

# Black Lives Matter Protests and Voter Turnout\*

Cameron Kimble<sup>†</sup>    Leanne Fan<sup>‡</sup>    Kasey Zapatka<sup>§</sup>    Kevin Morris<sup>¶</sup>

May 05, 2021

## Abstract

In the summer of 2020, Americans took to the streets in larger numbers than ever before. Following the police murders of George Floyd and Breonna Taylor, an enormous multiracial coalition voiced their dissatisfaction with the state of policing in the United States. These mass protests took place shortly before a presidential election, and the incumbent loudly voiced his disdain for protesters and their political message. This paper explores one aspect of the impact of the Black Lives Matter (BLM) movement. Using a national voter file and data on protest location, we aim to estimate the causal impact of physical proximity to a BLM protest on voter turnout. Our pilot studies using Georgia, North Carolina, and Ohio point to a complicated relationship that demands further investigation.

---

\*The authors thank XX for their feedback and support. All errors are our responsibility.

<sup>†</sup>Research and Program Associate, Brennan Center for Justice (cameron.kimble@nyu.edu)

<sup>‡</sup>PhD Student, Department of Sociology, CUNY Graduate Center (lfan@gradcenter.cuny.edu)

<sup>§</sup>PhD Candidate, Department of Sociology, CUNY Graduate Center (kzapatka@gradcenter.cuny.edu)

<sup>¶</sup>PhD Student, Department of Sociology, CUNY Graduate Center (kmorris@fgradcenter.cuny.edu)

# Data and Design

We use a variety of data sources and empirical approaches to understand the electoral consequences of the 2020 protests in support of the Black Lives Matter movement. We are interested ultimately in the *causal* effect of protest exposure on these beliefs and behaviors, complicating our empirical approach. It seems highly probable that factors associated with protest formation are also associated with other political behaviors. Put differently, demonstrating a correlation between protest exposure and voting behavior might simply point to a third factor influencing them both. An example might be helpful: it seems possible that areas with large Black populations were politicized by the Trump Administration’s response to the COVID-19 pandemic, which disproportionately impacted Black communities. This politicization—which may have occurred prior to the murder of George Floyd—could have increased an area’s propensity to protest *and* increased their likelihood of voting in November. Thus, any correlation between protest and turnout would be due not to the protests themselves, but the underlying politicization.

To estimate the causal effect of protests on voter turnout, we leverage the known fact that protests are less likely to develop in inclement weather [CITE]. We use an instrumental variables (IV) approach that allows us to leverage random fluctuations in rainfall, the resulting protest formation, and eventual turnout and political behavior. IV setups are a common way for identifying causal relationships in the social sciences [CITE; CITE; CITE].

IV models do, however, have one very strict assumption: namely, the *exclusion restriction*. In other words, we must assume that our exogenous variable (here, rainfall) is unrelated to our dependent variable in any way other than through our endogenous variable (here, protest formation). If there is reason to believe that rainfall influenced turnout in November or political beliefs, our estimates cannot determine the causal relationship. There is reason to be particularly concerned with a violation of this assumption when using rainfall: as Mellon (2021) shows, nearly 200 journal articles have used rainfall as an instrument in recent years.

It is incumbent upon us, then, to demonstrate that the use of rainfall does not violate the exclusion restriction.

To improve the validity of our approach, we do not use *overall* rainfall, but rather relative rainfall during the period in which the protests occurred. George Floyd was murdered on May 25, and protests began in the days that followed, and peaked the weekend of June 6 (Buchanan, Bui, and Patel 2020). As such, our instrumental variable is the amount of rainfall that fell between May 26 and June 7, 2020, *relative* to the amount of rainfall that fell in that period between 2000 and 2019. This avoids a rainfall measure that simply captures generally rainier parts of the country, whose political orientation may be different. It is still possible, however, that relative rainfall in late May and early June could influence our dependent variables through other avenues. This seems especially likely if places with high relative rainfall in May and June *also* had high rainfall in late October—such high rainfall could then reduce turnout [CITE]. However, as we demonstrate in the Supplemental Information, this is not the case. Relative rainfall in our period is only weakly associated with relative rainfall in late October ( $r = 0.07$ ). We thus conclude that our approach satisfies the exclusion restriction.

## Data

Our data come from a variety of sources. To estimate relative rainfall, we turn to the National Oceanic and Atmospheric Administration (NOAA), which collects detailed weather data from around the country. NOAA records the estimated daily rainfall at some 13,000 locations around the country. We use the `rnoaa` (Chamberlain 2021) package to download and process this data. The weather sites include each site’s latitude and longitude.

Our data on protest formation comes from the U.S. Crisis Monitor, a project of the Armed Conflict Location & Event Data Project (ACLED) and the Bridging Divides Initiative (BDI)

at Princeton University.<sup>1</sup> The U.S. Crisis compiled geocoded data on Black Lives Matter protests from around the United States throughout 2020. Using this data, we identify over 3,800 Black Lives Matter protests occurring between May 26 and June 7. These protests occurred in all 50 states and Washington, DC.

## Voter File Data

To explore the relationship between protest exposure and individual-level turnout, we leverage the national registered voter file provided by data vendor L2. L2 collects the registered voter file from each state in the country and includes a host of self-reported and modelled information such as age, partisan affiliation, gender, and race / ethnicity. Importantly, they also indicate whether each voter participated in the 2020 general election, as well as past contests. These records are geocoded, and are mapped to home Census block groups. Taken as a whole, we have the individual-level voter records for the more than 180 million registered voters in the United States. We calculate each voter’s exposure to relative rainfall by assigning them the relative rainfall of the nearest weather station. Their exposure to protest is calculated as the distance between their home address and the nearest Black Lives Matter protest. In addition to testing the relationship between distance to the closest protest (instrumented by rainfall), we control for individual- and neighborhood-level characteristics.

The individual-level registered voter file data allow us to test our first hypothesis:

*H1:* Exposure to protest [increased / decreased] turnout. This was [especially / less] true for Black voters.

There are some limitations to using the registered voter file to test for turnout effects. Most importantly, by looking at participation among registered voters, we may be “selecting” on the dependent variable (See Nyhan, Skovron, and Titiunik 2017). In other words, registration *itself* is a form of political participation. If protest organizers registered many new voters

---

<sup>1</sup>See <https://acleddata.com/special-projects/us-crisis-monitor/>.

at the protests—an approach that did apparently occur [CITE]—but only a relatively small share of these new registrants actually voted, the net result may have been *lower* turnout among registered voters but *higher* turnout among eligible citizens.

To sidestep this potential problem, we also aggregate the number of ballots cast up from the individual-level registered voter file to the Census [block group? tract?] level (Morris 2021, 2020). The Census Bureau makes estimates of the citizen voting-age population available at this level. This provides a denominator for turnout that is unbiased by differential registered voters. By testing whether the distance between the center of a [tract / block group] to a protest (again instrumented by the relative rainfall at the weather station closest to the [tract / bg]) is related to that neighborhood’s turnout, we estimate the causal effect of protest exposure on turnout among eligible citizens—and not just registered voters.

DEALING WITH SPATIAL DEPENDENCIES TKTKTKTKTK [KASEY]

## Survey Data

While the administrative data gives us exceptional coverage of the relationships between protest and turnout across the country, turnout is a relatively blunt measure of political participation. Moreover, it seems possible that the BLM protests could have shifted voters’ *beliefs* about topics such as the police without bringing new participants into the voting booth.

To understand the effect of protest exposure on political beliefs we turn to national survey data. We use two national surveys: the American National Election Study 2020 Time Series data (ANES) and the Cooperative Election Study (CES). Both are widely used among sociologists and political scientists to understand the political orientation of the American electorate. In addition to a rich set of data about individuals’ political beliefs, these surveys also collect information about respondents’ age, race, family income, and other characteristics. Each record includes the respondent’s home ZIP code. ADD A LITTLE MORE

## ABOUT SURVEYS / SAMPLE SIZE / ETC

Once again, our empirical strategy relies on using rainfall as an instrument for protest exposure. In the case of the national survey data, however, there are not respondents from all ZIP codes in the country. As such, we run our IV models in two steps. We begin by predicting the distance from the center of each ZIP code in the country to the closest protest as a function of rainfall. To do so, each ZIP code is assigned the relative rainfall of the closest weather station (or the mean relative rainfall of all weather stations in the ZIP code). We then measure the distance from the center of the ZIP code to the closest BLM protest (this is coded as 0 miles if a protest occurred anywhere inside the ZIP code). The predicted distance for each ZIP code is then merged with the national survey data.

## Constructing Political Attitudes in the ANES

TKTKTKTKTK [LEANNE]

## Constructing Political Attitudes in the CES

TKTKTKTKTK [LEANNE]

## References

- Buchanan, Larry, Quoc Trung Bui, and Jugal K. Patel. 2020. “Black Lives Matter May Be the Largest Movement in U.S. History.” *The New York Times: U.S.*, July 3, 2020. <https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html>.
- Chamberlain, Scott. 2021. *Rnoaa: 'NOAA' Weather Data from R*. <https://CRAN.R-project.org/package=rnoaa>.
- Mellon, Jonathan. 2021. “Rain, Rain, Go Away: 176 Potential Exclusion-Restriction Violations for Studies Using Weather as an Instrumental Variable.” SSRN Scholarly Paper ID 3715610. Rochester, NY: Social Science Research Network. <https://doi.org/10.2139/ssrn.3715610>.
- Morris, Kevin. 2020. “Neighborhoods and Felony Disenfranchisement: The Case of New York City.” *Urban Affairs Review*, May, 1078087420921522. <https://doi.org/10.1177/1078087420921522>.
- . 2021. “Turnout and Amendment Four: Mobilizing Eligible Voters Close to Formerly Incarcerated Floridians.” *American Political Science Review*, 1–16. <https://doi.org/10.1017/S0003055421000253>.
- Nyhan, Brendan, Christopher Skovron, and Rocío Titiunik. 2017. “Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach.” *American Journal of Political Science* 61 (3): 744–60. <https://doi.org/10.1111/ajps.12288>.