Elite Cues and Mass Non-compliance*

Zachary P Dickson[§] & **Sara B Hobolt**[†]

[§] London School of Economics z.dickson@lse.ac.uk
[†] London School of Economics s.b.hobolt@lse.ac.uk

Abstract

Elite cues have been shown to influence public attitudes and behaviors. But to what extent can political elites encourage mass non-compliance simply by naming the geographic location of a subset of the population? In this paper, we study the effects of elite cues on non-compliant behavior during the COVID-19 pandemic, focusing on a series of controversial tweets sent by US President Donald Trump calling for the "liberation" of Minnesota, Virginia and Michigan from COVID-19 restrictions at the height of the pandemic. Leveraging the fact that these messages referred to three specific US states, we adopt a generalized difference-in-difference approach relying on spatial variation to identify the causal effects of the targeted cues. Our analysis finds that the President's messages inspired mass non-compliance in the forms of an increase in public mobility, a decrease in adherence to stay-at-home restrictions, and an increase in criminal activity in the days following. Treatment effect heterogeneity further demonstrates Trump's outsized influence on mobility in Republican counties and on criminal activity among White individuals. Our findings have broad implications for our understanding of the capacity of elites to motivate mass non-compliant behavior.

Keywords— Elite cues, political behavior, COVID, Trump, crime, social media

*__Acknowledgements:__ *The authors are grateful for the generous support from the Volkswagen Foundation via The COVIDEU Project grant.*

1 Introduction

In moments of acute crisis, political elites rely on high levels of cooperation from citizens to manage the potential negative consequences. Governments typically take strong executive action to achieve specific public policy ends – such as the restoration of public order or the management of imminent physical risk – and they depend on the public to comply with their policies and follow their guidance. Yet, in a polarized political environment, can elite messages have the opposite effect and lead to greater public non-compliance?

During the COVID-19 pandemic, studies have shown that consistent and unified government messaging and public trust in governments led to higher levels of compliance with health-related measures (Jørgensen, Bor, and Petersen 2021; Jørgensen et al. 2021; Anderson and Hobolt 2022; Bargain and Aminjonov 2020). These findings are in line with a rich body of literature demonstrating that elite cues can have significant effects on citizens' political behavior and attitudes (Lupia and McCubbins 1998; Samuels and Zucco 2014; Zaller and Feldman 1992; Lupia 1994; Brader and Tucker 2012). A consistent finding is that citizens tend to follow the cues of their preferred party or politicians when political elites are polarized. In the context of the COVID-19 pandemic, evidence from the United States suggests that different positions taken by Democratic and Republican party leaders on the threat posed by COVID-19 and the appropriate crisis response led to stark differences in both support for and compliance with COVID-related measures (Gadarian, Goodman, and Pepinsky 2022; Allcott et al. 2020; Grossman et al. 2020; Roberts and Utych 2021; Bisbee and Lee 2022; Green et al. 2020). Just as consistent elite messaging and public trust have been seen to lead to higher levels of public compliance, it has been argued that the partisan polarization of the issue in the United States, led by the Trump administration, was a contributing factor leading to low levels of public compliance with public health measures and higher numbers of COVID-19 cases and associated deaths (Gadarian, Goodman, and Pepinsky

2022; Hamilton and Safford 2021).

This raises the question of whether specific elite messages can change people's behavior and encourage non-compliance, even to the extent that people would break the law during an acute crisis. There is a growing body of evidence suggesting that Republicans and Democrats behaved differently during the crisis (Gadarian, Goodman, and Pepinsky 2022; Hamilton and Safford 2021; Roberts and Utych 2021), yet it is more difficult to pinpoint whether specific elite messages caused differences in behavior. In this article, we are interested in whether the polarizing messages by President Trump caused Republicans to be less compliant with COVID-related social distancing measures, and the degree to which non-compliant behavior transformed into more extreme activities in opposition to government interventions.

To test this, we analyze the effect of a series of controversial tweets sent by President Trump calling for the "liberation" of Minnesota, Virginia and Michigan from COVID-19 restrictions at the height of the first wave of the pandemic in 2020 when stay-at-home orders were in place across the US. We leverage the fact that these messages referred to specific states, as this allows us to adopt a spatial difference-in-difference approach to estimate the causal effects of elite cues on non-compliance behavior among in-group partisans in targeted areas. We can thus differentiate between the effect of the messages in targeted and non-targeted states, and between predominantly Republican (red) and predominantly Democrat (blue) counties.

In support of our expectations, we find that Trump's calls to action increased non-compliant behavior in the form of human mobility in the days following the tweets. We document an increase in mobility in the Trump-targeted Republican counties even in relation solely to Republican counties elsewhere around the country, illustrating the localized and exclusive effects of the cues. This is in line with recent work that has demonstrated an effect of elite cues (tweets) on mobility during the crisis (Grossman et al. 2020; Bisbee and Lee 2022). We then turn attention to the spill-over effects of the polarizing cues, demonstrating that Trump's tweets were not only associated with anti-government rebellion and violent rhetoric on Twitter, they were also followed by an increase in crime in the targeted states. We document this increase exclusively among White individuals, illustrating the partisan and heterogeneous effects of President Trump's calls to action.

This paper thus contributes to the literature on elite cues in three important ways. First, we examine the causal effect of elite cues on behavior in a crisis by employing a spatial difference-in-difference approach to compare behaviors among in-group partisans in targeted areas with in-group partisans in non-targeted areas. Second, we focus on actual observed behaviors of citizens during the crisis, rather than self-reported survey responses. To capture non-compliance during the pandemic, we focus on localized data on social mobility when stay-at-home orders were in place.

Finally, after modeling the key topics using responses to President Trump's messages on social media using hierarchical topic modeling based on the transformers architecture, we provide a more nuanced picture of how these messages were received. Our analysis reveals a strong association with anti-government, radical and violent rhetoric. This link between the tweets and radical non-compliant behavior then guides our further analysis of criminal behavior, which we find increased in the ten-days following the cues in the three targeted states and for White individuals only. The findings thus illustrate the substantive effects that polarizing elite messages on social media can have on real world behavior, even when such behavior is potentially costly to the individual. Worryingly, they also imply that such elite messages can lead to more radical forms of non-compliance, such as those witnessed on January 6, 2021 at the United States Capitol. Our findings therefore have broader implications for how elites can undermine compliance with and respect for local institutions and the rule of law.

2 Elite Cues and Public Compliance

"Trump's practice of charismatic populism portrayed him as uniquely knowledgeable, with a particular authority that other politicians and health leaders lacked...he demanded the media spotlight" write Gadarian, Goodman and Pepinsky in their authoritative account of the politics of the pandemic in the US, concluding "Trump's decisions made the pandemic worse" (Gadarian, Goodman, and Pepinsky 2022, 273–74). Studies have argued that Trump's leadership worsened the outcome of the pandemic in the US in a number of ways, including encouraging less social distancing (Roberts and Utych 2021; Grossman et al. 2020; Bisbee and Lee 2022), reducing mask wearing (Hahn 2021), and undermining trust in science agencies (Hamilton and Safford 2021; Gadarian, Goodman, and Pepinsky 2022).

While there is no doubt that Trump was a highly unusual leader and conspicuous media presence, these claims about how he shaped pandemic outcomes raise broader questions about the ways in which political leaders can affect outcomes in a crisis. In this paper, we are not focusing on the effect of policy choices, but more specifically on the extent to which elite messaging can influence the behavior of citizens. Particularly, we are interested in identifying the causal effect of specific messages opposing COVID-related restrictions on citizens' adherence to such restrictions and, in turn, on more radical instances of non-compliance. We are thus presenting a hard test case of whether elite cueing is effective, since we are examining not whether a politician can engender support for their policies, but rather whether they can encourage citizens to act in ways that go against the official rules and guidance, even when there are potential costs associated with breaking the rules, including personal health risks.

There is a large body of literature demonstrating that elite cues can have significant effects on citizens' behavior and attitudes, as well as their support for public policies (Tesler 2012; Zaller and Feldman 1992). Messages from political actors are among the most widely available and influential information shortcuts in politics and individuals respond to cues based on their perceived credibility and trustworthiness (Arceneaux 2008; Arceneaux and Johnson 2013; Lupia 1994; Lupia and McCubbins 1998; Leeper and Slothuus 2014). In a polarized political context, citizens will often interpret cues from the perspective of in-groups and out-groups. Specifically, the literature shows that citizens tend to follow the cues of their preferred party or politicians (Nicholson 2012; Brader and Tucker 2012; Samuels and Zucco 2014). In the US, partisan identities are powerful social identities that provide a lens through which people observe the world (Green, Palmquist, and Schickler 2004; Iyengar and Simon 2000). Partisan cues thus shape how citizens perceive policies and the political world and have been shown to activate partisan biases even on traditionally nonpartisan issues (Druckman 2001; Kam 2005).

In what ways did elite cues matter during the pandemic? We might expect elite cues to be particularly important in moments of heightened uncertainty as was the case in early-2020 at the start of the pandemic. As with any crisis situation, the pandemic presented citizens with the daunting challenge of navigating a new, complex and changing information environment. In crises in general, citizens will often "rally 'round the flag" of the leader, and evidence from the first phase of the pandemic suggests that citizens around the world became more supportive of and receptive to their political leaders (De Vries et al. 2021; Baekgaard et al. 2020; Bol et al. 2021; Lupu and Zechmeister 2021). However, unlike much of the world – where mainstream politicians sought to present a united front in response to the pandemic (Barari et al. 2020; Hensel et al. 2022; Anderson and Hobolt 2022; De Vries et al. 2021) – the response in the US was deeply politicized and polarized along partisan lines, with different positions taken by Democratic and Republican party leaders both on the threat posed by COVID-19 and the appropriate response (Allcott et al. 2020; Roberts and Utych 2021; Bisbee and Lee

2022). For example, Green et al. (2020) analyze the rhetoric surrounding COVID-19 by Congress members and show that while Democrats highlighted the public health threat, Republicans placed greater emphasis on China and businesses. Likewise, Gadarian, Goodman, and Pepinsky (2022) describe President Trump's response to the pandemic as particularly polarizing and divisive in their comprehensive reflection of the influence of COVID-19 on the United States.

In such a polarized environment, we would expect the effect of elite cues to be conditioned by the partisanship of the receiver. In other words, we would expect Republican partisans to be more receptive to the messages of Republican politicians, such as Donald Trump, whereas Democrat partisans will be less receptive and may even shift their opinion in the opposite direction. Indeed, studies have shown that that Republican partisans were generally less supportive of COVID-19 measures and less likely to comply (Allcott et al. 2020; Gadarian, Goodman, and Pepinsky 2022). Survey evidence shows that Democrats are typically more likely to see COVID-19 as a major threat and more supportive than Republicans in their stated support of and willingness to comply with these measures (Van Green and Tyson 2020; Gadarian, Goodman, and Pepinsky 2022). Moreover, more Republican counties typically display lower levels of compliance with social distancing measures than Democratic counties (Roberts and Utych 2021; Bisbee and Lee 2022)

Yet, while the evidence reveal clear differences in partisan attitudes and behaviors in the US during the crisis, it is challenging to examine empirically whether these differences are due to elite rhetoric during the pandemic. Some studies have made important contributions to examining the role of elite cues during the pandemic. Bursztyn et al. (2020) use county-level variation in television consumption of two Fox News programs—Sean Hannity and Tucker Carlson—and find that differences in viewership predict differences in COVID-related health outcomes. Two other studies examine the effect of elite cues on compliance of social distancing rules more directly. Grossman et al. (2020) shows that a governor's tweets encouraging social distancing have a meaningful impact on social distancing behaviors, and the effect is larger in Democrat-leaning counties. Similarly, Bisbee and Lee (2022) reveal that the partisan gap in Americans' social distancing behaviors is exaggerated by President Trump's pronouncements on the seriousness of the virus. They leverage changes in Trump's evaluation of the pandemic revealed in his tweets to show an increase in mobility in Republican-leaning counties when Trump issues anti-lockdown tweets (including the liberate tweets, examined in this paper).

2.1 Trump's Messages

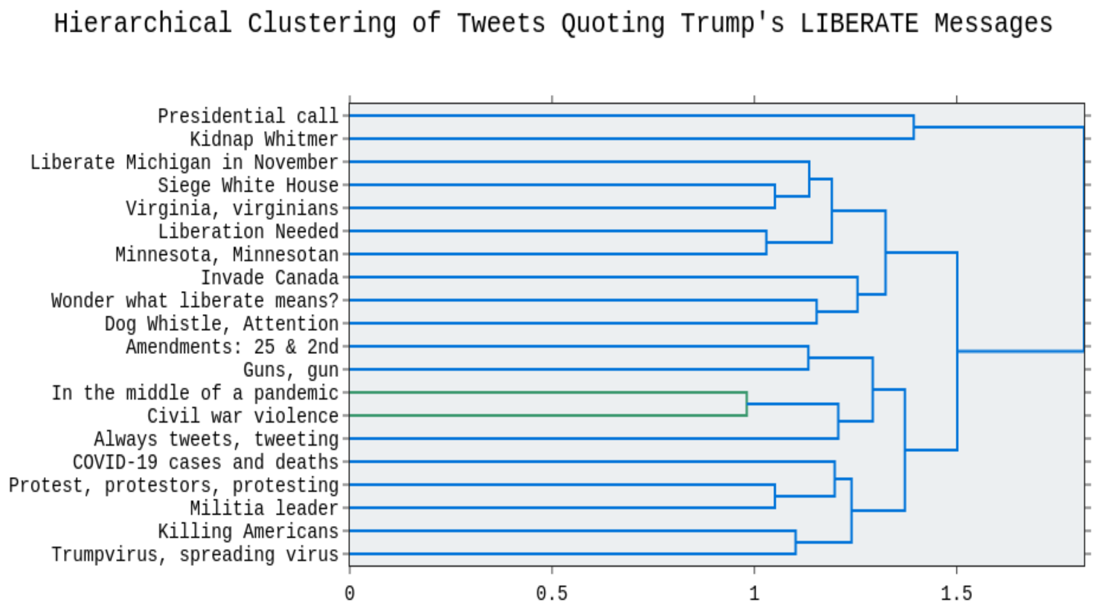
We build on these studies, and the wider literature on elite cueing, to examine the localized effect of elite rhetoric. We leverage the fact that Trump's liberate tweets referred to three specific US states, so we can adopt a spatial difference-in-difference approach to examine the causal effect on compliance in those states compared to non-targeted states. Specifically, our focus is on President Trump's calls for the "liberation" of Michigan, Minnesota and Virginia at the height of the initial outbreak of COVID-19. On April 17th, 2020, Trump broadcast three separate messages to his 80+ million Twitter followers that read as follows: "LIBERATE MICHIGAN", "LIBERATE MINNESOTA", and "LIBERATE VIRGINIA" (Collins and Zadrozny 2020).[1] At that point in time, and in the surrounding days, each of the three states targeted by Trump were under stay-at-home mandates from state governments in order to slow the spread of the COVID-19 virus.

How were these messages interpreted by the media and the public? According to media reporting at the time, Trump's calls for liberation were widely seen as encouraging citizens to disobey the stay-at-home orders in place, and even as an incitement of

1. President Trump's full message to Virginia was "LIBERATE VIRGINIA, and save your great 2nd Amendment. It is under siege!"

violence and rebellion (Collins and Zadrozny 2020; Fallows 2020). For example, former Assistant Attorney General for National Security Mary McCord argued that "it's not at all unreasonable to consider Trump's tweets about "liberation" as at least tacit encouragement to citizens to take up arms against duly elected state officials of the party opposite his own" (McCord 2020). We can also observe the ways in which the messages were received by examining the replies of individuals who engaged with the calls for liberation on Twitter. Hundreds of thousands of Twitter users liked, shared and replied to the liberate messages. To explore the nature of the responses, we created hierarchical topic models using every message that "quote-tweeted" one of the three liberate messages.[2]

Figure 1: Hierarchical Topic Model of LIBERATE Quote Tweets



2. Our topic models included 143,171 quote tweets. Details of the topic models, including text pre-processing steps and the libraries used are available in Appendix A. Quote tweets are Twitter messages that re-share an original message along with new commentary from the user. Quote tweets allow users to re-frame, dissent, or express their approval of a message.

Figure 1 presents the top-20 topics and their hierarchical groupings.[3] Several of the topics appear to identify the President's messages as calls for protest and even violence, calling for "Liberation Needed", "Guns" and "Protest."[4] Interestingly, however, several of the topics appear to consist of messages that are in opposition to the President's calls for liberation. For example, the topics "Dog Whistle, Attention" and "Trumpvirus, spreading virus" are not likely to originate from individuals who agree with the sentiment of the president's messages. Rather, they are likely to be condemnations given anticipated follow-on effects from the messages.

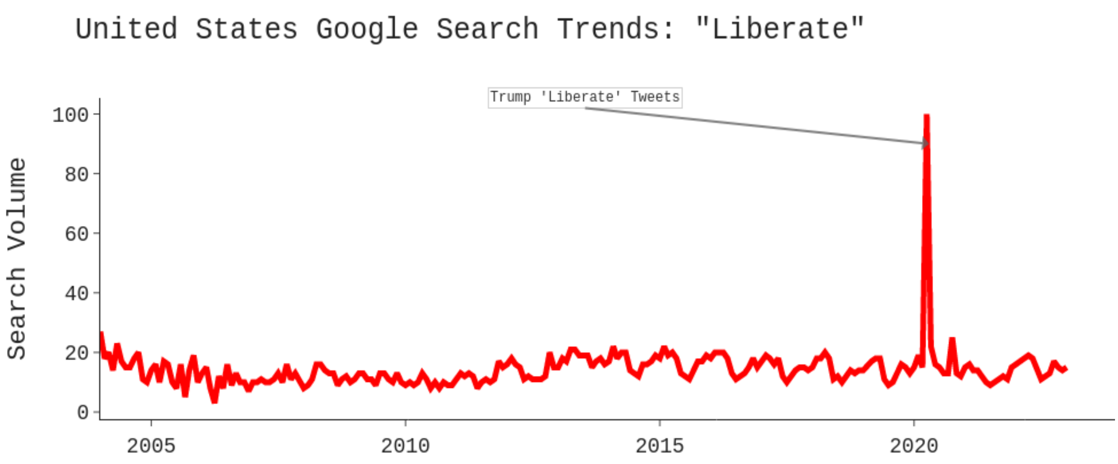
Twitter is not representative of the entire US population and Twitter discourse can differ from that of the public. We know, however, that the President's messages were widely covered by prominent media outlets such as *NBC News*, *The Washington Post* and *The Atlantic* (Collins and Zadrozny 2020; Fallows 2020; McCord 2020). Moreover, we analyzed Internet trend data based on Google search history to gain further confirmation that the President's messages were widely received by the public. We examined the search history of the word *liberate* using Google Trends (Google 2020) as presented in Figure 2.[5] The data indicate that at no other point in the nearly twenty-year history during which Google has tracked search data was the term "liberate" searched more frequently than in April 2020.

3. The topics are interpreted but the exact topic words, along with additional word representations and further visualizations are provided in Appendix A.

4. "Kidnap Witmer" is especially damning, considering that several members of the far-right Wolverine Watchmen militia were later arrested and convicted of various terrorism charges for planning to kidnap Michigan Governor Gretchen Witmer and violently overthrow the state government because of 'tyrannical COVID-19 restrictions' (Smith 2022).

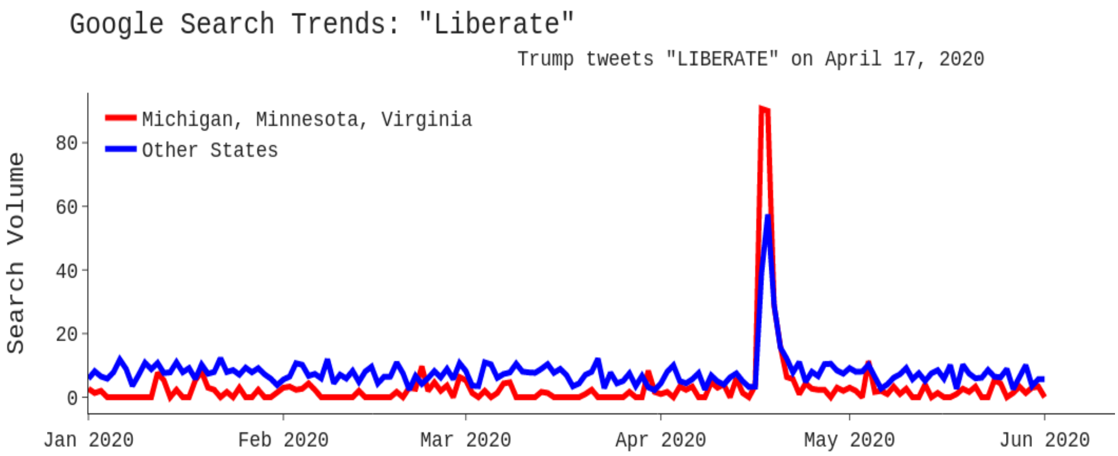
5. Google Trends data are normalized and scaled according to time period and geography in order to represent the relative popularity of a search term on a range between 0 and 100 (Google 2020).

Figure 2: US Historical Google searches for 'liberate'



Google Trends data further allow for segmenting the data by geographic location and allow for daily search history when narrowing down the time frame to include only the first five months of 2020. We present those data in Figure 3, after separating Minnesota, Virginia and Michigan from the remaining 47 US states. From the figure, we can see that daily search trends increased dramatically the day that Trump sent the messages. Search trends confirm that indeed the messages had a national effect, but that the effect was considerably more concentrated in the three states that were mentioned explicitly by the president.

Figure 3: Google Search Trends for "Liberate"



Given that Trump's liberate tweets targeted three specific states, and were so widely read and commented upon, we can examine the causal effect on compliance in those states compared to non-targeted states. In line with the elite cueing literature, our expectation is that non-compliance increased in the targeted states, but only among those individual receptive to the in-group messages of Trump, namely Republican partisans. This leads us to the following hypotheses:

> *H1:* Individuals in states targeted by Trump's messages are less compliant with COVID-19 stay-at-home orders in the days following the tweets than individuals in states that were not targeted in the messages.

> *H1a:* The effects of Trump's messages on non-compliance are observed in predominantly Republican counties.

We go one step further to examine the effect of the tweets not just on compliance with social distancing measures, but also with more generalized non-compliant behavior, i.e. criminal activities more broadly. Evidence suggests that COVID-19 crime rates fell in the first phase of the pandemic, mainly attributed to the stay-at-home-orders in place that led to a drop in the type of minor offences which are typically committed in the community in peer groups (Stickle and Felson 2020; Boman and Gallupe 2020). Studies in criminology have suggested that the lockdowns altered the social dynamics often associated with minor offending, as individuals (often young males) had less access to the peer groups in which criminal behavior often occurs (Boman and Gallupe 2020; Lopez and Rosenfeld 2021).

We would thus expect that if Trump's tweet encouraged people not to comply with the stay-at-home orders, this could also spill over into other criminal activities – such as disorderly conduct, vandalism, destruction of property and assault – in the targeted states. Moreover, since we have shown in our analysis of responses to the tweets that the cues were generally associated with disaffection with state and local authorities and institutions, with clear indications of non-compliant and even violent behavior, this may also have triggered a greater propensity for criminal behavior among some individuals. Again, we expect the effect of such non-compliant behavior only among individuals most receptive to Trump's messages, i.e. Republican partisans and Trump supporters. Since we do not have individual-level partisanship or county-level data for criminal activity (see below), we use another proxy for partisanship available in the data, namely race. This is a crude measure, of course, but the literature consistently shows that whites are more likely to be Republican partisan and Trump supporters compared to non-Whites (Sides, Tesler, and Vavreck 2017). For example, a Pew Research Center study shows that only 6% of black voters and 28% of Hispanic voters supported Trump in 2016 compared to 54% of whites (and 62% of white men) (Doherty, Kiley, and Johnson 2018). In other words, non-white voters in general are unlikely to be as receptive to Trump's cues, and we would therefore expect a heterogeneous treatment effect where the targeted messages of Trump matters most for crime rates among white Americans. This leads to our final hypothesis:

> *H2:* Individuals in states targeted by Trump's messages are more likely to commit crimes in the days following the tweets than individuals in states that were not targeted in the messages. This effect is likely to be most pronounced for Whites compared to Non-Whites.

3 Research Design

3.1 Data and Variables

To estimate the effects of Trump's messages on mass non-compliance, we focus on three different outcomes of non-compliance measured daily: mobility, stay-at-home compliance, and criminal behavior. Daily mobility and stay-at-home compliance data are available at the county level and are based on mobile phone locations. Our primary source of mobility comes from Meta's (previously Facebook) Data for Good project (Meta 2022). The movement range data tracks daily movement through the Facebook application and were released to researchers and public health experts to aid in combating the spread of COVID-19. There are two types of data available from Meta: movement range data and 'stay put' data. The movement range data measures how much people move around in a given area. The stay put data is calculated using the fraction of the population that remains in a single location for the entire day.[6] Both measurements of mobility are calculated using the average values from the month before lockdowns and restrictions were put in place (February 2020). Meta's mobility data are especially valuable for our analysis, because in combination they provides measures of both the extent to which individuals traveled in a given county, as well as the percent

6. More on the methodology of the mobility data is available in this blog from Meta Research.

of the population that remained in a single location for the day. We refer to the latter of the two as compliance with stay-at-home measures.

For measures of criminal activity, we rely on arrest data from the FBI's National Incident-Based Reporting System (NIBRS) (FBI 2022). The data are available daily for US states and include information on the type of crime for which the arrest was made, as well as demographic characteristics of the offender. 45 US states (and the District of Columbia) reported arrests in 2020 to NIBRS, including the three states we are interested in that were targeted by the President's messages.

We specifically focus on four crimes that are linked to non-compliance with stay-at-home orders (Stickle and Felson 2020; Boman and Gallupe 2020) as well as any disorderly activity in response to the President's calls for liberation: Assault (simple and aggravated), Disorderly Conduct, and Destruction/Damage/Vandalism of Property. We present descriptive statistics for arrests for these crimes in Appendix G.

3.2 Identification and Empirical Strategy

The focus of our analysis is the extent to which President Trump's cues motivated non-compliant behavior. At the time of the Trump's messages, the majority of states around the country had issued stay-at-home mandates to curb the spread of COVID-19. Our identification strategy exploits the spatially targeted nature of the calls for liberation. Because Trump's messages targeted Minnesota, Virginia and Michigan and only these states were mentioned in the messages, we treat the three states as the treatment group, while considering remaining US states as the control group.

Although the cues were directed specifically to the citizens of Michigan, Virginia and Minnesota, the President's messages were seen widely, which is evidenced by the widespread national media coverage and the dramatic increase in online search behavior in the rest of the country, shown above. Therefore, our identification strategy ensures that our estimates capture the effect of the *targeted cues* rather than the effect of cues more broadly. For this reason, our analysis likely underestimates the effect of the cues in totality because individuals were still exposed to the treatment even if their locations were not targeted specifically.

We rely on a generalized difference-in-difference estimation strategy. The key identifying assumption for our strategy is that the President's decision to target Minnesota, Michigan and Virginia was not based on characteristics of those states that make them more or less susceptible to his messages. In other words, the assumption made is that the treated and control units would follow parallel paths in the outcomes we observe and in the absence of the treatment.[7] The main threat to identification is that President Trump was responding to events that were already occurring or other state-level characteristics. If this were to be the case, then such characteristics could confound the relationship between the treatment and our outcomes of interest. A specific scenario that threatens our identification is that Trump targeted the three states because of protest activity in those states.[8] Therefore, we ran several additional analyses to check that state-level characteristics, such as anti-lockdown protests, COVID-19 cases, or violent crime were not higher in Michigan, Virginia and Minnesota compared to other states in the period leading up to the tweets, and therefore are unlikely to have influenced Trump's decision to target the states.

We test for different state-level characteristics by attempting to predict the targeted states using COVID-19 conditions (cases and deaths), state-wide protest activity, and violent crime as inputs. This includes regressing a binary indicator identifying the states

7. Formally, for every $g \in \{s, n\}$ and $t \in \{1, 2\}$, let $Y_{g,t}(1)$ and $Y_{g,t}(0)$ denote potential outcomes of group g at period t with and without the treatment, respectively. The parallel trends assumption requires that the untreated expected outcome is the same for both the treated and untreated groups: $E[Y_{s,2}(0) - Y_{s,1}(0)] = E[Y_{n,2}(0) - Y_{n,1}(0)]$.

8. Indeed, we find media coverage suggesting that Trump's calls for liberation came "after small protests by Trump supporters broke out in a handful of states, many of which were fueled by anti-vaccination and anti-government groups" (Collins and Zadrozny 2020).

targeted on each of the potential characteristics Trump could feasibly be responding to around the time of the messages on April 17th. We try a number of specifications[9] using panel regressions and find no evidence that protest activity, COVID-19 characteristics, or violent crime in Minnesota, Virginia or Michigan predict the three states in relation to the rest of the country.[10]

4 Non-compliance – Mobility

To test Hypotheses 1 and 1a, we consider the effects of the cues on mobility, focusing on the time period one week before and after President Trump's messages were sent on April 17th. Our focus is limited to this two-week time period because it was important to include only times during which citizens were under stay-at-home mandates. Even after narrowing our time horizon to this period, there were several states that we do not include in the analysis either because they did not issue (mandatory) stay-at-home orders or because the state's orders concluded before the seven days following April 17th. However, 40 of the US states (and Washington DC) met the inclusion criteria for the analysis. Figure 4 presents each state with the time periods during which citizens were under stay-at-home orders.

9. Our regressions rely on data from Crowd Counting Consortium (CCC) project (Pressman and Chenoweth 2022). The CCC project collects and makes publicly available data on marches, protests, strikes, demonstrations and other political crowd formations. The crowd data is collected by crawling local newspapers and television sites for information indicating political mobilization. More information on the data and the regression results are presented in Appendix B.

10. One likely potential explanation for why Trump targeted each of the three states is that each state was narrowly won by Trump's competitor, Hilary Clinton in the 2016 US Presidential election. Should this be the case, there are no logical reasons to expect that a narrow margin of victory for Trump would confound the relationship between treatment and our outcomes of interest. Moreover, the fact that each of the three targeted states were won by the Democratic nominee for president in 2016 strongly suggests that there are fewer partisans that are receptive to Trump's cues, which means that the effect of the cues are more conservative in the three targeted states compared to states with higher levels of support for the president.

Figure 4: US State Stay-at-Home Orders in 2020

Bars indicate duration of state stay-at-home orders. The dotted line indicates the date at which President Trump called for the liberation of Michigan, Virginia and Minnesota. The shaded red area indicates the time period of focus for the analysis on Mobility. States with missing bars did not issue (mandatory) stay-at-home orders.

4.1 Mobility Results

We first estimate the cumulative effects of the targeted cues at the county level. We include a number of relevant county- and state-level characteristics, including daily county-level COVID-19 cases and deaths.[11] Our model can be formalized using the following equation:

$$Y_{i,j,t} = \alpha + \theta Targeted_{i,j,t} + \delta_t + \zeta_i + X_{i,j,t} + \epsilon_{i,j,t} \tag{1}$$

where Y is mobility in county i in state j at time t. X is a matrix of county-level, time-variant controls such as COVID-19 cases and deaths. $Targeted$ is simply a binary indicator that takes the value of 1 on April 17th onward in counties in targeted states and 0 in all other cases. δ and ζ are fixed effects parameters for time and county, respectively. These parameters absorb time-invariant county-level characteristics that are unobserved, such as differences in employment, income, racial composition, and other fixed county characteristics. The parameter of interest is θ, which captures the effect of the cues on mobility in targeted counties in Minnesota, Virginia and Michigan ($Targeted = 1$) in relation to mobility in the rest of the counties in the country that were under the same stay-at-home conditions after the messages.

Using the above model, we present the results from six specifications. In the first set of results presented in Table 1, we use as the outcome variable change in mobility. These estimations capture the effect of the targeted cues on mobility in the following 7 days in the targeted counties. In the second set of results presented in Table 2, the outcome variable is the fraction of a given county's population that stayed at home for the day (Meta's "stay put" measure). These estimations capture the effect of the targeted cues

11. Covariate balance between treated and control counties, which includes the control variables used in the mobility models, is presented in Appendix C.

on county-level compliance with stay-at-home orders.

In both Table 1 and Table 2, model 1 presents the estimates for the effect of Trump's cues on all counties within Minnesota, Michigan and Virginia. In model 2, we restrict the analysis to include only treated counties in which a majority voted for the Democratic nominee in 2016 (i.e. "Blue counties"). In this model, we use as the control group only counties which also have a Democratic majority in other states around the country and who were also under stay-at-home orders. In model 3, we repeat the previous process, using only "Red" counties for the treated and control groups.

Table 1: Effect of "Liberate" Cues on Mobility

	Mobility		
	(1)	(2)	(3)
Targeted cues	2.284*	1.005	2.706**
	(0.906)	(0.631)	(0.854)
County FEs	✓	✓	✓
Time FEs	✓	✓	✓
Control variables	✓	✓	✓
Num.Obs.	29 064	6132	22 932
R2	0.764	0.825	0.714
R2 Adj.	0.745	0.810	0.692
R2 Within	0.017	0.022	0.018
R2 Within Adj.	0.017	0.019	0.017
RMSE	5.00	4.36	5.12

* $p < 0.05$, ** $p < 0.01$

Standard errors are clustered by state and time and provided in parentheses. Full table results, including estimates for all uninterpreted control variables, are provided in Appendix D.

Table 2: Effect of "Liberate" Cues on Stay-at-home Compliance

	Stay-at-home compliance		
	(1)	(2)	(3)
Targeted cues	-1.128^*	-0.660	-1.336^*
	(0.502)	(0.438)	(0.476)
County FEs	✓	✓	✓
Time FEs	✓	✓	✓
Control variables	✓	✓	✓
Num.Obs.	29 064	6132	22 932
R2	0.883	0.904	0.869
R2 Adj.	0.874	0.896	0.859
R2 Within	0.020	0.022	0.026
R2 Within Adj.	0.020	0.020	0.025
RMSE	1.91	1.89	1.90

* $p < 0.05$

Standard errors are clustered by state and time and provided in parentheses. Full table results including estimates for all control variables are provided in Appendix D.

In Table 1, there is a positive and significant cumulative effect of the cues on mobility in the following seven days in the targeted states (model 1). In essence, this means that mobility increased in counties in Minnesota, Virginia and Michigan from April 17–April 23 in comparison to mobility around the country during the same time period and in counties that were under the same stay-at-home guidelines. In model 2, which looks at only Democratic majority counties, we also see an increase in mobility; however, the effect is much smaller and cannot be reliably estimated. In model 3, which looks at only Republican counties, the magnitude of the effect is the greatest of the three models, indicating that Red counties in targeted states had a greater response to Trump's cues and in line with Hypothesis 1 and 1a.

Table 2 provides the estimates of the effect of the cues on compliance with stay-at-home orders. Similar to the previous results, there is a clear indication that the cues caused a change in the percentage of the population that stayed at home and complied with restrictive measures in the targeted states. Although we cannot accurately estimate the effect of the targeted cues in the Blue counties, similar to the models looking at mobility, the cues had the greatest effect on reducing stay-at-home compliance in Red counties. Taken in combination, the results from Table 1 and Table 2 indicate that President Trump's cues led to both an increase in mobility and a reduction in the population complying with stay-at-home orders in the following week.

4.2 Event Study Analysis

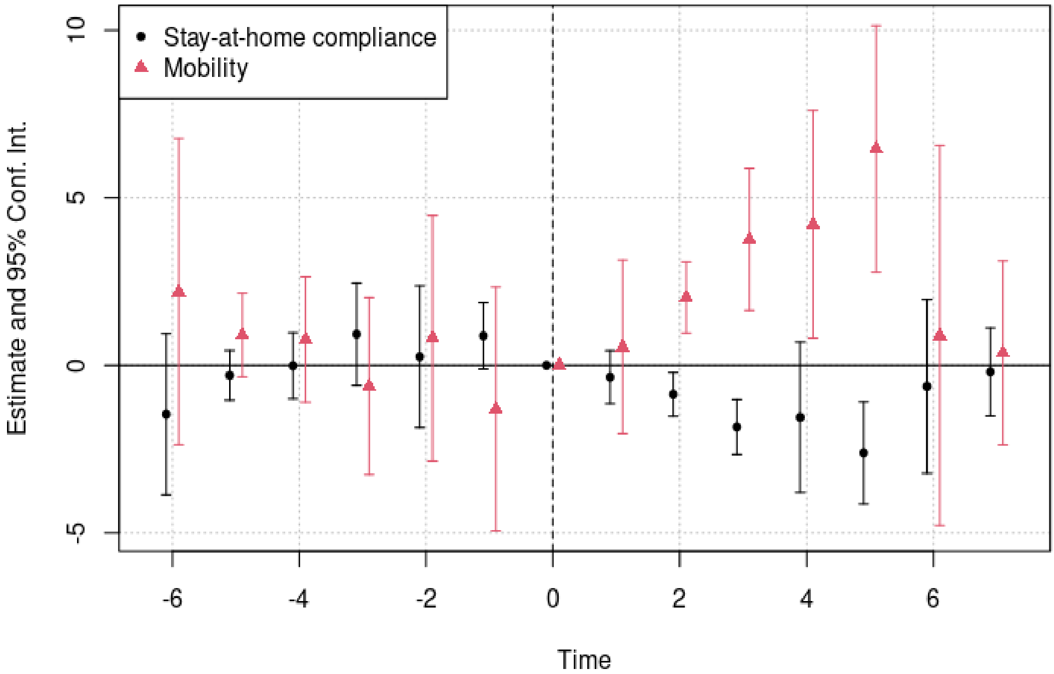
We also conducted our analysis using an event study specification, which has several advantages. First, it allows for visual inspection of pre-treatment parallel trends. Because our identification strategy relies on the assumption that there are parallel trends in mobility between counties in the targeted states (Michigan, Virginia and Minnesota) and the rest of the nation's counties that were under the same restrictions in the time leading up to the Trump's tweets, the event study allows for visual inspection of this assumption. Second, the event study provides a temporal understanding of the ways in which the treatment materialized in the days following. The event study model can be formalized as follows:

$$Y_{i,j,t} = \alpha + \sum_{\tau=-7}^{7} \theta_\tau Targeted_{j,t} + \delta_t + \zeta_i + X_{i,j,t} + \epsilon_{i,j,t} \tag{2}$$

where Y is mobility in county i in state j at time t. τ indicates the leads and lags of the treatment period. In the case that τ is greater than zero, the model captures the dynamic treatment effect of the cues. Whereas when τ is less than zero, the results allow for inspection of a pre-treatment trends between the treated and control counties. In the models, the omitted reference period is $\tau = -1$ (the day before the messages were sent).

Figure 5 presents the daily effects of the cues on mobility and on stay-at-home compliance in the seven days before and after treatment. We use the model 3 specification for each outcome, which estimates the effects of the cues in Republican counties using only Republican counties around the country as the control group. Reassuringly, both outcomes meet the parallel trends assumption in the time leading up to Trump's messages. Following the cues, mobility increases in near-linear fashion for the following few days, peaking on April 21 before returning to similar levels as other Red counties on the 22nd and 23rd. The compliance estimates indicate a similar pattern but in reverse, with compliance sharply decreasing in the following five days before reverting to mean levels on April 22nd and 23rd.

Figure 5: Event Study – Effect of Cues on Mobility



4.3 Robustness – Mobility

We take a number of steps to ensure the robustness of the results of the effects of the elite cues on mobility. Meta is not the only provider of daily mobility data at the county level for the United States, so we first replicate our mobility analysis using data from an alternative source: Google. Google provides Community Mobility Reports data (Google 2020), which estimates daily mobility for US counties according to the type of mobility of the user. For example, data are available for mobility resulting from transit, parks, grocery and pharmacy, workplace, and retail and recreation. Given the aims of our analysis and the desire to estimate the effect of the cues on non-compliant behavior, we focus on mobility associated with retail and recreation. We additionally create an aggregated measurement of mobility by combining all types of the mobility offered in the Google data. Estimations from our replication with the Google mobility data confirms the results of our substantive conclusions in the primary analysis. Full details of our estimation process, the Google Mobility data, and the results are available in Appendix E. Additionally, we also try an alternative estimator to estimate the effect of the cues on mobility for robustness. In Appendix F, we use a first-difference estimator for inference. Results from these estimations similarly confirm the substantive results of our analysis.

5 Non-compliance – Crime

We observe a clear effect from the President's messages on mobility in the subsequent days. The effect is relatively short-lived, as would be expected, and mobility returns to mean levels quickly thereafter. We now consider the effect of the cues on crime, with a similar expectation of a sharp and short-lived increase following the President's messages in the targeted states (Hypothesis 2). As identified by the media and shown in our analysis of interactions with the messages on Twitter, calling for "liberation" has specific connotations and likely inspired non-compliant behavior, especially with local and state authorities.

Following the same research design as previously, we examine the extent to which the President's cues increased non-compliant behavior in the form of crime ten days before and after the messages were sent and in the same states that were under mandatory stay-at-home orders. As discussed, we rely on daily arrest data from NIBRS (FBI 2022) as a proxy for criminal behavior. Arrest data are available according to the crime that is suspected. We focus on four crimes – Assault (simple and Aggravated), Disorderly conduct, and Destruction/Damage/Vandalism of Property – and present descriptive statistics for these crimes in Appendix G.

There are numerous systemic factors that affect crime levels and the extent to which individuals are arrested for the crimes they commit (Wikström 2006). We address time-invariant factors by including fixed-effects for state and time characteristics in our models. Additionally, we include in our models daily COVID-19 cases and deaths, as the pandemic and lockdowns have been shown to affect criminal behavior (Boman and Gallupe 2020). Moreover, as shifting temperatures have been shown to drive crime in the US (Baryshnikova, Davidson, and Wesselbaum 2021), we also include average daily temperature measured at the state level.[12]

Given that we estimate arrest rates at the state level and therefore have relatively fewer control units, we rely on matrix completion methods for inference. Matrix completion methods resemble generalized synthetic control methods (Abadie, Diamond, and Hainmueller 2010, 2015), but generally outperform in a number of configurations (Athey et al. 2021). Matrix completion imputes missing potential outcome data to create control outcomes, against which the causal effects of a binary treatment can be recovered in panel data settings. For estimation, we include the equivalent to fixed effects for state and time (days) and we rely on non-parametric bootstrapping and a five-fold cross-validation procedure to select the optimal λ hyperparameter for inference (Xu and Liu 2022).

We present dynamic estimates for the effect of the cues on arrests of white individuals in Figure 6, given our expectation that they will be more responsive to Trump's cues compared to non-whites, as they are on average more likely to be Republican partisans (Sides, Tesler, and Vavreck 2017). Similar to the dynamic estimates of mobility and stay-

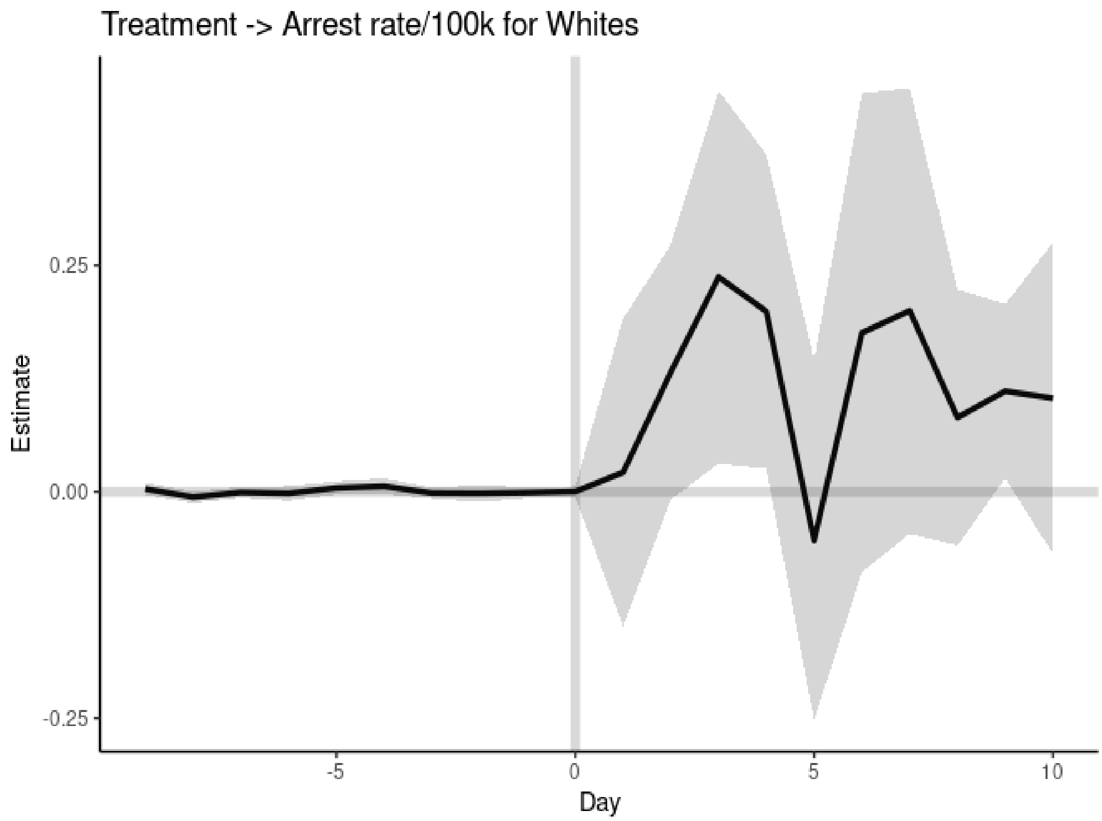
12. Temperature measurements are made at the location of the largest airport in each state.

at-home compliance, there is a rapid and dramatic increase in the two-days following the messages (days 2 and 3). On the fourth day – Monday April 20th – there is a sharp decrease, which is followed by generally higher levels of crime over the remaining week. The cumulative ATT for the ten days following the messages indicate an increase of 12.8 percent[13] [CI 95% 0.015–0.226]. In substantive terms, this amounts to approximately 55 additional arrests of white individuals for crimes related to assault, disorderly conduct and vandalism/destruction of property.[14]

13. The estimated coefficient is 12.1 (also see Appendix H). To interpret the coefficient in percentage terms, we exponentiate the coefficient $e^{\theta} = 1.128$.

14. A one standard deviation increase in arrests amounts to approximately 25 additional arrests. The cummulative effect over the ten days following the messages in substantive terms is ~ 55 additional arrests. This "back-of-the-envelope" calculation is conducted as follows: there were 483 total arrests of whites in the three states and $e^{\theta} = 1.128$, therefore $\sim 55 = 483 - (483/1.128)$. Complete results, including CATT and ATT estimates, are provided in Appendix H.

Figure 6: Treatment Effect: Arrest rate/100k for Whites



ATT estimates for the effect of targeted cues on the arrest rate for white individuals for crimes related to assault, disorderly conduct and vandalism/destruction of property. $\alpha = 0.10$.

5.1 Robustness – Crime

We consider a number of alternative potential scenarios that might explain the increase in crime we observe following the President's liberation tweets. First, it is possible that arrests for crime increased more generally in the time following the messages, rather than being specific to white individuals in the targeted states. We consider this scenario by modeling the same estimations using arrests for the same crimes, but with respect to alternative racial groups. Specifically, we replicated the previous analysis using arrests of Blacks and all arrests at the state level (regardless of race). In these estimations, we similarly used the same compositions for treated and control groups. For example, in the estimations of arrests of Blacks, we use as the control group Black arrests elsewhere in non-targeted US states. The results show no increase in arrests for Blacks in the targeted states, nor for the entirety of the targeted states following the President's messages. These results are presented in Appendix J and Appendix I, respectively.

We also consider whether our results were sensitive to modeling or measurement decisions. First, we examine the possibility that there was simply a sharp, one-day increase in crime by Whites following the President's messages, which accounted for the entire cumulative treatment effect estimate. Therefore, we re-estimate the treatment effects using a two-day arrest average of arrests for the crimes we are interested in. Using panel regressions with fixed state and date effects, we estimate results using both a one-day and two-day strategy. In the regressions, we also consider different combinations of the control variables. The results, presented in Appendix K, indicate some variation between the one- and two-day strategies. In the daily models, the strongest estimate (in terms of magnitude) suggests an increase of 9.7 percent and includes the relevant control variables. In the two-day models, the strongest estimate indicates an increase of 16.8 percent and also includes the same control variables. In all configurations, the treatment effects can be reliably estimated with and without the controls, providing reassurance that our results are not dependent on key assumptions made in the analysis. These results lend strong support to our primary matrix completion strategy, which provided an estimate that essentially splits the difference in calculating an effect size increase of 12.8 percent, providing reassurance that our estimates are not based on the decision to use daily arrest rates or on our specific modeling strategy.

6 Discussion

During crises, such as pandemics, elites play an important role in restoring calm and order and mobilizing support for policy responses. The literature has shown that elite cues can increase support for policies, and that these effects are most pronounced among partisan supporters (Lupia 1994; Brader and Tucker 2012; Nicholson 2012; Anderson and Hobolt 2022; Jørgensen et al. 2021). In this paper, we ask whether political elites can persuade their supporters to *disobey* the rules and guidance put in place in response to a crisis, even at considerable personal risk. Our study examines the effect of elite cues on changes in actual behavior, rather than merely shifts in attitudes, in the context of the polarized and politicized environment of the COVID-19 pandemic in the US.

Specifically, we have analyzed the effects of President Donald Trump's controversial tweets that called for the "liberation" of Minnesota, Virginia and Michigan from COVID restrictions at the height of the pandemic when stay-at-home orders were in place across the US. We show that there was considerable public interest in the tweets, especially in the targeted states. Furthermore, our analysis of the interactions with the tweets on social media reveal associations with protest, disobedience and even violence. Leveraging the fact that Trump's messages referred to specific states, we adopt a spatial difference-in-difference approach to estimate the causal effects of these cues on non-compliant behavior among in-group partisans in targeted areas.

Our findings demonstrate that Trump's calls to action led to higher levels of non-compliant behavior in Republican counties in the targeted states in the days following the tweets: there was a marked increase in mobility in the Trump-targeted Republican counties when comparing with Republican counties around the country, despite the parallel trends in mobility in the days leading up to the controversial tweets. We expand the focus to investigate the spill-over effects of such polarizing elite cues, demonstrating that Trump's tweets were not only associated with anti-government rebellion and violent rhetoric on Twitter, they also resulted in an increase in crime in the targeted states. Our findings show that this increase in criminal activity was only among white individuals and not among the entire targeted states or specific racial groups such as Blacks, which is in line with our expectation that these cues would primarily affect Republican partisans.

These results thus contribute to our understanding of elite cues by demonstrating the effects of elite messages on actual behavior, even behavior associated with potential personal risk. They are in line with recent important studies illustrating the effect of U.S. politicians' social media messages on mobility (Grossman et al. 2020; Bisbee and Lee 2022), but our study also highlights the substantive effects of the specific 'liberation' messages on in-group partisans in targeted states. Moreover, we go one step further in exploring the associations of Trump's messages and acts of rebellion and violence, by analyzing social media responses to the messages as well as showing the spill-over effects on criminal activity in targeted states.

As with any study, there are several limitations to our findings and the degree to which they generalize to other contexts. First, the capacity for elites to motivate non-compliant and/or criminal behavior is likely conditional on a number of other factors that are specific to the US context under President Trump and amidst a pandemic. There is no doubt that Donald Trump is a highly unique politician and communicator who has an unprecedented ability to reach a wide audience with his social media communication (Gadarian, Goodman, and Pepinsky 2022). Moreover, his calls for liberation occurred at a time when state governments had placed extraordinary restrictions on civil liberties, further polarizing the US electorate along political lines and likely aiding Trump's persuasive powers among Republican partisans.

Second, while the effects of Trump's messages on increased mobility and criminal activity are robust, they were also relatively small and short-lived. Yet, this is unsurprising given that the intervention was a single set of tweets and that the comparison is between targeted and non-targeted states. While the increase in non-compliant behavior in response to these tweets may not in and of itself be cause for concern, the great worry is that a sustained campaign by politicians like Donald Trump seeking to undermine respect for rules and regulations among supporters can have even greater effects on non-compliant behavior and further fracture support for and trust in core democratic institutions. As an example, the continuing messaging to undermine trust in the outcome of the 2020 Presidential election has not only shifted attitudes among some Republican partisans, but has also culminated in violent action on January 6, 2021. This study thus contributes to our understanding of how elites can mobilize supporters in opposition to established rules, norms and institutions when operating in a highly polarized context.

Software

Software utilized but not referenced within the main text.

1. Figure 1, Figure 2, Figure 3 and Figure 4: (Plotly 2015)

2. Figure 5: (Berge, Krantz, and McDermott 2021)

3. Figure 6: (Waskom 2021)

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program." *Journal of the American Statistical Association* 105 (490): 493–505.

———. 2015. "Comparative politics and the synthetic control method." *American Journal of Political Science* 59 (2): 495–510.

Aktay, Ahmet, Shailesh Bavadekar, Gwen Cossoul, John Davis, Damien Desfontaines, Alex Fabrikant, Evgeniy Gabrilovich, Krishna Gadepalli, Bryant Gipson, Miguel Guevara, et al. 2020. "Google COVID-19 community mobility reports: anonymization process description (version 1.1)." *arXiv preprint arXiv:2004.04145.*

Allcott, Hunt, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang. 2020. "Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic" [in en]. *Journal of Public Economics* 191 (November): 104254.

Anderson, Christopher J, and Sara B Hobolt. 2022. "Creating compliance in crisis: messages, messengers, and masking up in Britain." *West European Politics* 46 (2): 300–323.

Arceneaux, Kevin. 2008. "Can partisan cues diminish democratic accountability?" *Political Behavior* 30:139–160.

Arceneaux, Kevin, and Martin Johnson. 2013. *Changing Minds or Changing Channels?: Partisan News in an Age of Choice* [in en]. Chicago Studies in American Politics. Chicago, IL: University of Chicago Press, August.

Athey, Susan, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens, and Khashayar Khosravi. 2021. "Matrix completion methods for causal panel data models." *Journal of the American Statistical Association* 116 (536): 1716–1730.

Baekgaard, Martin, Julian Christensen, Jonas Krogh Madsen, Kim Sass Mikkelsen, et al. 2020. "Rallying around the flag in times of COVID-19: Societal lockdown and trust in democratic institutions." *Journal of Behavioral Public Administration* 3 (2).

Barari, Soubhik, Stefano Caria, Antonio Davola, Paolo Falco, Thiemo Fetzer, Stefano Fiorin, Lukas Hensel, Andriy Ivchenko, Jon Jachimowicz, Gary King, et al. 2020. "Evaluating COVID-19 public health messaging in Italy: Self-reported compliance and growing mental health concerns." *MedRxiv,* 2020–03.

Bargain, Olivier, and Ulugbek Aminjonov. 2020. "Trust and compliance to public health policies in times of COVID-19." *Journal of Public Economics* 192:104316.

Baryshnikova, Nadezhda, Shannon Davidson, and Dennis Wesselbaum. 2021. "Do you feel the heat around the corner? The effect of weather on crime." *Empirical Economics,* 1–21.

Berge, Laurent, Sebastian Krantz, and Grant McDermott. 2021. *R Package 'fixest'.*

Bisbee, James, and Diana Da In Lee. 2022. "Objective facts and elite cues: partisan responses to covid-19." *The Journal of Politics* 84 (3): 1278–1291.

Bol, Damien, Marco Giani, André Blais, and Peter John Loewen. 2021. "The effect of COVID-19 lockdowns on political support: Some good news for democracy?" [In en]. Publisher: John Wiley & Sons, Ltd, *European Journal of Political Research* 60, no. 2 (May): 497–505.

Boman, John H, and Owen Gallupe. 2020. "Has COVID-19 changed crime? Crime rates in the United States during the pandemic." *American Journal of Criminal Justice* 45:537–545.

Brader, Ted, and Joshua A. Tucker. 2012. "Following the Party's Lead: Party Cues, Policy Opinion, and the Power of Partisanship in Three Multiparty Systems." *Comparative Politics* 44 (4): 403–420.

Bursztyn, Leonardo, Aakaash Rao, Christopher P Roth, and David H Yanagizawa-Drott. 2020. *Misinformation during a pandemic.* Technical report. National Bureau of Economic Research.

Collins, Ben, and Brandy Zadrozny. 2020. *In Trump's 'LIBERATE' Tweets, Extremists See a Call To Arms. NBC News.*

De Vries, Catherine E., Bert N. Bakker, Sara B. Hobolt, and Kevin Arceneaux. 2021. "Crisis signaling: how Italy's coronavirus lockdown affected incumbent support in other European countries" [in en]. Publisher: Cambridge University Press, *Political Science Research and Methods* 9, no. 3 (July): 451–467.

Doherty, Carroll, Jocelyn Kiley, and Bridget Johnson. 2018. "An examination of the 2016 electorate, based on validated voters." *Pew Research Center.*

Druckman, James N. 2001. “On the Limits of Framing Effects: Who Can Frame?” Publisher: The University of Chicago Press, *The Journal of Politics* 63, no. 4 (November): 1041–1066.

Fallows, James. 2020. “Trump Tweets ‘Liberate’.” Accessed on: 2022-12-27, *The Atlantic* (April).

FBI. 2022. *National Incident-Based Reporting System (NIBRS), Federal Bureau of Investigation.*

Gadarian, Shana Kushner, Sara Wallace Goodman, and Pepinsky. 2022. *Pandemic Politics: The Deadly Toll of Partisanship in the Age of COVID.* Princeton University Press.

Google. 2020. “Google Trends: Search Term: *“liberate”*.” Accessed on: 2022-02-17 (April).

Green, Donald P, Bradley Palmquist, and Eric Schickler. 2004. *Partisan hearts and minds: Political parties and the social identities of voters.* Yale University Press.

Green, Jon, Jared Edgerton, Daniel Naftel, Kelsey Shoub, and Skyler J Cranmer. 2020. “Elusive consensus: Polarization in elite communication on the COVID-19 pandemic.” *Science Advances* 6 (28): eabc2717.

Grootendorst, Maarten. 2022. “BERTopic: Neural topic modeling with a class-based TF-IDF procedure.” *arXiv preprint arXiv:2203.05794.*

Grossman, Guy, Soojong Kim, Jonah M. Rexer, and Harsha Thirumurthy. 2020. "Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States." *Proceedings of the National Academy of Sciences* 117, no. 39 (September): 24144–24153.

Hahn, Robert A. 2021. "Estimating the COVID-related deaths attributable to president Trump's early pronouncements about masks." *International Journal of Health Services* 51 (1): 14–17.

Hamilton, Lawrence C, and Thomas G Safford. 2021. "Elite cues and the rapid decline in trust in science agencies on COVID-19." *Sociological Perspectives* 64 (5): 988–1011.

Hensel, Lukas, Marc Witte, A Stefano Caria, Thiemo Fetzer, Stefano Fiorin, Friedrich M Götz, Margarita Gomez, Johannes Haushofer, Andriy Ivchenko, Gordon Kraft-Todd, et al. 2022. "Global behaviors, perceptions, and the emergence of social norms at the onset of the COVID-19 pandemic." *Journal of Economic Behavior & Organization* 193:473–496.

Iyengar, Shanto, and Adam F. Simon. 2000. "New Perspectives and Evidence on Political Communication and Campaign Effects." *Annual Review of Psychology* 51 (1): 149–169.

Jørgensen, Frederik, Alexander Bor, Marie Fly Lindholt, and Michael Bang Petersen. 2021. "Public support for government responses against COVID-19: Assessing levels and predictors in eight Western democracies during 2020." *West European Politics* 44 (5-6): 1129–1158.

Jørgensen, Frederik, Alexander Bor, and Michael Bang Petersen. 2021. "Compliance without fear: Individual-level protective behaviour during the first wave of the COVID-19 pandemic." *British Journal of Health Psychology* 26 (2): 679–696.

Kam, Cindy D. 2005. "Who toes the party line? Cues, values, and individual differences." *Political Behavior* 27:163–182.

Leeper, Thomas J., and Rune Slothuus. 2014. "Political Parties, Motivated Reasoning, and Public Opinion Formation." *Political Psychology* 35 (S1).

Lopez, Ernesto, and Richard Rosenfeld. 2021. "Crime, quarantine, and the US coronavirus pandemic." *Criminology & Public Policy* 20 (3): 401–422.

Lupia, Arthur. 1994. "Shortcuts versus encyclopedias: Information and voting behavior in California insurance reform elections." *American Political Science Review* 88 (1): 63–76.

Lupia, Arthur, and Arthur D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* [In en]. Cambridge University Press.

Lupu, Noam, and Elizabeth J Zechmeister. 2021. "The early COVID-19 pandemic and democratic attitudes." *PloS One* 16 (6): e0253485.

McCord, Mary. 2020. *Trump's 'LIBERATE MICHIGAN!'Tweets Incite Insurrection. That's Illegal. Washington Post.*

Meta. 2022. *Movement Range Maps.* Data for Good at Meta (previously Facebook).

Nicholson, Stephen P. 2012. “Polarizing Cues” [in en]. *American Journal of Political Science* 56 (1): 52–66.

Plotly. 2015. “Collaborative Data Science.” https://plot.ly.

Pressman, Jeremy, and Erica Chenoweth. 2022. *Crowd Counting Consortium.*

Roberts, Damon C, and Stephen M Utych. 2021. “Polarized social distancing: Residents of Republican-majority counties spend more time away from home during the COVID-19 crisis.” *Social Science Quarterly* 102 (6): 2516–2527.

Samuels, David, and Cesar Zucco. 2014. “The Power of Partisanship in Brazil: Evidence from Survey Experiments.” *American Journal of Political Science* 58 (1): 212–225.

Sides, John, Michael Tesler, and Lynn Vavreck. 2017. “The 2016 US election: How Trump lost and won.” *Journal of Democracy* 28 (2): 34–44.

Smith, Mitch. 2022. *Jury Convicts Men of Supporting Plot to Kidnap Michigan Governor. The New York Times.* https://www.nytimes.com/2022/10/26/us/michigan-wolverine-watchmen-trial.html. Accessed on: 2023-02-27.

Stickle, Ben, and Marcus Felson. 2020. “Crime rates in a pandemic: The largest criminological experiment in history.” *American Journal of Criminal Justice* 45 (4): 525–536.

Tesler, Michael. 2012. “The Spillover of Racialization into Health Care: How President Obama Polarized Public Opinion by Racial Attitudes and Race” [in en]. *American Journal of Political Science* 56 (3): 690–704.

Twitter. 2021. *Academic research product track API.*

Van Green, Ted, and Alec Tyson. 2020. "5 Facts about Partisan Reactions to COVID-19 in the US." *Pew Research Center.*

Waskom, Michael L. 2021. "Seaborn: statistical data visualization." *Journal of Open Source Software* 6 (60): 3021.

Wikström, Per-Olof H. 2006. "Individuals, settings, and acts of crime: Situational mechanisms and the explanation of crime." *The Explanation of Crime: Context, Mechanisms and Development,* 61–107.

Xu, Yiqing, and Licheng Liu. 2022. *gsynth: Generalized Synthetic Control Method.* R package version 1.2.1.

Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science* 36 (3): 579–616.

Part I

Appendix: Elite Cues and Mass Non-compliance

Table of Contents

A Topic Models

To give greater understanding of how the messages were interpreted, we collected all the available quote tweets using the Twitter V2 API (Twitter 2021). Our analysis focused on messages that quote-tweeted the original three messages, as well as the messages that then re-tweeted those messages as well. In total, we collected 143,171 quote tweets. Text pre-processing included removing all usernames and URLs from the text. We then removed all messages that did not include 40 or more characters after the URLs and usernames were removed. The logic behind this step was to remove messages that did not include substantive text and instead were used as retweets simply to either tag another user or share an accompanying URL as part of the quote tweet. Additionally, we removed duplicate messages because many of the messages were simply retweets of retweets. Following these steps, our analysis included 14,875 tweets.

The hierarchical topic models were constructed using BERTopic Grootendorst (2022). BERTopic calculates the topic-term matrix (c-TF-IDF matrix) and then we apply a centroid linkage function to approximate the cosine distances between hierarchical clusters. Prior to reduction, the model identified 189 topics. We then reduced this number to 100. We additionally present results after reducing the number of topics to 10 below.

A.0.1 Un-interpreted topics from Hierarchical Topic Model

In Figure 1 in the main text, the top-20 topics were interpreted by the authors. Below, we present those topics (in their uninterpreted form) in the same order in which they appear in the figure in the main text.

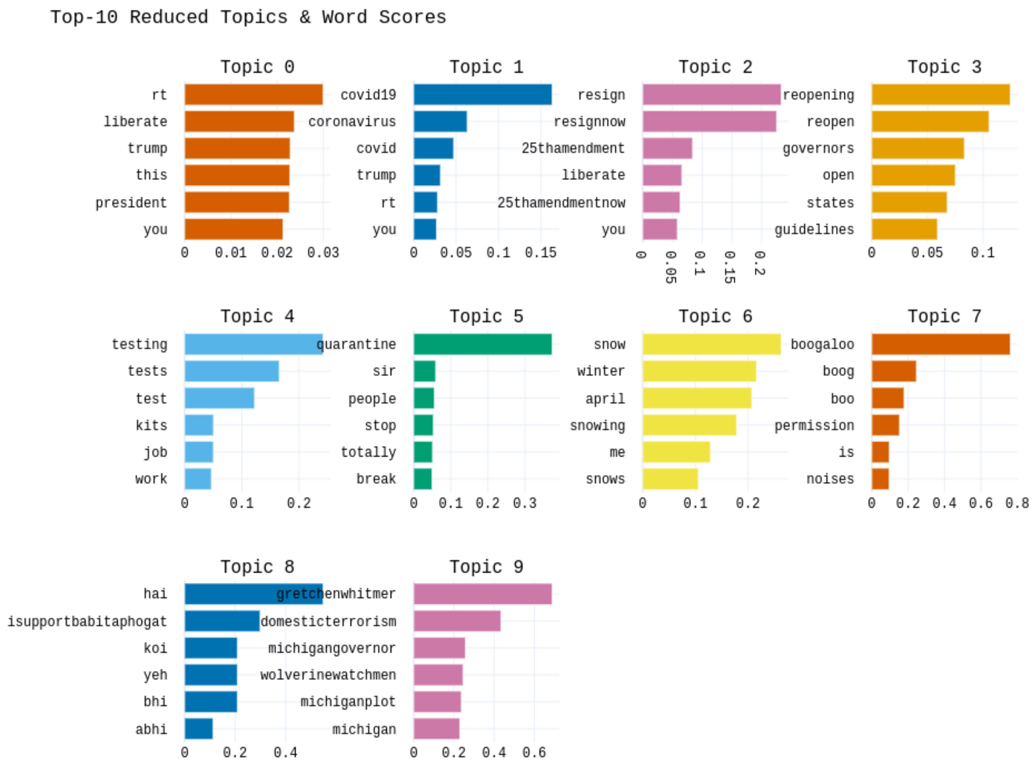
Table A1: Raw Topic Descriptions

Topic	Topic number
presidente si appelle	1
kidnap whitmer plot	18
november michigan liberate	19
white house siege	2
virginia virginian virginias	3
liberated need liberation	12
minnesota minnesotan minnesotas	4
canada invaded liberate	9
means liberate wonder	17
coupled attn whistles	10
amendment 25th 2nd	6
guns gun covid19	16
pandemic global middle	5
civil war violence	0
tweet tweets tweeting	11
cases deaths lowest	8
protest protestors protesting	14
governors militia leader	7
killed americans kill	15
virus trumpvirus spread	13

A.1 Top-ten Topics from Quote Tweets

The following presents the top-ten topics after further reducing all 189 topics identified down to ten. For readers less familiar with the American political landscape, Topic 7 refers to the "Boogaloo Bois", a far-right anti-government extremist group. Another notable and ominous topic is number 9, which mentions the "Wolverine Watchmen" and Gretchen Whitmer, the governor for the state of Michigan. More than a dozen individuals identifying as the Wolverine Watchmen were later arrested in October 2020 for plotting to kidnap Whitmer and violently overthrow the state government. Several members involved have since pleaded guilty of conspiracy to commit kidnapping, gun crimes and other terrorist-related charges.

Figure A1: Reduced Topics and Word scores



B Exogeneity Assumption

We examine whether Trump was simply responding to increased protest activity in the three states he targeted by regressing a binary treatment indicator identifying the three states on the number of protests in the 30 days surrounding the time of the messages. We use protest data available from the Crowd Counting Consortium (Pressman and Chenoweth 2022). More information on the Crowd Counting Consortium can be found on the CCC homepage or in this article in *The Washington Post*.

Our regressions are formalized using the following equation:

$$Y_{i,j} = \alpha + \beta \text{protests}_{i,j} + \delta_t + \zeta_i + \epsilon_{i,j} \tag{3}$$

Where Y is a treatment indicator for state i at time t, and protests is the corresponding number of (log) protests. δ and ζ are fixed effects parameters for time and state, respectively. We cluster standard errors by state and time.

Exogeneity Assumption – Results

We present regressions results using the design articulated above with a number of different state level characteristics to test our assumption that the president's decision to target Minnesota, Virginia and Michigan is as good-as-random conditional on observable factors. We present regression results below using daily state protests (model 1), COVID-19 cases (model 2), COVID-19 deaths (model 3), and violent crime (model 4) to try to predict Trump's decision to target the three states. The results demonstrate that the estimates are not statistical differentiable from zero when using any of the four state-level characteristics in the time leading up to President Trump's messages.

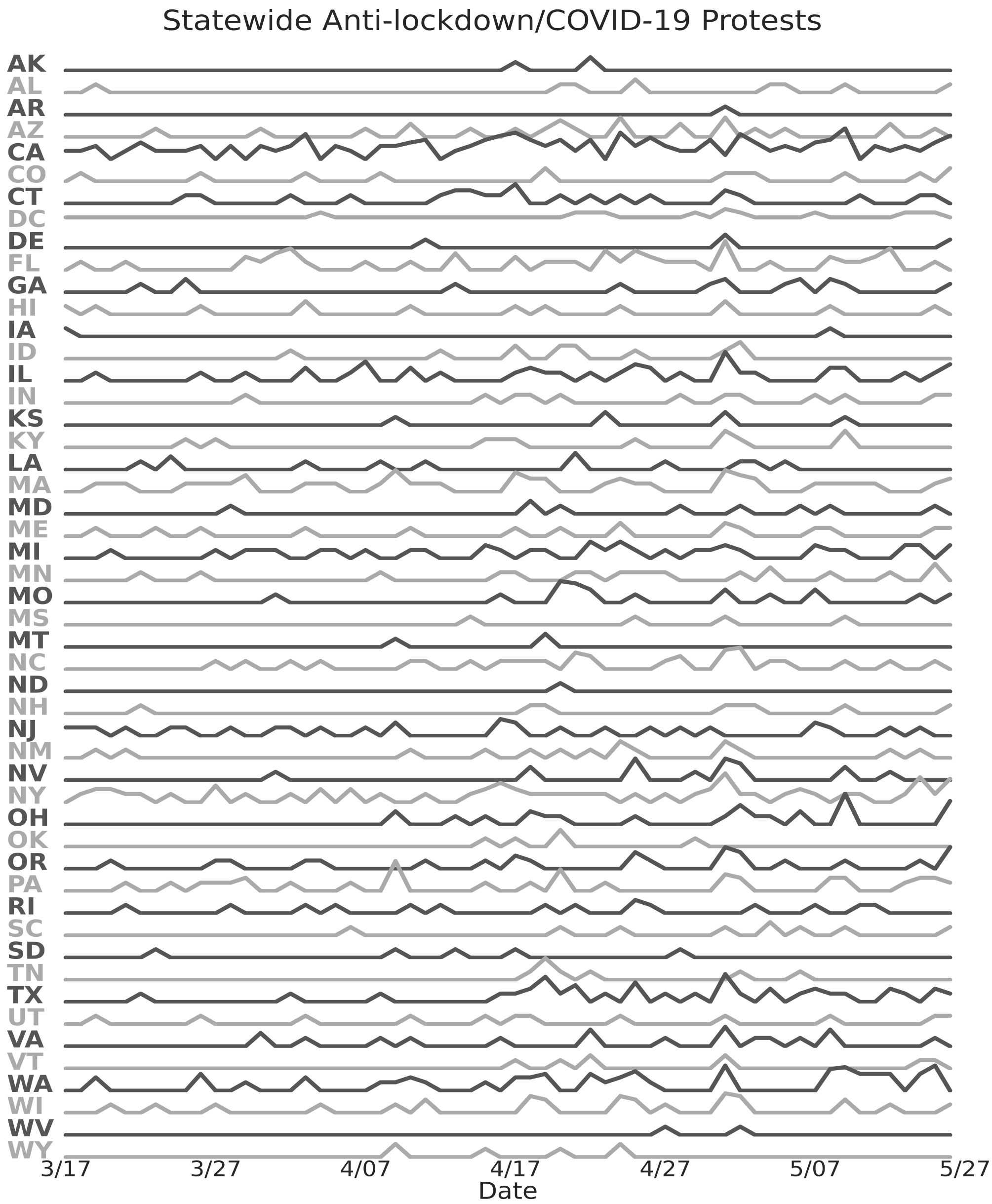
A point we make in the article's footnotes (10) is that the most likely explanation for why Trump targeted each of the three states are viewed as battlegrounds for Trump's 2020 Presidential Election bid. Each state was narrowly won by Trump's competitor, Hilary Clinton in the 2016 US Presidential election. Moreover, each of the three states had Democratic governors at the time, which also may have contributed to Trump's desire to target these states specifically. Should our assumptions be correct, there are no logical reasons to expect that a narrow margin of victory for Trump, or Trump's desire to win the states in the 2020 Presidential Election, would confound the relationship between treatment and our outcomes of interest. Furthermore, the fact that each of the three targeted states were won by the Democratic nominee for president in 2016 strongly suggests that there are fewer partisans that are receptive to Trump's cues, which means that the effect of the cues may be diminished in the three targeted states compared to states with higher levels of support for the president.

Table A2: Predicting Targeted States with State Characteristics

	1	2	3	4
Anti-lockdown protests	3.138e-05			
	(0.0004)			
(log) COVID-19 cases		0.1308		
		(0.1510)		
(log) COVID-19 deaths			0.0463	
			(0.1166)	
(log) Violent crime per 100k				-0.0035
				(0.0050)
Dep. Variable	Targeted	Targeted	Targeted	Targeted
No. Observations	3060	3000	3000	2688
R-squared	3.316e-07	0.0055	0.0011	9.346e-05
R-Squared (Within)	2.314e-05	0.0290	0.0215	-0.0005
R-Squared (Between)	5.694e-05	-4.8204	-0.1064	0.0122
R-Squared (Overall)	3.948e-05	-2.3149	-0.0403	0.0058
F-statistic	0.0010	15.920	3.2638	0.2412
P-value (F-stat)	0.9751	0.0001	0.0709	0.6234

Std. Errors are clustered by state and date and are reported in parentheses

Figure A2: Statewide Daily Anti-lockdown/COVID-19 Protests



Values presented in natural logarithmic form. Daily statewide anti-lockdown protests include protests in which 'COVID-19' was a listed issue for the protest. Data source: Crowd Counting Consortium Pressman and Chenoweth (2022).

C Covariate Balance – Mobility

Table A3: Covariate balance between treated and control counties

	Control ($N = 27837$)		Treatment ($N = 1227$)	
	Mean	Std. Dev.	Mean	Std. Dev.
Mobility (pre-treatment)	-26.4	10.0	-30.4	9.8
Stay-at-home (pre-treatment)	26.6	5.5	29.0	5.6
Income/capita	10.95	0.24	11.02	0.25
Unemployment	1.95	0.27	1.94	0.29
Income (within state)	4.50	0.20	4.45	0.25
County population	10.67	1.24	10.44	1.08
White percentage	4.23	0.19	4.33	0.08
Black percentage	2.60	0.70	2.66	0.42
Over 65	2.80	0.09	2.77	0.01
Median age	3.66	0.06	3.68	0.02
COVID cases (county)	3.26	1.78	2.78	1.58
COVID deaths (county)	0.82	1.20	0.53	1.01
COVID cases (state)	8.99	1.16	8.83	1.05
Biden vote share	0.3	0.2	0.3	0.1
Trump vote share	0.5	0.2	0.4	0.2

D Full results from Table 1 and Table 2

D.1 Full Results Table from Mobility Analysis

Below presents the full results from the mobility analysis. The time-invariant county characteristics were interacted with a time variable.

Table A4: Effect of Cues on County Mobility

	1	2	3
Treatment effect	2.284	1.005	2.706
	(0.906)	(0.631)	(0.854)
County COVID cases	−0.248	−0.357	−0.258
	(0.275)	(0.406)	(0.252)
County COVID deaths	−0.239	−0.029	−0.247
	(0.301)	(0.451)	(0.308)
State COVID cases	0.392	0.040	0.864
	(0.496)	(0.252)	(0.525)
County population	−0.260	−0.060	−0.244
	(0.175)	(0.163)	(0.682)
Income	−0.352	−0.471	−0.296
	(0.481)	(0.468)	(0.470)
Unemployment	−0.104	−0.086	−0.114
	(0.168)	(0.170)	(0.191)
Income within state	0.375	0.433	0.339
	(0.500)	(0.447)	(0.491)
County population	−0.006	−0.003	−0.009
	(0.015)	(0.024)	(0.016)
White percentage	0.082	−0.202	0.185
	(0.318)	(0.309)	(0.286)
Black percentage	−0.093	−0.157	−0.059
	(0.108)	(0.092)	(0.118)
Percent over 65	−0.924	−0.631	−1.001
	(0.343)	(0.430)	(0.396)
Median age	1.363	1.446	1.269
	(0.904)	(0.500)	(1.107)
Num.Obs.	29 064	6132	22 932
R2	0.764	0.825	0.714
R2 Adj.	0.745	0.810	0.692
R2 Within	0.017	0.022	0.018
R2 Within Adj.	0.017	0.019	0.017
RMSE	5.00	4.36	5.12
FE: County	X	X	X
FE: Time	X	X	X

Standard errors are clustered by state and presented in parentheses. All time-invariant variables are interacted with a time fixed effects integer. All count and proportional data are estimated in log form.

Table A5: Effect of Cues on Stay-at-home Compliance

	1	2	3
Treatment effect	−1.128	−0.660	−1.336
	(0.502)	(0.438)	(0.476)
County COVID cases	0.165	0.312	0.160
	(0.081)	(0.175)	(0.084)
County COVID deaths	0.093	0.189	0.068
	(0.140)	(0.194)	(0.126)
State COVID cases	−0.283	−0.425	−0.561
	(0.417)	(0.926)	(0.353)
County population	0.283	0.214	0.228
	(0.105)	(0.090)	(0.353)
Income	0.060	0.045	0.092
	(0.199)	(0.161)	(0.205)
Unemployment	0.014	−0.020	0.049
	(0.079)	(0.048)	(0.093)
Income within state	−0.194	−0.161	−0.209
	(0.184)	(0.132)	(0.198)
County population	−0.006	−0.010	−0.002
	(0.009)	(0.007)	(0.008)
White percentage	−0.053	0.024	−0.099
	(0.135)	(0.160)	(0.140)
Black percentage	−0.010	0.014	−0.032
	(0.033)	(0.026)	(0.036)
Percent over 65	0.180	0.000	0.283
	(0.297)	(0.246)	(0.367)
Median age	−0.249	−0.286	−0.222
	(0.659)	(0.467)	(0.780)
Num.Obs.	29 064	6132	22 932
R2	0.883	0.904	0.869
R2 Adj.	0.874	0.896	0.859
R2 Within	0.020	0.022	0.026
R2 Within Adj.	0.020	0.020	0.025
RMSE	1.91	1.89	1.90
FE: County	X	X	X
FE: Time	X	X	X

Standard errors are clustered by state and presented in parentheses. All time-invariant variables are interacted with a time fixed effects integer. All count and proportional data are estimated in log form.

E Robustness Check – Google Mobility

For robustness, we replicate the primary analysis using data from Google's Community Mobility Reports (Aktay et al. 2020). The Google Community Mobility Reports data consist of aggregated and anonymized daily mobility data. These data were similarly created with the aim of aiding public health officials in combating COVID-19. Mobility data were calculated daily for each US county using the median daily value from the respective location's five-week period in January 2020 (January 3 – February 6). Daily county values are then provided as the percentage change in mobility from the respective area's median value. These data are especially informative given that mobility for various activities are available. For each US county, daily data are available for human mobility resulting from retail and recreation, grocery and pharmacy, transit and transportation, workplace mobility and residential mobility.

As with the primary analysis, we use two different measures of mobility. First, we consider mobility amounting from retail and recreation. We consider retail and recreation mobility as a key indicator of individuals' compliance with stay-at-home orders because it is discretionary and therefore indicates a clear violation of the orders. Second, we create an aggregate measurement of mobility by combining the five mobility measures provided by Google to create a single mobility outcome.

In our analysis, we only include states under stay-at-home orders and we do not include counties for which data is unavailable for any of the days within the time-frame of the analysis. Our model can be formalized using the following equation:

$$Y_{i,j,t} = \alpha + \theta Targeted_{i,j,t} + \delta_t + \zeta_i + X_{i,j,t}\beta_1 + W_{i,j}\beta_2 + \epsilon_{i,j,t} \tag{4}$$

Using the above specification, we estimated six models. Similar to the primary analysis, model one estimates the effect of the cues on all the counties within the treated states. Model 2 estimates the effect of the cues on Republican counties (with Republican counties as the control), and model three estimates the effect of the cues on democratic counties (with Democratic counties as the control).

Table A6: Effect of "Liberate" Cues on Retail and Recreational mobility

	Retail and recreation mobility		
	(1)	(2)	(3)
Treatment effect	2.577	1.201	3.866*
	(1.591)	(1.167)	(1.392)
COVID-19 cases	-0.481	-0.455	-0.435
	(0.292)	(0.335)	(0.336)
COVID-19 deaths	-0.343	-0.012	-0.637*
	(0.241)	(0.266)	(0.229)
Time × presidential vote	-0.454***	-0.577***	-0.368***
	(0.066)	(0.115)	(0.054)
Time × income	-1.045***	-1.074***	-1.023***
	(0.244)	(0.263)	(0.261)
Time × unemployment	-0.205+	-0.242*	-0.146
	(0.106)	(0.089)	(0.134)
Time × log income state	0.630*	0.643*	0.647*
	(0.231)	(0.261)	(0.256)
Time × population	-0.026***	-0.028***	-0.020*
	(0.006)	(0.007)	(0.008)
Time × white percentage	0.588+	0.619	0.552
	(0.300)	(0.398)	(0.332)
Time × black percentage	0.049	0.067	0.029
	(0.068)	(0.079)	(0.067)
Time × percent over 65	0.103	0.412	0.005
	(0.672)	(0.726)	(0.759)
Time × median age	-1.517	-1.556	-1.643
	(1.199)	(0.974)	(1.629)
Num.Obs.	22346	9357	12989
R2	0.896	0.915	0.872
R2 Adj.	0.891	0.911	0.865
R2 Within	0.135	0.167	0.070
R2 Within Adj.	0.134	0.166	0.069
RMSE	4.48	4.16	4.58
FE: county	X	X	X
FE: time	X	X	X

$+p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001$

Standard errors are clustered by state and provided in parentheses.

Table A7: Effect of "Liberate" Cues on aggregate mobility

	Aggregate mobility		
	(1)	(2)	(3)
Treatment effect	13.487	14.498	13.416*
	(8.462)	(12.152)	(4.962)
COVID-19 cases	-2.242	-2.322	-2.068+
	(1.401)	(2.668)	(1.195)
COVID-19 deaths	-0.783	1.062	-1.943+
	(1.624)	(1.955)	(1.095)
Time × presidential vote	-0.672*	-1.771**	0.099
	(0.316)	(0.527)	(0.401)
Time × income	-4.127*	-4.469*	-3.893**
	(1.616)	(1.973)	(1.189)
Time × unemployment	-0.449	-0.573	-0.261
	(0.403)	(0.351)	(0.521)
Time × log income state	3.713*	4.053*	3.557**
	(1.410)	(1.697)	(1.120)
Time × county population	0.014	-0.048	0.104*
	(0.011)	(0.048)	(0.046)
Time × white percentage	-1.710	-2.476	-1.302
	(1.127)	(1.789)	(0.924)
Time × black percentage	-0.471	-0.768+	-0.316
	(0.364)	(0.398)	(0.343)
Time × percent over 65	-3.381	-4.816	-2.209
	(2.336)	(3.645)	(1.823)
Time × median age	2.170	2.994	0.455
	(3.369)	(3.764)	(3.602)
Num.Obs.	22346	9357	12989
R2	0.762	0.764	0.721
R2 Adj.	0.749	0.751	0.706
R2 Within	0.020	0.022	0.014
R2 Within Adj.	0.019	0.021	0.013
RMSE	29.14	31.89	26.64
FE: county	X	X	X
FE: time	X	X	X

$+p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001$

Standard errors are clustered by state and provided in parentheses.

F First Difference – Mobility

F.1 Replication of Mobility Analysis using First-Difference

This section replicates the analysis of the effect of the cues on mobility (Table 1) and stay-at-home compliance (Table 2) in targeted counties using first difference model specifications and Meta Mobility data (Meta 2022).

Table A8: Effect of "Liberate" Cues on mobility using First Difference models

	Mobility		
	Full state	Blue counties	Red counties
Effect of targeted cues	1.110*	−0.370	1.703**
	(0.473)	(0.849)	(0.564)
County	✓	✓	✓
Time	✓	✓	✓
Control variables	✓	✓	✓
Num.Obs.	29 064	6132	22 932
Num.Obs. in estimation:	26988	5694	21294
R2	0.413	0.442	0.415
R2 Adj.	0.410	0.428	0.411

* p < 0.05, ** p < 0.01, *** p < 0.001

Standard errors clustered by state and day and provided in parentheses.

Table A9: Effect of "Liberate" Cues on stay-at-home compliance using First Difference models

	Stay-at-home compliance		
	Full state	Blue counties	Red counties
Effect of targeted cues	−0.631***	0.134	−0.895***
	(0.179)	(0.366)	(0.207)
County	✓	✓	✓
Time	✓	✓	✓
Control variables	✓	✓	✓
Num.Obs.	29 064	6132	22 932
Num.Obs. in estimation:	26988	5694	21294
R2	0.721	0.707	0.730
R2 Adj.	0.720	0.700	0.728

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Standard errors clustered by state and day and provided in parentheses.

G Descriptive Statistics and Covariate Balance

G.1 Crosstabs – Race and Offence

The following presents crosstabs for race of the offender and the offence type. The arrests included occurred during the time period from 2020-04-07–2020-04-27.

Table A10: Crosstabs – Race and Offence

Race		Aggravated Assault	Destruction/Damage/Vandalism of Property	Disorderly Conduct	Simple Assault	All
Native American	N	103	72	10	191	376
	% row	27.4	19.1	2.7	50.8	100.0
Asian	N	64	37	2	151	254
	% row	25.2	14.6	0.8	59.4	100.0
Black or African American	N	529	411	94	689	1723
	% row	30.7	23.9	5.5	40.0	100.0
Pacific Islander	N	13	8	2	33	56
	% row	23.2	14.3	3.6	58.9	100.0
Unknown	N	83	58	11	189	341
	% row	24.3	17.0	3.2	55.4	100.0
White	N	709	640	185	845	2379
	% row	29.8	26.9	7.8	35.5	100.0
All	N 1501	1226	304	2098	5129	
	% row	29.3	23.9	5.9	40.9	100.0

Note: According to NIBRS (FBI 2022), "Pacific Islander" also includes Native Hawaiians. "Native American" includes what the FBI calls "American Indian or Alaska Native".

G.2 Crosstabs – Arrests by State

The following presents the number of arrests by state and according to the offence committed. The arrests included occurred during the time period from 2020-04-07–2020-04-27.

Table A11: Crosstabs – Arrests by State

State	Aggravated Assault	Destruction/Damage/Vandalism of Property	Disorderly Conduct	Simple Assault	All
Alabama	14	10	0	35	59
Arizona	28	37	25	48	138
Arkansas	46	30	1	49	126
Colorado	51	43	16	61	171
Connecticut	30	32	22	45	129
Delaware	31	29	0	45	105
District of Columbia	5	4	0	10	19
Georgia	43	42	6	50	141
Hawaii	13	4	0	57	74
Idaho	27	28	6	38	99
Illinois	14	14	5	30	63
Indiana	43	15	8	46	112
Iowa	51	36	2	47	136
Kansas	39	32	9	47	127
Kentucky	43	32	0	48	123
Louisiana	38	29	0	44	111
Maine	11	17	0	31	59
Maryland	16	11	2	30	59
Massachusetts	55	35	20	66	176
Michigan	57	42	10	66	175
Minnesota	44	33	0	76	153
Mississippi	23	12	0	45	80
Missouri	45	42	6	47	140
Montana	40	29	4	53	126
Nebraska	15	16	5	40	76
Nevada	19	15	1	40	75
New Hampshire	17	19	4	36	76
New Mexico	49	31	6	58	144
New York	24	26	0	35	85
North Carolina	52	44	13	60	169
North Dakota	24	15	1	50	90
Ohio	47	41	10	63	161
Oklahoma	52	36	2	56	146
Oregon	48	44	15	63	170
Pennsylvania	1	0	1	10	12
Rhode Island	18	20	10	36	84
South Carolina	45	40	25	48	158
South Dakota	45	18	0	64	127
Tennessee	48	43	7	59	157
Utah	40	50	14	76	180
Vermont	23	9	1	25	58
Virginia	45	42	7	61	155
Washington	73	68	7	99	247
West Virginia	29	13	0	41	83
Wisconsin	58	54	46	63	221
Wyoming	3	1	1	9	14
All	1582	1283	318	2206	5389
	29.4%	23.8%	5.9%	40.9%	100.0

G.3 Covariate Balance

The table below presents covariate balance between the treated and control states during the time period from 2020-04-07–2020-04-27. Only offences by White individuals are considered in the number of arrests/crimes below. Not all covariates presented below were used in the model because they are time-invariant and therefore are absorbed in the fixed effects.

Table A12: Covariate Balance Between Targeted and Untargeted States – Full dataset

	Control		Treatment			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	Std. Error
Arrests	10.7	13.9	17.5	19.2	6.8	1.4
Arrests/100k	0.3	0.3	0.2	0.2	-0.1	0.0
(log) Arrests/100k	-2.0	1.3	-2.4	1.4	-0.4	0.1
Total State arrests	64.3	54.4	78.8	40.4	14.5	3.0
State Temp.	14.8	7.0	10.0	5.6	-4.8	0.4
COVID-19 cases	11400.0	33891.0	15205.0	13482.3	3805.1	3195.7
COVID-19 deaths	616.3	2360.5	1056.6	1187.5	440.4	273.3
(log) COVID-19 cases	8.2	1.3	9.2	1.1	1.0***	0.2
(log) COVID-19 deaths	4.7	1.7	6.3	1.2	1.6***	0.3
State population	5004560.2	3945936.3	8071486.7	1841490.3	3066926.4***	427292.0
(log) State population	15.1	0.9	15.9	0.2	0.8***	0.1
(log) population density	4.6	1.2	4.9	0.5	0.4**	0.1
(log) Unemployment rate	2.0	0.3	2.0	0.1	0.0	0.0
(log) Income/capita	10.4	0.1	10.5	0.1	0.1***	0.0
(log) Black perc.	1.9	1.1	2.5	0.5	0.6***	0.1
(log) Hispanic perc.	2.1	0.8	2.1	0.3	0.0	0.1
(log) White perc.	4.3	0.2	4.3	0.1	0.0	0.0
(log) Male perc.	3.9	0.0	3.9	0.0	0.0**	0.0
(log) Median age	3.7	0.1	3.7	0.0	0.0	0.0
(log) Over 65 perc	2.8	0.1	2.7	0.0	-0.1***	0.0
GINI index	0.5	0.0	0.5	0.0	0.0+	0.0

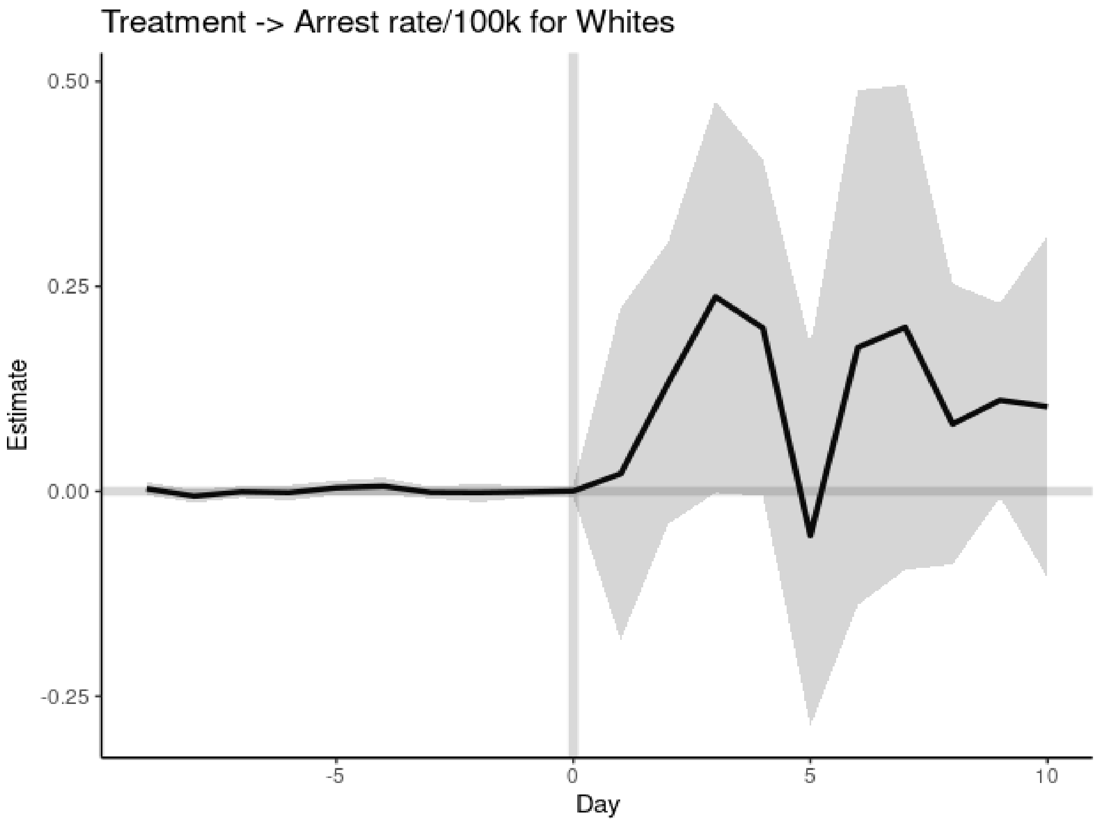
Note: All measurements are made daily. State demographic characteristics are not included in statistical estimates because they do not vary over time in the analysis.

H Full Results from Crime Analysis

H.1 Full results from Crime Analysis

The following includes full results from matrix completion estimation of the primary crime results presented in the article. The estimates are made using the Gsynth library in R (Xu and Liu 2022).

Figure A3: Treatment Effect: Arrest rate/100k for Whites



ATT estimates for the effect of targeted cues on the arrest rate for white individuals for crimes related to assault, disorderly conduct and vandalism/destruction of property. Grey area represents 95% confidence intervals.

H.1.1 ATT: Arrest rate/100k

Table A13: ATT: Arrest rate/100k for Whites

	Estimate	S.E.	CI.lower	CI.upper	p.value
ATT.avg	0.121	0.054	0.015	0.226	0.025

H.1.2 Cummulative ATT: Arrest rate/100k for Whites

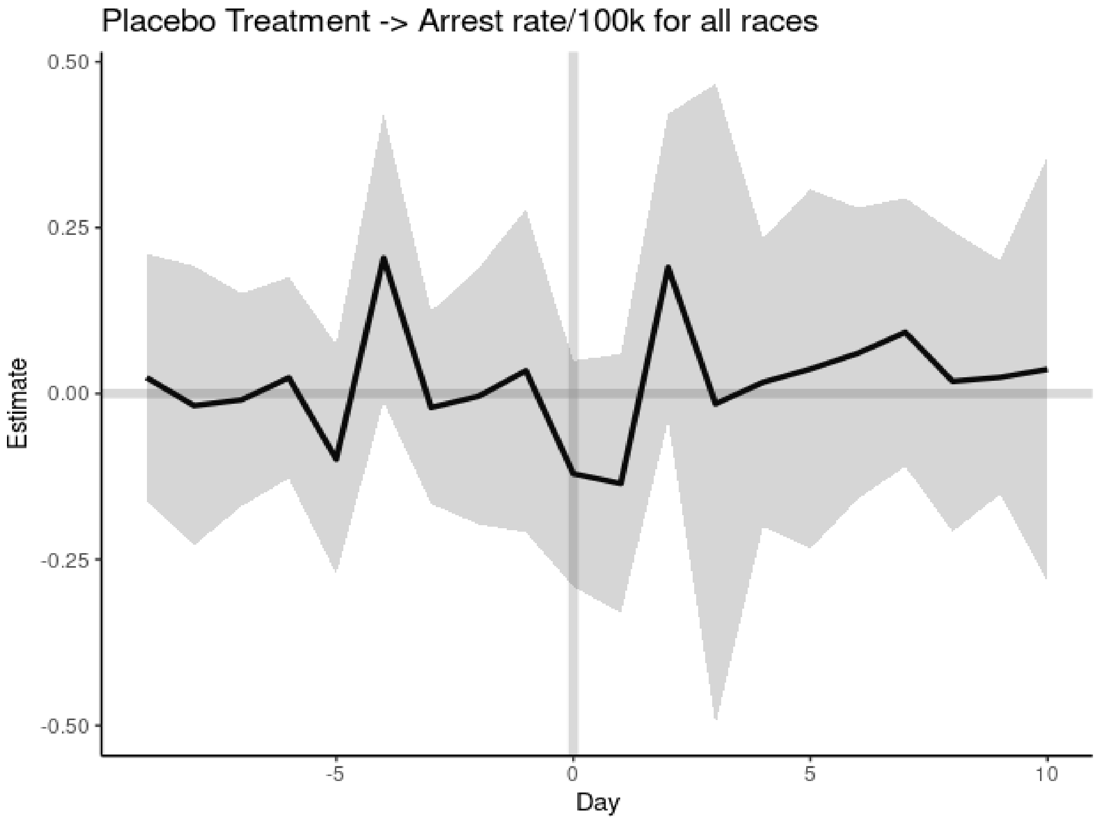
	CATT	S.E.	CI.lower	CI.upper	p.value
0	0.00	0.00	-0.01	0.01	0.95
1	0.02	0.10	-0.19	0.21	0.89
2	0.15	0.16	-0.17	0.46	0.35
3	0.39	0.16	0.09	0.69	0.01
4	0.59	0.20	0.20	0.98	0.00
5	0.54	0.20	0.13	0.93	0.01
6	0.71	0.30	0.11	1.27	0.02
7	0.91	0.43	0.04	1.68	0.04
8	0.99	0.44	0.08	1.78	0.03
9	1.11	0.46	0.16	1.94	0.02
10	1.21	0.54	0.13	2.21	0.02

I Robustness: Arrest rate/100k for all races

I.1 Full results from Robustness Check

The following results replicate the primary analysis using arrests for the same crimes but include all races.

Figure A4: Placebo Treatment: Arrest rate/100k for all races



I.1.1 ATT: Arrest rate/100k for all races

Table A14: ATT: Arrest rate/100k for all races

	Estimate	S.E.	CI.lower	CI.upper	p.value
ATT.avg	0.027	0.049	-0.069	0.123	0.578

I.1.2 Cummulative ATT: Arrest rate/100k for all races

Table A15: CATT: Arrest rate/100k for all races

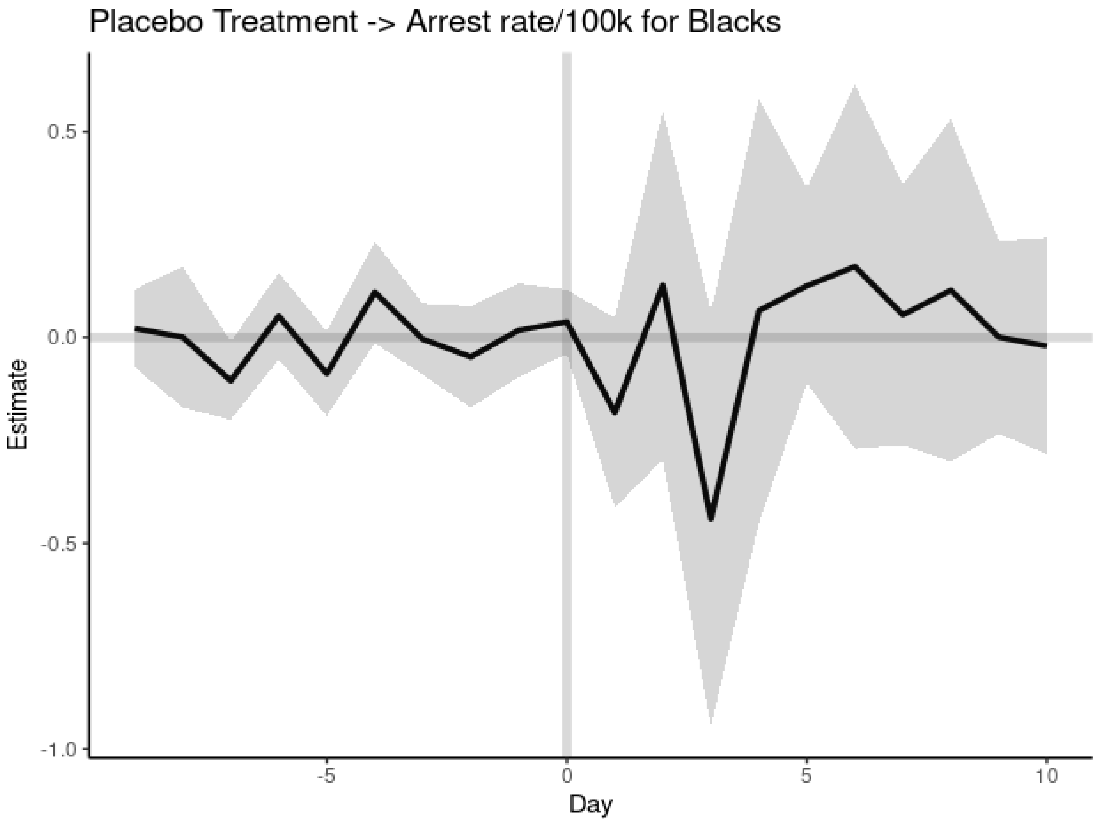
	CATT	S.E.	CI.lower	CI.upper	p.value
0	-0.12	0.09	-0.28	0.06	0.18
1	-0.26	0.13	-0.50	0.00	0.05
2	-0.11	0.18	-0.45	0.23	0.59
3	-0.13	0.28	-0.67	0.45	0.65
4	-0.11	0.28	-0.67	0.47	0.69
5	-0.07	0.28	-0.58	0.52	0.79
6	-0.00	0.32	-0.61	0.66	0.97
7	0.09	0.36	-0.64	0.80	0.76
8	0.11	0.41	-0.74	0.90	0.74
9	0.13	0.44	-0.79	0.96	0.73
10	0.17	0.52	-0.89	1.18	0.70

J Robustness: Arrest rate/100k for Blacks

J.1 Full results from Robustness Check

The following results replicate the primary analysis using arrests for the same crimes but include arrests from Blacks only.

Figure A5: Placebo Treatment: Arrest rate/100k for Blacks



J.1.1 ATT: Arrest rate/100k for Blacks

Table A16: ATT: Arrest rate/100k for Blacks

	Estimate	S.E.	CI.lower	CI.upper	p.value
ATT.avg	0.002	0.077	-0.150	0.153	0.984

J.1.2 Cummulative ATT: Arrest rate/100k for Blacks

Table A17: CATT: Arrest rate/100k for Blacks

	CATT	S.E.	CI.lower	CI.upper	p.value
0	0.04	0.04	-0.04	0.11	0.40
1	-0.15	0.13	-0.43	0.10	0.25
2	-0.02	0.22	-0.46	0.41	0.85
3	-0.46	0.26	-1.01	-0.01	0.04
4	-0.39	0.36	-1.07	0.33	0.32
5	-0.27	0.39	-0.99	0.51	0.56
6	-0.10	0.53	-1.12	0.96	0.86
7	-0.04	0.62	-1.17	1.22	0.99
8	0.07	0.66	-1.05	1.53	0.91
9	0.07	0.73	-1.17	1.66	0.89
10	0.05	0.76	-1.26	1.67	0.96

K TWFE Estimations of Arrest Rate for Whites

K.1 TWFE estimations

For robustness, we estimated the results using the two-day average of the arrest rate. We estimate the following fixed effects equation:

$$Y_{i,t} = \alpha + \theta Targeted_{i,t} + \delta_t + \zeta_i + W_{i,t} + \epsilon_{i,t} \tag{5}$$

Where Y is violent crime in state i at time t. W is a matrix of time varying, state-specific characteristics such as COVID-19 cases, COVID-19 deaths and daily average temperature measured at the state level. $Targeted$ is a dummy that takes the value of 1 at the day on which President Trump called for the liberation of Michigan, Minnesota and Virginia and thereafter, and is 0 in all other cases. δ and ζ are fixed effects parameters for time and state, respectively. The parameter of interest is θ, which estimates the cumulative difference in arrests within states and across time.

We present cumulative estimates of the cues on arrests of Whites for crimes related to assault, disorderly conduct and destruction/damage/vandalism of property below. The outcome in the models below measure arrests using a two-day average. The first model includes results without controls, while the remaining models present results with different configurations of the control variables. In model four which includes all the considered controls, the estimate suggest a cumulative increase of 15.6 percent in the arrest rate for crimes related to assault, disorderly conduct and destruction/damage/vandalism of property by white individuals in Michigan, Virginia and Minnesota following Trump's calls for the liberation of those states.

K.1.1 Two-day results

Table A18: Effect of Cues on Two-day Arrest Rate of Whites

	Arrest Rate			
	(1)	(2)	(3)	(4)
Targeted cues	0.138*	0.158*	0.119***	0.156*
	(0.046)	(0.061)	(0.023)	(0.060)
(log) COVID cases		−0.150		−0.245
		(0.218)		(0.184)
(log) COVID deaths		−0.044		0.009
		(0.159)		(0.099)
(log) Temp			0.040	−0.004
			(0.037)	(0.022)
(log) State Arrests				0.853***
				(0.127)
Num.Obs.	311	311	311	311
R2	0.949	0.949	0.949	0.965
R2 Adj.	0.940	0.940	0.940	0.959
R2 Within	0.005	0.009	0.009	0.319
R2 Within Adj.	0.001	−0.003	0.002	0.306
AIC	183.6	186.3	184.2	73.7
BIC	348.1	358.3	352.4	253.2
RMSE	0.28	0.28	0.28	0.23
FE: State	X	X	X	X
FE: Time	X	X	X	X

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Standard errors are clustered by state and date, and presented in parentheses.

K.1.2 Daily results

In the table below, we present estimates using TWFE models. In the estimates, the outcome is (log) daily White arrests for crimes related to assault, disorderly conduct and destruction/damage/vandalism of property. The estimates present the cumulative effect of the cues on arrests.

Table A19: Effect of Cues on Daily White Arrest Rate

	(1)	(2)	(3)	(4)
Targeted cues	0.097+	0.088+	0.074***	0.093***
	(0.052)	(0.044)	(0.015)	(0.016)
(log) COVID cases		−0.022		−0.010
		(0.111)		(0.085)
(log) COVID deaths		0.060		0.042
		(0.094)		(0.050)
(log) Temp			0.042*	0.014
			(0.019)	(0.017)
(log) State Arrests				0.758***
				(0.083)
Num.Obs.	777	777	777	777
R2	0.889	0.889	0.890	0.918
R2 Adj.	0.879	0.879	0.880	0.910
R2 Within	0.001	0.002	0.006	0.258
R2 Within Adj.	0.000	−0.002	0.004	0.253
AIC	707.5	711.1	705.5	484.5
BIC	1005.5	1018.4	1008.1	801.1
RMSE	0.35	0.35	0.35	0.30
FE: State	X	X	X	X
FE: Time	X	X	X	X

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Standard errors are clustered by state and date, and presented in parentheses.