc Vautrey, Erica Warner, Susan Wootton, and Esther Duflo. 2021a. "Doctors' and Nurses' Social Media Ads Reduced Holiday Travel and COVID-19 Infections: A Cluster Randomized Controlled Trial". Technical report National Bureau of Economic Research. http://www.nber.org/papers/w29020.

Breza, Emily, Fatima Cody Stanford, Marcella Alsan, Burak Alsan, Abhijit Banerjee, Arun G Chandrasekhar, Sarah Eichmeyer, Traci Glushko, Paul Goldsmith-Pinkham, Kelly Holland, Emily Hoppe, Mohit Karnani, Sarah Liegl, Tristan Loisel, Lucy Ogbu-Nwobodo, Benjamin A. Olken, Carlos Torres, Pierre-Luc Vautrey, Erica Warner, Susan Wootton, and Esther Duflo. 2021b. "Effects of a large-scale social media advertising campaign on holiday travel and COVID-19 infections: a cluster randomized controlled trial." *Nature medicine* 27 (9): 1622–1628.

Broockman, David E, and Donald P Green. 2014. "Do online advertisements increase political candidates' name recognition or favorability? Evidence from randomized field experiments." *Political Behavior* 36 (2): 263–289.

Bueno de Mesquita, Ethan, and Mehdi Shadmehr. 2023. "Social norms and social change." *Quarterly Journal of Political Science* 18 (3): 339–363.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott. 2020. "Misperceived social norms: Women working outside the home in saudi arabia." *American economic review* 110 (10): 2997–3029.

Byker, Tanya, Caitlin Myers, and Maura Graff. 2019. "Can a social media campaign increase the use of long-acting reversible contraception? Evidence from a cluster randomized control trial using Facebook." *Contraception* 100 (2): 116–122.

Campos-Mercade, Pol, Armando N Meier, Florian H Schneider, and Erik Wengström. 2021. "Prosociality predicts health behaviors during the COVID-19 pandemic." *Journal of public economics* 195: 104367.

Cantoni, Davide, David Y Yang, Noam Yuchtman, and Y Jane Zhang. 2019. "Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement." *The Quarterly Journal of Economics* 134 (2): 1021–1077.

Durante, Ruben, Luigi Guiso, and Giorgio Gulino. 2021. "Asocial capital: Civic culture and social distancing during COVID-19." *Journal of public economics* 194: 104342.

Green, Donald P., Adam Zelizer, and David Kirby. 2018. "Publicizing Scandal: Results from Five Field Experiments." *Quarterly Journal of Political Science* 13 (3): 237-261.

Green, Donald P, Adam Zelizer, and Winston Lin. 2021. "How Effective Are Radio Messages Aimed at Reducing Teen Births Among Latinas? A Randomized Controlled Trial." *Population Research and Policy Review* 40 (4): 695–722.

Hornik, Robert, Lela Jacobsohn, Robert Orwin, Andrea Piesse, and Graham Kalton. 2008. "Effects of the national youth anti-drug media campaign on youths." *American Journal of Public Health* 98 (12): 2229–2236.

Larsen, Bradley J., Timothy J. Ryan, Steven Greene, Marc J. Hetherington, Rahsaan Maxwell, and Steven Tadelis. 2023. "Counter-stereotypical messaging and partisan cues: Moving the needle on vaccines in a polarized United States." *Science Advances* 9 (29): eadg9434.

Lin, Tong, and Xiaoli Nan. 2022. "A Scoping Review of Emerging COVID-19 Health Communication Research in Communication and Media Journals." *Health Communication*: 1–12.

Longshore, Douglas, Bonnie Ghosh-Dastidar, and Phyllis L Ellickson. 2006. "National Youth Anti-Drug Media Campaign and school-based drug prevention: Evidence for a synergistic effect in ALERT Plus." *Addictive behaviors* 31 (3): 496–508.

Moehring, Alex, Avinash Collis, Kiran Garimella, M Amin Rahimian, Sinan Aral, and Dean Eckles. 2023. "Providing normative information increases intentions to accept a COVID-19 vaccine." *Nature Communications* 14 (1): 126.

Richardson, Amanda, Jennifer Cullen, Paul Mowery, Kristen McCausland, and Donna Vallone. 2011. "The path to quit: how awareness of a large-scale mass-media smoking cessation campaign promotes quit attempts." *Nicotine & Tobacco Research* 13 (11): 1098–1105.

Sly, David F, Gary R Heald, and Sarah Ray. 2001. "The Florida "truth" anti-tobacco media evaluation: design, first year results, and implications for planning future state media evaluations." *Tobacco control* 10 (1): 9–15.

Wakefield, Melanie A, Barbara Loken, and Robert C Hornik. 2010. "Use of mass media campaigns to change health behaviour." *The Lancet* 376 (9748): 1261–1271.

Wakefield, Melanie, Yvonne Terry-McElrath, Sherry Emery, Henry Saffer, Frank J Chaloupka, Glen Szczypka, Brian Flay, Patrick M O'Malley, and Lloyd D Johnston. 2006. "Effect of televised, tobacco company–funded smoking prevention advertising on youth smoking-related beliefs, intentions, and behavior." *American Journal of Public Health* 96 (12): 2154–2160.

Wan, William, and Brittany Shammas. 2020. "Why health officials are terrified of a pandemic Christmas." *The Washington Post*.

Wust, MaryKate. 2020. "Safety During the Holiday Season: How to Modify Your COVID-19 Celebrations." *Penn Medicine News*.

The Minimal Effects of Public Health Campaigns on Travel During the COVID-19 Pandemic

Supplemental Information is intended for online publication only

Contents

A. Treatment Mailers in Publicity Condition

Figure A1: Mailer with 50% Dosage Publicity

Figure A2: Mailer with 100% Dosage Publicity

B. Demographic Summary of All Treatment Conditions

Table B1: **Balance of Control and Treated Census Block Groups on Key Covariates, all Treatment Conditions**

	Control	50% Private	100% Private	50% Public	100% Public
% Female	53.1	53.5	53.0	52.9	53.4
($\hat{SE}$)	(0.1)	(0.3)	(0.3)	(0.3)	(0.3)
Mean Age (in Years)	49.7	49.4	50.3	49.6	50.8*
	(0.1)	(0.4)	(0.4)	(0.4)	(0.4)
Mean Income ($ in '000s)	71.4	74.2	82.7	69.3	75.4
	(1.1)	(6.6)	(6.7)	(5.2)	(6.2)
% African American	29.6	30.7	23.6	29.6	27.6
	(0.6)	(3.5)	(3.3)	(3.7)	(3.6)
% Hispanic	22.5	20.3	20.9	21.1	20.3
	(0.4)	(2.3)	(2.0)	(2.4)	(2.3)
N CBGs	3200	100	100	100	100
N Individuals (in '000s)	3,177.5	99.0	97.5	97.2	98.1

Two-tailed p-values indicated at p < .05 (*) and p < .01 (**).

C. Estimated Treatment Effects: Alternative Outcomes

Table C1: **Estimated Effects of Treatment on Social Distancing: Cell Phone Mobility**

	Any Treatment	50% Private	100% Private	50% Public	100% Public
Median Distance Travelled (2.8 km)					
	0.4	-1.4	1.4	1.0	0.8
	(0.8)	(0.4)	(1.4)	(1.4)	(2.1)

Two-tailed p-values indicated at $p < .05$ (*) and $p < .01$ (**). More conservative of two-way clustered and CBG-clustered standard errors and p-values displayed.

D. Estimated Treatment Effects: Spillover Effects on Untreated Households in Treated CBGs

Table D1: **Estimated Spillover Effects of Treatment on Social Distancing: Survey**

Est. Effects on Untreated Households in condition...	50%-Private	50%-Public
Holiday Shopping Online (vs. In-Person) $\widehat{\text{ATE}}$	3.9	2.2
($\widehat{\text{SE}}$)	(3.6)	(4.3)
Visited Bar or Restaurant	-0.7	0.7
	(2.4)	(2.6)
Visited House of Friend or Extended Family Member	-6.5*	1.3
	(3.0)	(3.8)
Wear a Mask in Public	1.0	-1.9
	(1.8)	(2.0)
Celebrate Holidays in Group	-1.2	-1.4
	(1.7)	(1.8)
See Neighbors Gathering for Holidays	0.9	-2.5
	(2.4)	(2.0)
Recall Mailer	-0.9	0.8
	(2.4)	(2.8)

Two-tailed p-values indicated at $p < .05$ (*) and $p < .01$ (**). Standard errors and associated p-values clustered at CBG-level.