OSF **REGISTRIES**

Police Violence and the Health of Black Infants

Public 0

Study Information

Title

Provide the working title of your study. It is helpful if this is the same title that you submit for publication of your final manuscript, but it is not a requirement.

Police Violence and the Health of Black Infants

Authors

The author who submits the preregistration is the recipient of the award money and must also be an author of the published manuscript. Additional authors may be added or removed at any time.

Joscha Legewie

Research Questions

Please list each research question included in this study.

What is the effect of maternal exposure to police killings of armed and unarmed Black Americans in the residential environment on the health of black infants in terms of birth weight and gestation age (preterm birth)? How does this effect compare to police killings involving white and Hispanic suspects and how does it differ by mother's race?

Contributors (/x8by4/contributors)

Joscha Legewie (/6emzu)

Description

No description given.

Registration type

Prereg Challenge

Date registered

October 2, 2018

Date created

October 2, 2018

Registered from

osf.io/bthkg (/bthkg)

Category

Project

Registration DOI

Hypotheses

For each of the research questions listed in the previous section, provide one or multiple specific and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here.

I expect a negative effect of maternal in utero exposure to police killings of unarmed Black Americans in the residential environment on the health of black infants in terms of birth weight and gestation age.

I expect that the size this effect declines with spatial distance to the police killing.

I expect a smaller or no effect of maternal in utero exposure to police killings of armed Black Americans in the residential environment on the health of black infants in terms of birth weight and gestation age.

I expect a smaller or no effect of maternal in utero exposure to police killings of White and Hispanic Americans in the residential environment on the health of black infants in terms of birth weight and gestation age.

Sampling Plan

Existing Data

Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation. Please do not hesitate to contact us if you have questions about how to answer this question (prereg@cos.io).

- Registration prior to accessing the data

Explanation of existing data

If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study. The purpose of this question is to assure that the line between confirmatory and exploratory analysis is clear.

This study relies on two data sources. The first are birth certificates for all births in California from 2007 to 2017 assembled by the Center for Health Statistics and Informatics at the California Department of Public Health. These birth records include every birth in California (about half a million per year) with information on the date of birth, birth weight, the obstetric estimate of gestation age, and maternal and paternal characteristics such as full names, race, education and residential address. For the geocoding of addresses on birth records, I will use the ArcGIS Desktop geocoder.
The project PI submitted a data application July 5, 2018. The application was approved on August 8, 2018 and the data is currently being prepared for delivery by Center for Health Statistics and Informatics at the California Department of Public Health. As of today, the PI and other project members had no access to data on California birth records.

The second data source is incident-level data on policing killings in California between 2005 and 2017. This data relies on Fatal Encounters (Burghart 2016) and our own extensive coding. The coding procedure is described below. The coding is currently in progress and will be finalized over the next weeks (see attached file for details on coding process). The result of this procedure is a comprehensive database on officer-involved killings in California between January 2005 and December 2017. The database includes information on the name of the victim, the date of the incident, the geographic location of the incident in terms of x and y coordinates, the race and age of the victim and whether the victim was armed or unarmed.

Data collection procedures

This study relies on two data sources. The first are birth certificates for all births in California from 2007 to 2017. Access to this data is provided by the Center for Health Statistics and Informatics at the California Department of Public Health. The second is incident-level data on all policing killings in California between 2005 and 2017. To collect comprehensive information on police killings, I will rely on Fatal Encounters with my own extensive coding to fill in missing information, correct errors and add additional variables to the database. The goal of Fatal Encounters is to "create a comprehensive national data-base of people who are killed through interactions with police" from 2000 to today (Burghart 2016). It is the most comprehensive data collection project that tracks officer-involved killings over multiple years. The Fatal Encounters archive relies on paid researchers, public records requests and crowd-sourced data. As of September 2018, the archive considers the information complete from January 2005 to December 2017. I will use Fatal Encounters to create a comprehensive dataset on officer-involved killings in California based on four steps.

First, I downloaded FE data on April 8, 2018 and restricted the database to all 2,553 incidents in California between 2005 and 2017.

Second, I use paid researchers to code the type of incident and restrict the cases to "intentional or purposeful police killing" and "unintentional police killing but result of extremely reckless, reckless or negligent use of force". Specifically, I follow Legewie and Fagan (2016) to define officer-involved killings or deaths as "any interaction with the police where the officer uses force and the person dies during or immediately after the interaction. This includes cases that result in death as a consequence of being shot, beaten, restrained, pepper sprayed, tasered, or otherwise harmed by police officers, whether on-duty or off-duty. ... This includes acts where one or more police officers set in motion a chain of events that leads to the death of a suspect or another individual if the original act involved the use of police force. These cases might capture high speed chases or instances where a police officer kills a bystander when

shooting at a suspect. The definition excludes (a) suicides, (b) accidents caused by suspects themselves (e.g. a fleeing suspect who causes a deadly car crash), and (c) police-caused accidents unrelated to the use of force (e.g. a car crash under normal traffic conditions that is not related to the vehicular pursuit of a suspect)." (Legewie and Fagan 2016). At the time of writing this pre-registration, the coding process is not complete. Out of the first 1,143 cases, 743 cases were classified as "Intentional or purposeful", 40 as "Unintentional but result of extremely reckless, reckless or negligent use of force" and the remaining 360 cases were excluded from the database. Excluded cases include duplicates, an incident in which an off-duty police officer shot and killed his wife during a domestic dispute, a case in which a suspect successfully fled to a forest area from the police and was found dead several days later from drowning in a nearby swamp, several cases in which the suspect kills bystanders during a car chase, and several cases in which the deaths are accidental or self-inflicted.

Third, I will verify and completed missing information for individual victims on the following variables: name, age, gender, race, date of the incident, geographic location, and names of the involved police agencies. As part of this verification process, I also added a variable "armed/unarmed" to the database. While the Fatal Encounters database is comprehensive and well-maintained, there are occasional errors and cases with missing information. Most important, information on victim race is missing for about 17.4% of all cases in California and considered unreliable for others. To address this problem, we use two independent coders from Amazon Mechanical Turk to verify and complete the information for all cases. We gave each coder the victim's name, date of the incident, and the state in which it occurred. They were asked to collect information on the race of the victim, the involved police agencies, and the city in which the incident occurred from newspaper sources. A student research assistant cross-validated all cases for which the information from Fatal Encounters and our MTurk coders were not identical. This cross-validation and the comparison between the two independent coders strengthens the quality and accuracy of the data verification process.

Finally, I will use the Google Geocoding API to convert the address of the incident into geographic coordinates (latitude and longitude).

The result of this procedure is a comprehensive database of all officer-involved killings in California between January 2005 and December 2017. The database includes information on the name of the victim, the date of the incident, the geographic location of the incident in terms of x and y coordinates, the race and age of the victim and whether the victim was armed or unarmed.

(optional)

- No files selected

Sample size

Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?

The approximate sample size is 5 million births in California and about 1,500 police killings.

Sample size rationale

This could include a power analysis or an arbitrary constraint such as time, money, or personnel.

N/A

Stopping rule

If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.

N/A

Variables

Manipulated variables

N/A

(optional)

- No files selected

Measured variables

OUTCOME VARIABLES
The analysis will focus on two outcome variables related to infant health: birth weight measured in grams and gestational age measured in weeks based on the obstetric estimate of gestation. These variables are common measures on infant health. Low birth weight and preterm birth are related to infant mortality (Callaghan et al. 2006; Centers for Disease Control and Prevention 2002; Kramer et al. 2000; McCormick 1985) and long-term outcomes such as cognitive development, test scores, ADHD and others (Bhutta et al. 2002). In alternative specifications, I will focus on high risk infants and examine binary indicators for low birth weight (<2,500 grams) and preterm delivery (<37 completed weeks of gestation).

TREATMENT INDICATOR/MAIN INDEPENDENT VARIABLE
To estimate the effect of maternal exposure to police killings, I will construct two terms separately for police killings of armed and unarmed Black Americans. The first is an indicator variable Near that measures proximity of mother's residence at time of birth to any police killing of an armed/unarmed Black American between 2005 and 2017 independent of the timing of the birth and incident. It is coded as 1 if any police killing occurred within a specific distance k and 0 otherwise regardless of when the incident occurred. There is no a priori or theoretical way to derive the correct value of the distance parameter k. Instead, the spatial scale of a potential effect

is an empirical question. Hence, I will estimate models with different values of k ranging from 0.25 to 5 or even more kilometers. I will report multiple specifications to show at which distance a potential effect fades and disappears. Distances will be calculated using great circle distance based on the `st_distance` function in the R package `sf` (which uses the `geod_inverse` function from `proj.4` when the ` proj.4` version is larger than 4.8.0). The Near indicator measures whether at least one police killing occurred within the relevant distance.

The second is a set of 9 indicator variables that measures the timing of maternal exposure to police killings of armed/unarmed Black Americans within a specific distance k. The 9 indicators distinguish between incidents that happened in different time period before birth (similar to lead terms), exposure while in utero by trimester of gestation (the key treatment indicators) and different time periods after conception (similar to lagged terms). All 9 indicator variables are binary and coded as 1 if at least one police killing occurred during the specific time interval and within k kilometers of mother's residence and 0 otherwise. The three lead indicators are born 18 to 12, 12 to 6 and 6 to 1 months before the incident. They are designated as D_j where j ranges from -3 to -1. The effect of the three indicators should be zero because birth occurs before the incident. The terms help us to evaluate the plausibility of the design. The three treatment indicators measure exposure while in utero by trimester of gestation within k kilometers of mother's residence. They are defined as exposure during the 1st trimester D_0, 2nd trimester D_1 and 3rd trimester D_2 of gestation. The three terms are the key treatment indicator and estimate the effect of maternal exposure to police killings by trimester of gestation. Finally, the three lagged terms are defined as conceived 1 to 6, 6 to 12 or 12 to 18 months after a police killing within k kilometers of the incident. While birth outcomes might be affected by exposure to police killings prior to pregnancy, infants are not exposed in utero. Previous research and the most plausible mechanism linking maternal stress to birth outcomes suggest that the effect of the lagged terms should be small. Note that multiple of the 9 indicator variables for the timing of exposure can be coded as 1 if multiple

police killings occurred within k kilometer. This specification makes it possible to examine changes in birth outcomes of black infants before and after police killings of armed/unarmed Black Americans.

CONTROL VARIABLES
Sex of child coded as 0 for male and 1 for female, mother's age (categorized as less than 20, 20–29, 30–34, or 35 years or older), mother's education (categorized as less than high school, high school, some college, college, advanced degree, or missing), and child parity (first, second, third, fourth born or higher, parity missing).

SIBLINGS
To identify full siblings in the sample, I will follow the procedure used by Liu, King, and Bearman (2010) based on California birth records from 1997 to 2007. In particular, I will exact match mother's date of birth, father's date of birth, and the first letter of the mother's maiden name for all children born between 2000 and 2017.

(optional)

- No files selected

Indices

N/A

(optional)

- No files selected

Design Plan

Study type

Please check one of the following statements

- Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, “natural experiments,” and regression discontinuity designs.

Blinding

Blinding describes who is aware of the experimental manipulations within a study. Mark all that apply.

- No blinding is involved in this study.

Study design

Difference-in-Difference (DD) approach with and without sibling comparison

(optional)

- No files selected

Randomization

If you are doing a randomized study, how will you randomize, and at what level?

N/A

Analysis Plan

Statistical models

Note: The attachment includes a formatted version of the model description.

The goal of the analysis is to estimate the effect of maternal exposure to police killings of armed and unarmed Black Americans

on birth outcomes for Black infants. Estimating the effect of police killings is challenging because police killings are not random. They are linked to crime and other neighborhood characteristics, which might also affect birth outcomes. To overcome this challenge, I rely on a Difference-in-Difference (DD) approach with and without sibling comparison (Angrist and Pischke 2008; Meyer 1995). The estimation strategy and model specification is motivated by previous research on the effect of Hydraulic fracturing on infant health (Currie, Greenstone, and Meckel 2017), and the effect of police killings (Ang 2018) and police surges (Legewie and Fagan 2018) on student test scores.

All models will be estimated using a generalization of the within (fixed-effect) estimator for multiple high dimensional categorical variables using the lfe package in R (Gaure 2013). In particular, I will use the within transformation for the neighborhood and year-by-month fixed-effect term and for the mother fixed effect term in the model with sibling comparison.

DIFFERENCE-IN-DIFFERENCE MODEL
The difference-in-difference approach leverages millions of birth records from California to compare changes in birth outcomes for Black infants in exposed areas born in different time periods before and after police killings of Black Americans to changes in birth outcomes for control cases in unaffected areas. In particular, I distinguish nine time period for children who are born 18-12, 12-6 and 6-1 months before a police killing, those who are exposed during the first, second and third trimester in utero, and those who are conceived in the 1-6, 6-12 and 12-18 months after an incident. Formally, I plan to estimate the following regression model with clustered standard errors on the neighborhood level to address potential serial correlation problems (Bertrand, Duflo, and Mullainathan 2004):

$$y_{icjt}=\pi_{cj}+\eta_t+\gamma_c \text{ year}+\beta_1 \text{Near}_i+\sum_{r=-3}^{5} (\delta_r \text{Exposure}_{(i,r)}) + \beta_2 X_{icjt} + \varepsilon_{icjt}$$

where the dependent variables are birth outcomes for mother i in

county c in census tract j and at year-by-month time period t (e.g. January 2014). The model includes a stable neighborhood effect π_cj that controls for mean differences in birth outcomes across census tracts, and a year-by-month time effect η_t that captures differences in birth outcomes over years and months (e.g. January 2014) that are constant across all births such as seasonal fluctuations or general time trends. In addition, the model includes county-specific, linear time trends γ_c year that captures different temporal trends in birth outcomes across California's 58 counties. The vector X_ijt consists of maternal and child characteristics. It includes sex of child coded as 0 for male and 1 for female, mother's age (categorized as less than 20, 20–29, 30–34, or 35 years or older), mother's education (categorized as less than high school, high school, some college, college, advanced degree, or missing), and child parity (first, second, third, fourth born or higher, parity missing).

To estimate the effect of maternal exposure to police killings, the model includes two terms that will be constructed separately for police killings of armed and unarmed Black Americans. The first is an indicator variable Near_i that measures proximity of mother's residence at time of birth to any police killing of an armed/unarmed Black American between 2005 and 2017 independent of the timing of the birth and incident. It is coded as 1 if any police killing occurred within a specific distance k and 0 otherwise regardless of when the incident occurred. There is no a priori or theoretical way to derive the correct value of the distance parameter k. Instead, the spatial scale of a potential effect is an empirical question. Hence, I will estimate models with different values of k ranging from 0.25 to 5 or even more kilometers. I will report multiple specifications to show at which distance a potential effect fades and disappears. Distances will be calculated using great circle distance based on the `st_distance` function in the R package `sf` (which uses the `geod_inverse` function from `proj.4` when the ` proj.4` version is larger than 4.8.0). The Near_i indicator measures whether at least one police killing occurred within the relevant distance.

The second is a set of 9 indicator variables that measures the timing of maternal exposure to police killings of armed/unarmed Black

Americans within a specific distance k (see Table 1 for a description of the 9 terms). The 9 indicators distinguish between incidents that happened in different time period before birth (similar to lead terms), exposure while in utero by trimester of gestation (the key treatment indicators) and different time periods after conception (similar to lagged terms). All 9 indicator variables are binary and coded as 1 if at least one police killing occurred during the specific time interval and within k kilometers of mother's residence and 0 otherwise. The three lead indicators are born 18 to 12, 12 to 6 and 6 to 1 months before the incident. They are designated as D_j where j ranges from -3 to -1. The effect of the three indicators should be zero because birth occurs before the incident. The terms help us to evaluate the plausibility of the design. The three treatment indicators measure exposure while in utero by trimester of gestation within k kilometers of mother's residence. They are defined as exposure during the 1st trimester D_0, 2nd trimester D_1 and 3rd trimester D_2 of gestation. The three terms are the key treatment indicator and estimate the effect of maternal exposure to police killings by trimester of gestation. Finally, the three lagged terms are defined as conceived 1 to 6, 6 to 12 or 12 to 18 months after a police killing within k kilometers of the incident. While birth outcomes might be affected by exposure to police killings prior to pregnancy, infants are not exposed in utero. Previous research and the most plausible mechanism linking maternal stress to birth outcomes suggest that the effect of the lagged terms should be small. Note that multiple of the 9 indicator variables for the timing of exposure can be coded as 1 if multiple police killings occurred within k kilometer. This specification makes it possible to examine changes in birth outcomes of black infants before and after police killings of armed/unarmed Black Americans.

I will compare the estimates for maternal exposure to police killings of Black Americans with similar estimates for killings of white and Hispanic victims on the health of Black, white and Hispanic infants. For this purpose, I will construct separate exposure indicators for police killings involving Black, white and Hispanic victims and estimate the models separately for the sample of black, white and Hispanic infants.

TABLE 1: Indicator variables for maternal exposure to police killing

Term	Description
D_(-3)	Born 12-18 months before incident within k km
D_(-2)	Born 6-12 months before incident within k km
D_(-1)	Born 1-6 months before incident within k km
D_0	Exposure within k km during 1st trimester of gestation
D_1	Exposure within k km during 2nd trimester of gestation
D_2	Exposure within k km during 3rd trimester of gestation
D_3	Conceived/LMP 1-6 months after incident within k km
D_4	Conceived/LMP 6-12 months after incident within k km
D_5	Conceived/LMP 12-18 months after incident within k km

SIBLING COMPARISON

My second model will use a sibling comparison based on the following specification

$$y_{icjt}=\alpha_i+\pi_{cj}+\eta_t+\gamma_c \text{ year}+\beta_1 \text{Near}_i+\sum_{r=-3}^{5} (\delta_r \text{Exposure}_{(i,r)}) + \beta_2 X_{icjt}+\varepsilon_{icjt}$$

The individual-level fixed effect term α_i accounts for all observed and unobserved characteristics of the mother that are constant over time. As a consequence, the vector of covariates now only includes child's sex and excludes all characteristics that are constant or almost constant over time.

PLAUSIBILITY OF ESTIMATION STRATEGY AND CRIME AS AN ALTERNATIVE EXPLANATION

The core assumption of the DD approach is that in the absence of police killings changes in birth outcomes of exposed children would have been the same as changes in birth outcomes of children in control areas (common trend assumption). The estimates of δ_r for the lag ($r< 0$) and to some extend lead ($r> 2$) terms make it possible to evaluate the common trend assumption indirectly.

However, time-specific, neighborhood-level factors that are related to police killings and birth outcomes still threaten the identification strategy. Most importantly, unobserved trends in crime rates or

policing may account for both police shootings and changes in birth outcomes. From this perspective, a potential effect on birth outcomes is not caused by exposure to officer-involved killings but instead by changes in crime rates that increase police activity and officer-involved killings. While the model accounts for county-specific linear trends in birth outcomes and the specification of the exposure variable focuses on the nine months of pregnancy, local crime waves (as opposed to general trends) that coincide with police killings are a potential threat to my identification strategy particularly because previous research links crime with birth outcomes (Torche and Villarreal 2014). However, previous research also shows that police killings are unrelated to homicides, violent crimes and arrests after accounting for neighborhood and time fixed effects similar to the ones used in my model. In particular, Ang (2018) uses data from Los Angeles to examine the relation between police killings and homicides, violent crimes and arrests. He finds no evidence for changes in crime or arrests prior to or following police killings. This finding supports the argument that police killings are exogenous after conditioning on neighborhood and time fixed effects.

In addition, the comparison between the effect of maternal exposure to police killings of armed and unarmed Black Americans provides further evidence. In particular, the alternative explanation of a potential effect based on crime as a confounder has opposite implications for the type of police killings that are particularly consequential. If police killings affect birth outcomes, the effect of incidents that are perceived as unjustified involving a suspect that pose little threat should be more pronounced. However, if crime is an unobserved confounder that biases the results, justified police killings related to serious crime should have a larger effect on birth outcomes. I will exploit these contradicting implications to rule out crime as an alternative explanation. In particular, I will compare the effect of maternal exposure for incidents involving armed and unarmed suspect as a measure of the threat posed by suspects.

(optional)

- pre-analysis-plan-v3-statistical-models.pdf

(/project/x8by4/files/osfstorage/5bb3b72a6ca1d00016b9f7e1/)

Transformations

If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.

The control variables include sex of child coded as 0 for male and 1 for female, mother's age (categorized as less than 20, 20–29, 30–34, or 35 years or older), mother's education (categorized as less than high school, high school, some college, college, advanced degree, or missing), and child parity (first, second, third, fourth born or higher, parity missing).

Follow-up analyses

If not specified previously, will you be conducting any confirmatory analyses to follow up on effects in your statistical model, such as subgroup analyses, pairwise or complex contrasts, or follow-up tests from interactions? Remember that any analyses not specified in this research plan must be noted as exploratory.

N/A

Inference criteria

What criteria will you use to make inferences? Please describe the information you'll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?

Thresholds of 0.05 and 0.01 for p-values will be reported but I will use effect sizes, standard errors and p-values to evaluate the results.

Data exclusion

How will you determine which data points or samples (if any) to exclude from your analyses? How will outliers be handled?

I will restrict the sample to Black mothers and singleton births considering that twins and other multiples have lower birth weight and worse health at birth. In later steps of the analysis, I will compare the results to white and Hispanic mothers.

Missing data

How will you deal with incomplete or missing data?

All analysis will exclude cases with missing information on birth date, birth weight, gestation age, mother's race or mother's residential address.

Exploratory analysis

If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time. (optional)

N/A

Scripts

Upload an analysis script with clear comments

This optional step is helpful in order to create a process that is completely transparent and increase the likelihood that your analysis can be replicated. We recommend that you run the code on a simulated dataset in order to check that it will run without errors. (optional)

- No files selected

Other

Other

If there is any additional information that you feel needs to be included in your preregistration, please enter it here. (optional)

Ang, Desmond. 2018. “The Effects of Police Violence on Inner-City Students.” 75.

Angrist, Joshua D. and Jörn-Steffen Pischke. 2008. Mostly Harmless Econometrics: An Empiricist’s Companion. Princeton: Princeton University Press.

Bhutta, Adnan T., Mario A. Cleves, Patrick H. Casey, Mary M. Cradock, and K. J. S. Anand. 2002. “Cognitive and Behavioral Outcomes of School-Aged Children Who Were Born Preterm: A Meta-Analysis.” JAMA 288(6):728–37.

Burghart, Brian. 2016. “Fatal Encounters.”

Callaghan, William M., Marian F. MacDorman, Sonja A. Rasmussen, Cheng Qin, and Eve M. Lackritz. 2006. “The Contribution of Preterm Birth to Infant Mortality Rates in the United States.” Pediatrics 118(4):1566–73.

Centers for Disease Control and Prevention. 2002. “Infant Mortality and Low Birth Weight among Black and White Infants--United States, 1980-2000.” MMWR. Morbidity and Mortality Weekly Report 51(27):589–92.

Currie, Janet, Michael Greenstone, and Katherine Meckel. 2017. “Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania.” Science Advances 3(12):e1603021.

Gaure, Simen. 2013. “OLS with Multiple High Dimensional Category Variables.” Computational Statistics & Data Analysis 66:8–18.

Kramer, Michael S., Kitaw Demissie, Hong Yang, Robert W. Platt, Reg Sauvé, Robert Liston, and for the Fetal and Infant Health Study Group of the Canadian Perinatal Surveillance System. 2000. “The

Contribution of Mild and Moderate Preterm Birth to Infant Mortality." JAMA 284(7):843–49.

Legewie, Joscha and Jeffrey Fagan. 2016. "Group Threat, Police Officer Diversity and the Deadly Use of Police Force."

Legewie, Joscha and Jeffrey Fagan. 2018. "Aggressive Policing and the Educational Performance of Minority Youth."

McCormick, Marie C. 1985. "The Contribution of Low Birth Weight to Infant Mortality and Childhood Morbidity." New England Journal of Medicine 312(2):82–90.

Meyer, Bruce D. 1995. "Natural and Quasi-Experiments in Economics." Journal of Business & Economic Statistics 13(2):151–61.

Torche, Florencia and Andrés Villarreal. 2014. "Prenatal Exposure to Violence and Birth Weight in Mexico. Selectivity, Exposure, and Behavioral Responses." American Sociological Review 966–92.

(/) OSF **REGISTRIES**

Police Violence and the Health of Black Infants

Public 0

Summary

Summary

Provide a narrative summary of what is contained in this registration, or how it differs from prior registrations. (optional)

This document outlines four small revisions to the pre-analysis plan submitted on October 2, 2018, and available at https://osf.io/x8by4/register/565fb3678c5e4a66b5582f67. The revisions were written prior to obtaining access to geo-coded information on all births in California and therefore before the central treatment indicator – maternal exposure to police killings in the residential environment – could be constructed. The revisions are based on access to other information on birth records aside from geographic coordinates, which makes it possible to clarify some coding decisions. In particular, the PI received the data from the Center for Health Statistics and Informatics at the California Department of Public Health on October 10, 2018, and I expect to receive the geographic coordinates on October 22 or 23, 2018.

The pre-analysis plan outlines a strategy to estimate the effect of maternal exposure to police violence on infant health. This revision to the pre-analysis plan includes four changes:

1. The analysis will focus on birth records from 2007-2016 because the data delivery did not include data from 2017.

2. Sample restriction: California birth records include information on

Contributors (/exzd5/contributors)

Joscha Legewie (/6emzu)

Description

No description given.

Registration type

Open-Ended Registration

Date registered

October 19, 2018

Date created

October 19, 2018

Registered from

osf.io/bthkg (/bthkg)

Category

Project

Registration

all births in California and out-of-state births to California residents. I will restrict the analysis to singleton, in-state births to California residents. This excludes births in California to mothers who life in a different state and out-of-state births to California residents. Note that the pre-analysis plan already restricted the sample to singleton births. As discussed in the pre-analysis plan, the analysis also excludes cases with missing information on birth date, birth weight, gestation age, mother's age, mother's race/ethnicity or mother's residential ad-dress (the only addition to this list is mother's age, which is missing for 25 out of over 4 million births). With these restrictions, the sample size is 4,041,831 births (prior to excluding births with incorrect information on mother's residential address).

3. Recoding: Based on access to the documentation and part of the data, I want to clarify the coding for the following variables:

a. Mother's Race/Ethnicity: The three relevant racial/ethnic groups will be coded as non-Hispanic White, non-Hispanic Black and Hispanic. Mothers with multiple race categories are excluded from these three groups unless they are Hispanic.

b. Mother's education: The revised coding uses six categories for "less than high school", "high school degree", "some college or associate degree", "Bachelor's degree", "advanced degree" (master, doctorate, or professional degree), and "missing". The only change compared to the pre-analysis plan is the combination of the category "some college" and "associate degree".

4. Siblings: The approach to identify full siblings remains the same but I will exclude birth records with missing information on any of the variables used in the matching procedure (most importantly, father's birth date).

DOI

Affiliated institutions

This registration has no affiliated institutions

License

No license

Tags

No tags

Add a tag to enhar

Citation

osf.io/exzd5

g/blob/master/PRIVACY_POLICY.md) | Status (https://status.cos.io/) |
API (https://developer.osf.io/)
TOP Guidelines (http://cos.io/top/) | Reproducibility Project:
Psychology (https://osf.io/ezcuj/wiki/home/) |
Reproducibility Project: Cancer Biology
(https://osf.io/e81xl/wiki/home/)

(http://twitter.com/OSFramework)
(https://www.facebook.com/CenterForOpenScience/)
(https://groups.google.com/forum/#!forum/opensciencefframework)
(https://www.github.com/centerforopenscience)