**Read Me File for “Black Lives Matter Protests and Risk Avoidance: The Case of Civil Unrest During a Pandemic” (Dave, Friedson, Matsuzawa, Sabia, and Safford 2023)**

**Data**

The “Data” folder contains all the publicly available datasets that we used for our analysis, including:

* “cbg\_pop”: Census block group population data from the Census.
* “Masks”, “Reopen\_Policy”, “SIPO\_Expire”: Our data for reopening/COVID-19 policy control variables, which were collected from <https://docs.google.com/spreadsheets/d/1zu9qEWI8PsOI_i8nI_S29HDGHlIp2lfVMsGxpQ5tvAQ/edit#gid=973655443>
* “Pop\_Group.dta”: Data for age-specific population, collected from SEER.
* “states-daily”: State-level COVID-19 testing data collected from the COVID tracking project.
* “Weather”: County-level weather controls from NOAA.

Because of the restricted-use agreements, some of the proprietary data we used in our analyses cannot be posted. However, researchers may contact the following places to obtain the raw data that we used. Codes to clean the raw data are included in the folder.

* Social Distancing & Foot Traffic: <https://www.safegraph.com/>
* Age-specific COVID-19 Case: <https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Restricted-Access-Detai/mbd7-r32t>
* COVID-19 Hospitalization: <https://covid19researchdatabase.org/>
* BLM Protest: Please contact professor Joseph Sabia (jsabia@sdsu.edu) for this dataset

The authors are happy to provide any assistance necessary for users wishing to construct the final data set.

**Code**

The “code” directory contains two types of code. The first type is the code that needs to be executed in a specific order. These files are organized in a way that begins with a number indicating the order of execution (e.g. "01" should be run before "02", "02" should be run before "03", and so on), followed by a brief description of what the code does (e.g. " FT\_clean" indicates code that cleans the foot traffic data). The second type is the code that are called inside another do-files. These do-files are named “YYY”, which describes what the do-file intends to do (e.g. “ES\_Figure Code.do” creates a program that will plot our event study figure). Researchers do not need to execute the second type of files separately, as they will be automatically run when the first type of code is executed.

The Do-File “08\_Run\_TWFE\_Estimate.do” can be used to replicate the following tables and figures in the following order:

* Figures 2a-c. Event Study Analysis for Social Distancing Measure
* Figures 3a-c. Event Study Analysis for Foot Traffic Measure
* Table 1. Difference-in-Differences Estimates for Social Distancing Measure
* Table 2. Difference-in-Differences Estimates for Foot Traffic Measure
* Table 3. Heterogeneity in the Effects on Social Distancing
* Figure 3d. Event Study Analysis for COVID-19 case growth
* Table 4. Difference-in-Differences Estimates for COVID-19 case growth
* Table 5. Difference-in-Differences Estimates for COVID-19 case growth

The R Code “09\_Run\_CS\_Estimate.R” can be used to replicate the following figures in the following order:

* Figure 4. Callaway and Sant’Anna (2021) estimates on COVID-19 growth
* Figure 5. Callaway and Sant’Anna (2021) estimates on age-specific COVID-19 case growth
* Figure 6. Callaway and Sant’Anna (2021) estimates on age-specific COVID-19 hospitalization growth