NBER WORKING PAPER SERIES

POLICING THE POLICE:
THE IMPACT OF "PATTERN-OR-PRACTICE" INVESTIGATIONS ON CRIME

Tanaya Devi
Roland G. Fryer Jr

Policing the Police: The Impact of "Pattern-or-Practice" Investigations on Crime
Tanaya Devi and Roland G. Fryer Jr
NBER Working Paper No. 27324
June 2020
JEL No. J0,J15,K10

ABSTRACT

This paper provides the first empirical examination of the impact of federal and state "Pattern-or-Practice" investigations on crime and policing. For investigations that were not preceded by "viral" incidents of deadly force, investigations, on average, led to a statistically significant reduction in homicides and total crime. In stark contrast, all investigations that were preceded by "viral" incidents of deadly force have led to a large and statistically significant increase in homicides and total crime. We estimate that these investigations caused almost 900 excess homicides and almost 34,000 excess felonies. The leading hypothesis for why these investigations increase homicides and total crime is an abrupt change in the quantity of policing activity. In Chicago, the number of police-civilian interactions decreased by almost 90% in the month after the investigation was announced. In Riverside CA, interactions decreased 54%. In St. Louis, self-initiated police activities declined by 46%. Other theories we test such as changes in community trust or the aggressiveness of consent decrees associated with investigations -- all contradict the data in important ways.

Tanaya Devi
Harvard University
Cambridge, MA 02138
tanayadevi01@fas.harvard.edu

Roland G. Fryer Jr
Department of Economics
Harvard University
Littauer Center 208
Cambridge, MA 02138
and NBER
rfryer@fas.harvard.edu

A data appendix is available at http://www.nber.org/data-appendix/w27324

1 Introduction

Police use of force has become one of the most important and divisive social issues of our time. A spate of recent recordings of seemingly excessive force – from Laquan McDonald who was shot sixteen times in fifteen seconds while jaywalking in Chicago and ignoring police directives to stop; to Alton Sterling who was shot several times at point blank range while held down by two Baton Rouge police officers – has caused riots, peaceful and other forms of protests, and a national discussion on race and policing in America. "Black Lives Matter," a movement which campaigns against violence and systemic racism towards black people, has become a cultural idiom for the necessity of addressing police use of force and other forms of inequality in America.

A key tool the government uses to combat unconstitutional policing – which includes, but is not limited to, excessive force and racial bias – is called a "Pattern-or-Practice" investigation. Yet, how to investigate a police department with unknown bias is a challenging economic problem. Theoretically, there can be one of three effects. If the investigation scrutinizes systemic misconduct such as corruption, but does not influence everyday policing, then an investigation can eliminate unconstitutional behaviors without an impact on crime rates.[1] Conversely, if an investigation scrutinizes deficiencies that lead to more efficient policing (e.g. better data collection or optimization of current resources), then an investigation may lead to a reduction in crime. Finally, if a police investigation raises the costs of the type of policing that suppresses crime or, in any way, increases the price for individual officers of making type I errors (i.e. shooting someone who they believed to be carrying a weapon, but was not), then officers may substitute effort between tasks or reduce total effort in a way that causes crime to increase.

In this paper, we provide the first thorough empirical examination of the impact of federal and state "pattern-or-practice" investigations (hereafter referred to simply as investigations) on crime and policing – exploiting the staggered timing of investigations to credibly identify a series of causal effects. We collect four separate datasets for our main analysis: (1) data on pattern-or-practice investigations conducted by federal and state governments; (2) data on monthly crime rates from the FBI Uniform Crime Reports (UCR); (3) policing data gathered from 15 police departments, mainly through Freedom of Information Act (FOIA) requests; and (4) data from the decennial census used to identify synthetic control cities. In an effort to unearth potential mechanisms, we

---

[1]For instance, in the Newark Police Department, officers were accused of theft and corruption when they came into contact with individuals with large sums of money. Further, if the investigation fails to identify a pattern-or-practice of police misconduct – which has happened in 25 of the 73 cases – then this will likely have no impact on policing or crime.

also collected data on media attention for all cities in our sample.

Perhaps the simplest way to understand the pre-processed data is to estimate whether there are breaks (either in level or trend) in the month the investigation is announced. For homicides, there is neither a level nor trend break in the month of the event – suggesting little impacts of investigations on homicides in the raw data. Total crime demonstrates a small level break that is statistically significant.

We explore investigation specific heterogeneity by plotting the size of the level break for each investigation city (in the month of the event), by increasing size of the break. For the twenty-seven investigations that occurred in "principal" cities,[2] there were five with statistically positive level breaks in homicides, one with a statistically negative level break and twenty-one that are statistically zero.[3]

What was different between the five investigations that sustained a positive level break versus the other twenty-two that, on average, underwent no significant changes in homicide rates?

We correlate the size of the level break in crime (homicides and total crime), weighted by the precision of the estimate, with several demographics at the city level – total population, percent population between 15 and 24 years of age, percent of different races present in the city, median income, poverty rate, percent with public assistance income, percent of high school graduates and four year college graduates, isolation index, percent households with single parents, and unemployment rate. Of all demographics considered, three – total population, percent on public assistance income and percent with four year college graduate degrees – are significantly correlated with the size of the level break. After adjusting the p-values for multiple hypothesis testing, however, none of them are significant.

We also separate the investigations by the most recent event that seems to have caused the investigation. Five cities had investigations that quickly succeeded a "viral" incident where law enforcement officials used deadly force against an African-American civilian. These incidents caught national media attention and the cities witnessed protests and riots soon after. These cities are Baltimore, MD, Chicago, IL, Cincinnati, OH, Riverside, CA, and Ferguson, MO. Other cities' investigations were largely brought about by civilian complaints and allegations about excessive use of force.

---

[2]We limit to principal cities because matching the polcie departments to the correct city demographics is a process that is done manually, and identifying the police department that serves over 22,000 "census designated places" (which include smaller cities, towns, etc.) was infeasible.

[3]Throughout the paper we focus on homicide rates, given the reliability of these data relative to other crimes reported in the FBI Uniform Crime Reports (Jencks, 2001).

For the twenty-two investigations not preceded by viral incidents of deadly force, neither a level nor a trend break in homicide or total crime is observed – crime rates steadily decrease through the investigation month and continue for the two years after.

Investigations that shortly began after viral incidents of deadly force have a different pattern. There are large level breaks in both homicides and total crime. For the twenty-seven months before the investigation, homicides seem to be slightly increasing up to the month of the investigation. Then, there is a large level break (p-value = .000) in homicide rates, and, strikingly, no decline in homicide rates for two years following the announcement of the investigation. Total crime follows a similar, though distinct pattern. For the months preceding the investigation, total crime is decreasing and, again, demonstrates a large level break (p-value = .000) in the month the investigation is announced. Two years after the investigation, total crime continues to be markedly higher. Moreover, these level breaks still exist if we disaggregate crime rates to the bi-weekly, weekly, or daily level.

A concern with our analysis of the raw data is that other cities may have also experienced increases in crime rates in the post-investigation months (i.e. the increase in crime rates is a representation of a separate national trend.) To address this and help understand the cumulative impact of investigations, our main empirical strategy compares the average difference in felony crimes in investigation cities with the average difference in felony crimes for synthetic control cities constructed for each investigation city. Results using synthetic controls are consistent with what we find in the raw data.[4] Investigations not preceded by viral incidents of deadly force, on average, have a negative and statistically significant impact on homicide rates and total felony crime. The cumulative amount of crime that we estimate due to pattern-or-practice investigations in the two years after the announcement of an investigation is -6.39 (2.43) per 100,000 for homicides and -750.38 (431.05) per 100,000 for all felony crimes.

For investigations that were preceded by a viral incident of deadly force – Baltimore, Chicago, Cincinnati, Riverside and Ferguson – there is a marked increase in both homicide and total crime. The cumulative amount of crime that we estimate due to pattern-or-practice investigations in the two years after the announcement for this sample is 21.10 (5.54) per 100,000 for homicides and 1191.77 (429.50) per 100,000 for total felony crime. Put plainly, the causal effect of the investigations in these five cities – triggered mainly by the deaths of Freddie Gray, Laquan McDonald, Timothy

---

[4]Our findings are also consistent with other empirical models we tested such as Least Squares or Propensity Score Matching. See Tables 1 and 2.

Thomas, Tyisha Miller and Michael Brown at the hands of police – has resulted in 893 more homicides than would have been expected with no investigation and more than 33,472 additional felony crimes, relative to synthetic control cities.

To get a sense of how large this number is, the average number of fatal shootings of African-American civilians by police officers in Baltimore, Chicago, Cincinnati, Riverside and Saint Louis, per year, is 12.[5] Thus, even if investigations cured these cities of all future civilian casualties at the hands of police, it would take approximately 75 years to "break even." Our estimates suggest that investigating police departments after viral incidents of police violence is responsible for approximately 450 excess homicides per year. This is 2x the loss of life in the line of duty for the US Military in a year, 12.6x the annual loss of life due to school shootings, and 3x the loss of life due to lynchings between 1882 and 1901 – the most gruesome years.[6]

We explore the validity of our identification assumption by examining two alternative explanations for our findings. Perhaps the most natural alternative explanation is that partitioning the data based on viral incidents of deadly force on African-American civilians is more about the nature of police-civilian interactions which may be altered as a result of societal changes that succeed unsettling displays of police brutality. In this case, the stark results on investigations following infamous deadly force incidents are less about investigations and more about what happens anytime there is controversial use of force against an African-American civilian by police. To better understand this, we conduct parallel analyses for cities that had viral shootings but no investigation. There are no significant level breaks for homicides or total crime for viral shootings.

A second alternative explanation is that our estimates are obtained from a unique set of cities – rife with discrimination and tension between police and communities – and that any shootings of unarmed African-American men in these cities results in increases in crime whether or not they are accompanied by an investigation. Analyzing crime patterns after controversial shootings in Baltimore, Chicago, and Cincinnati in the 30 months leading up to the investigation but before the viral incident of deadly force that caused the investigation, yields no level breaks in crime.[7]

Still, some will wonder about the validity of our conclusions based on just five cities. If these five cities were an anamoly, the exact p-values calculated vis-a-vis permutation tests would render

---

[5]Authors' calculation is based on data from the Washington Post (Post 2019) using data from 2015 through 2018.

[6]Authors' calculation is based on data from the Congressional Research Service's 2019 report on Active-Duty Military Deaths (Congressional Research Service, 2019), CNN's 2019 report on school shootings (CNN, 2019), and statistics provided by the Archives at Tuskegee Institute at http://law2.umkc.edu/faculty/projects/ftrials/shipp/lynchingyear.html.

[7]Currently, we do not have data on officer involved shootings from Riverside Police Department for the investigation year or the years preceding it.

our estimates insignificant – they don't even border insignificance. Thus, given the magnitude of the results there seem to be two related conclusions one can draw from the data: (1) these estimates from five cities are significant but they are estimated on "bubbles" in time and future investigations, even ones that are conducted on the basis of discriminatory policing only, will not have such negative impacts; or (2) these estimates from five cities are significant and are likely predictive of the impact of future investigations into the forseeable future because the tension between minority communities and police has reached a tipping point. Our data cannot differentiate between these two, but will be known in the fullness of time.

Understanding the mechanism through which investigations increase crime is (necessarily) more speculative. The leading hypothesis for why investigations increase crime is that they cause an abrupt change in the quantity of policing activity. For investigations that were not preceded by viral incidents of deadly use of force, investigations had no impact on the quantity of policing. The same holds true for viral shootings that had no investigation. Chicago, Riverside, and St. Louis, where we have the most comprehensive data, show a distinctly different pattern. Using daily data on all recorded police-civilian contacts in Chicago, the day before the investigation was announced there were 1132 police-civilian contacts; 1079 a day after; 838 a week after; and 359 three weeks after. Using monthly data on pedestrian and traffic stops in Riverside, officers had 54% less contact with civilians in the month immediately following the investigation. Using biweekly data on police officers' self-initiated activities in St. Louis, officers conducted 7507 self initiated activities in the first half of the month of the investigation. This declined to 4770 self initiated activities four weeks later and 3482 activities ten weeks after the investigation was announced.[8] We also demonstrate that geographic divisions within Chicago and St. Louis that witnessed the largest reductions in police activity, after the investigation, also witnessed the largest increases in crime after the investigation (see Figure 11).

Other hypotheses we consider, such as decreased community trust and engagement or changes in the nature or extent of the consent decrees associated with investigations, all contradict the data in important ways.

The paper closest to ours is Mas (2006), who uses an event-study design to estimate the impact

---

[8]An often articulated reason for the decrease in the quantity of police-civilian contacts is officers' fear of national media. In almost every focus group we conducted with police officers, "I don't want to be the next 'viral' sensation," was a familiar refrain. The quantity of media differs substantially between cities investigated after incidents of deadly use of force (who average 45.86 articles in the month the investigation is announced) and cities investigated without viral incidents of deadly use of force (7.47 on average).

of final-offer arbitration for police unions on crime.[9] In the months after police officers lose in arbitration, arrest rates and average sentence lengths decline, and crime reports rise relative to when they win. And, these declines in performance are larger when the awarded wage is further from the police unions demand.

Another important contribution is Shi (2009) who tracks monthly felony and misdemeanor arrests and felony crimes before and after the infamous shooting of Timothy Thomas in April 2001, subsequent riots and the federal investigation into Cincinnati Police Departments in May 2001. He finds that average monthly felony arrests decreased by 19 percent while average misdemeanor arrests decreased by 27.8 percent after April 2001. He also determines that monthly felony crime increases by 16 percent after the riot. However, given that the regressions are conducted using a dummy variable that is 1 for all post riot months and 0 otherwise, it is hard to identify precisely whether the change in arrests and crimes are caused due to the riots or the investigation.

The next section provides a brief history of pattern-or-practice investigations and describes the data used in our analysis. Section 3 gives detail on our data and research design. Section 4 provides estimates of the impact of investigations on crime rates; section 5 explores alternative explanations for our findings. Section 6 attempts to unearth mechanisms that may explain why recent investigations increase crime. The final section concludes. There are two appendices: a data appendix which provides details on how we code variables and construct our samples and an online appendix that contains empirical results omitted from the main analysis.

2 Pattern-or-Practice Investigations

On May 3, 1991, America witnessed – after a high speed chase – Rodney King being beaten by four Los Angeles police officers while a dozen others watched. The officers were charged with assault with a deadly weapon and use of excessive force – all were acquitted. Within hours of the acquittal, the 1992 Los Angeles riots began. The rioting lasted six days – sixty three people were killed and thousands injured (Post, 1992).[10] An independent commission linked the beating of

---

[9]Our paper lies at the intersection of three literatures: an established literature in contract theory on regulation (e.g. Baron and Myerson 1982, Laffont and Tirole 1986, Prendergast 2002), a small literature in economics on the policing response to "oversight" (Shi 2009), and a burgeoning literature in criminology of the so-called "Ferguson Effect" (Shjarback, et al 2017 and Pyrooz, et al. 2016). It is also related to qualitative reports on individual investigations. Stone, et al. (2009) analyze policing in Los Angeles as a result of its investigation and is often used as an exemplar of the positive effects of federal investigations on policing and crime. Davis, et al. (2005) describes how the Pittsburgh police department responded to their investigation and subsequent consent decree.

[10]See Fryer (2020) for the effect of the Rodney King viral video and the acquittal of Los Angeles Police Department officers on crime rates.

King to institutional failure within the Los Angeles Police Department (LAPD) and Congress held hearings on how the federal government could do more to address police misconduct (Christopher Commission Report, 1991).

In 1994, Congress passed the Violent Crime Control and Law Enforcement Act which authorized the Attorney General to investigate and litigate cases involving a "pattern or practice of conduct by law enforcement officers" that violates the constitution or federal rights.[11] Under this authority, the Civil Rights Division of the Department of Justice may obtain a court order requiring state or local law enforcement agencies to address institutional failures that cause systemic police misconduct. Pattern-or-practice cases are investigated, litigated, and resolved by the Special Litigation Section of the Civil Rights Division of the Department of Justice.

A typical investigation has the following arc. The first step involves a process by which staff of the Civil Rights Division decide whether to open an investigation into a particular law enforcement agency. In making the decision the types of determinative questions are: (1) Would the allegations, if proven, establish a violation of the Constitution or federal laws?; and (2) would the allegations, if proven, constitute a pattern or practice, as opposed to a sporadic or isolated, violation of the Constitution or federal laws. In answering these questions, staff in the Civil Rights Division examine publicly available information as well as confidential information provided to the Division by witnesses and complainants which it regularly receives from affected community members and families, advocacy groups, prosecutors and defense attorneys, judges, legislators, and police officers with knowledge of misconduct. In some cases, law enforcement agencies or local government officials request to be investigated.

The second step is to prioritize among the set of viable investigations. The Civil Rights Division reports that many more jurisdictions meet the basic criteria to be investigated but that they do not have the resources to investigate them all (Civil Rights Division, 2017). A range of metrics are reportedly used to choose which cities to investigate, including whether the issues a city is dealing with are common across other law enforcement agencies and thus an investigation can provide a model of reform for other jurisdictions – or whether other tools, such as civil rights law suits aimed at individual officers, are better suited to address the issues in a particular law enforcement agency. Once an investigation is initiated, it involves the following steps.

First, immediately following the opening of an investigation, officials from the Civil Rights Division meet with the law enforcement leadership, local political leadership, police labor unions

---

[11]42 U.S.C. §14141 (a)

and affinity groups, and local community groups to explain the basis for the investigation, preview what the investigation will involve, and explain the next steps in the process. Second, officials review written policies, procedures, and training materials relevant to the scope of the investigation, through requests for documents shared with the law enforcement agency. Third, they review systems for monitoring and supervising individual officers, and for holding individual officers accountable for misconduct, including the handling of misconduct complaints, systems for reviewing arrests, searches, or uses of force, and officer disciplinary systems. Fourth, officials observe officer training sessions, participate in ride-alongs with officers on patrol in varying precincts or districts to view policing on the ground and obtain the perspective of officers on the job, and inspect police stations, including lock-up facilities. Fifth, officials analyze incident-related data (i.e. arrests and force reports, disciplinary records, misconduct complaints and investigations, and data documenting stops, searches, arrests and uses of force), often sampling data for large districts, and review the law enforcement agency's system for collecting and analyzing data to identify and correct problems. Sixth, they conduct interviews with police command staff and officers at all levels of authority in the department, both current and former, representatives of police labor organizations and other officer affinity groups, community groups and persons who have been victims of police misconduct, and local government leadership. In brief, during an investigation, the Department of Justice examines complaints, scrutinizes past data and contemporary interactions between law enforcement officials and civilians to determine if police departments have engaged in a pattern or practice of civil rights violations. Importantly, given the heightened scrutiny, if officers believe that any misstep can be amplified into a serious violation by the DOJ, then they might engage in fewer interactions to minimize such missteps.

The first pattern-or-practice investigation occurred in Torrance, CA in May 1995 for excessive use of force, unlawful stops, searches and seizures. The Civil Rights Division did not find any pattern-or-practice of unconstitutional policing – which has been the result in almost forty percent of investigations. The first investigation to find a pattern-or-practice (third overall) was initiated in Pittsburgh in April 1996. In January 1997, after almost a year-long investigation, the Civil Rights Division issued a "findings letter" to the city of Pittsburgh "finding a pattern or practice of excessive force, false arrests, and improper searches and seizures, grounded in a lack of adequate discipline for misconduct and a failure to supervise officers." The parties negotiated a resolution and jointly entered a court-ordered reform agreement overseen by an independent monitor that was in effect from April 1997 to September 2002, with ongoing monitoring through 2005. 61% of federal

investigations find a pattern or practice of unconstitutional policing and end with agreements via a consent decree or memorandum of agreement.

Since the initial investigation in Torrance, CA, there have been 69 federal investigations. Appendix Table 1 provides a list of all cases and a brief summary of the findings; 48 investigations investigated excessive use of force; another 15 investigated discriminatory policing. Other types of investigations include investigative holds – the practice of detaining people in jail without a warrant or probable cause – and gender bias in sexual assault cases.

The Violent Crime Control and Law Enforcement Act of 1994 limits the power to initiate investigations into police departments to the U.S Attorney General only. The only exception to this rule is California, which in 2000 became the first and only state to statutorily authorize its attorney general to address police misconduct (Maxwell and Solomon, 2018). Using the California Constitution (article V, section 13) and California Civil Code section 52.3, the California attorney general has been able to launch 4 pattern or practice investigations in the last 20 years. Two of these investigations – Riverside and Maywood – are already concluded and 2 of them – Kern County and Bakersfield – are still on-going (California Department of Justice, 2016). More details about these investigations are also included in Appendix Table 1.

3 Data, Research Design, and Descriptive Statistics

A. DATA

There are several sources of data used in our analysis: (1) data on pattern-or-practice investigations conducted by federal and state governments; (2) data on monthly crime rates from the FBI Uniform Crime Reports (UCR); (3) policing data gathered from 15 police departments, mainly through Freedom of Information Act (FOIA) requests; and (4) data from the decennial census used to identify synthetic control cities.[12] Policing data is collected to try and ferret out mechanisms. These data are a bit sporadic as they are only available for a small fraction of police departments that were investigated. We briefly describe each in turn. The data appendix contains more details.

Data on pattern-or-practice investigations were obtained from the Department of Justice and media reports.[13] Recall, there have been 73 investigations – 69 initiated by the U.S. Attorney

---

[12]Following Mas (2006), we attempted to use data from the Offender Based Transaction Statistics. These data track individuals arrested for felony crimes through the courts and, if convicted, the sentence. Unfortunately, this data exists for a subset of states between 1979 and 1990 only.

[13]Available at http://apps.frontline.org/fixingtheforce/ and http://oag.ca.gov/news

10

General and 4 by the California Attorney General. We drop 27 investigations because of missing data. For instance, Alabaster, AL Police Department did not report crime data to the FBI from 2000-2005 which coincides with the time period they were investigated. We drop 2 investigations because their event windows overlapped with a previous investigation in the same department and and 2 more because they investigated the department's canine unit and drug interdiction unit only. Thus, our final sample consists of 42 police departments; 15 of which are sheriffs' departments and cities that are not principal cities (i.e. small cities). In all results, we combine Ferguson, MO and St. Louis, MO, in part because Ferguson is not a principal city, and in part because both police departments were deeply interconnected in the melee that followed the fatal shooting of Michael Brown.[14]

Data on felony crime rates are available monthly on a city-level basis from the FBI's Uniform Crime Reports (UCR). We include criminal homicide, forcible rape, robbery, aggravated assault, burglary, and larceny. We do not include motor vehicle theft because it is sparsely available for more recent investigations. It is important to note: UCR data only includes those crimes reported to police. We focus on homicide rates, given the reliability of these data relative to other crimes reported in the FBI Uniform Crime Reports, but also show results for total felony crime.[15] UCR data is currently available through December 2016, although data for 2017 is available on an open website.[16]

To obtain policing data, we contacted 26 police departments to obtain data on police-civilian interactions, 911 call volume, and 911 response times. Data on police activity includes date, time of interaction, and location of interaction for a subset of cities. Data on 911 calls for service include date of call and for St. Louis only – response times and radio codes. Appendix Table 2 describes every request made of any police department and the outcome of that request.

We use city-level control variables at several points in our analysis, most importantly to calculate synthetic control cities using the procedure described in Abadie, et al. (2010). All city-level control variables are taken from the 1990 Census Long Form, the 2000 Census Long Form, and the 2011 ACS 5-year sample. In addition, we create three indices – race index, economic index, and education index – that are used as controls in our main specification. The variable used to calculate race index is percent minority (non-white). The variables included in economic index are median income,

---

[14]We also provide crime trends for only Ferguson in Appendix Figure 4.

[15]Jencks (2001) reports that homicide rates are better indicators of crime than other offenses as they are easy to define and hard to conceal.

[16]Thanks to Jacob Kaplan at the University of Pennsylvania for making the data publicly available at www.openiscpsr.org.

poverty rate, percent receiving public assistance income, unemployment rate, and percent single parent households. The education index consists of percent with at least a high school degree and percent with at least a 4-year college degree. We standardize all input variables across the distribution of principal cities and take the mean of all input variables to create any given index.

B. RESEARCH DESIGN

The models estimated in this paper are identified off the staggered timing of federal and state pattern or practice investigations. For each investigation, we construct an event window of length (μ_1, μ_2), which consists of the investigation month, the μ_1 months before the investigation and the μ_2 months after the investigation. In our typical specification, $\mu_1 = -27$ and $\mu_2 = 24$. The choice of μ_i is data-driven. The lower bound is limited by data constraints from Evangeline Parish, LA which is missing data in UCR more than 27 months prior to their investigation. The upper bound is limited by Chicago – the last investigation – which was announced on December 7, 2015.

Moreover, the units of μ_i are a bit arbitrary. Mas (2006) assumed μ_i are months. Ayres and Levitt (1998) assumes μ_i are yearly. Jacobson, et al (1993) use quarterly data. Our tradeoff is between finer data units which has the benefit of being more narrowly tailored to the exact timing of the event and precision in the outcome variable of interest. For instance, although we know the exact day that each investigation was announced, outcomes such as homicides are too few (e.g. noisy) at the daily level to be informative. For the majority of our specifications, μ_i is monthly, though for high frequency outcomes such as total crime or police-civilian interactions we can conduct analysis at the daily level. We ensure a balanced event study design by restricting our sample such that each city has an event window of precisely the same length.

We begin by analyzing the unprocessed data – estimating whether there are any breaks in trend or levels of crime rates in the month the investigation is announced. This approach is quite simple and transparent, but potentially problematic as it does not account for potential changes in national crime trends.

To address this, we compare the average difference in felony crime rates in investigation cities with the average differences in felonious crime rates in a synthetic city constructed for each investigation city (Abadie, et al 2010). The synthetic city – a weighted combination of cities leaving out the one considered – is constructed to mirror values of predictors of the outcome variables.[17] As

---

[17]To construct a synthetic control city, we first take the sample of all law enforcement agencies that have jurisdiction over a principal city and have non-missing data available over a full event window to create a "donor pool" – a set of comparison cities that are eligible to receive positive weights in the synthetic control city. We restrict this donor pool further to look most like the city that is investigated. This is done by finding a list of 400 cities that are closest to

predictors, we use lagged outcome variables (felony crime three, six, and nine months before the event) as well as the poverty rate, median household income, unemployment rate, percent white, black, and hispanic, total population, and percent of the population between ages 15-24 – taken from the previous decennial census closest to the start of the investigation (in 1990 or 2000) or the 2011 ACS 5-year estimates.[18] For each outcome, homicide or total crime, we construct different synthetic controls. Note well: results using synthetic controls for homicide (resp. total crime) when total crime (resp. homicide) is the outcome yield identical results (see Appendix Table 6).

Our main empirical model is similar to that implemented in Jacobson, et al (1993):

$$crime_{tc} = \boldsymbol{X}_c\boldsymbol{\beta} + \sum_{\tau \geq \mu_1}^{\mu_2} D_{\tau tc}\delta_\tau^S + \sum_{\tau \geq \mu_1}^{\mu_2} D_{\tau tc}Investigated_c\delta_\tau^I + \epsilon_{tc} \tag{1}$$

where $\tau \in \{-27, .., 24\}$. $D_{\tau tc}$ are an exhaustive set of indicator variables denoting months away from the investigation, $Investigated_c$ is a dummy variable that denotes whether the city was investigated or not and $\boldsymbol{X}_c$ is a set of city specific control variables that include the education, economic and race indices described in section 3. Notice, δ_τ^S gives the average outcome in month τ away from the investigation in synthetic cities while $\delta_\tau^S + \delta_\tau^I$ gives the average outcome for investigated cities. Crime is demeaned by the average number of crimes in that calendar month of the city in the pre-event period.[19] The exogeneity assumption here is that the $Investigated$ dummy is uncorrelated with the error term.

In order to conduct inference, we estimate the cumulative effect of investigations on homicide and total crime, separately, over each of the post-investigation announcement months, relative to differences in the pre-investigation period. Using (-27,24) windows, we fit the following the model to the data:

$$crime_{tc} = \alpha + \alpha^I Investigated_c + \boldsymbol{X}_c\boldsymbol{\beta} + \sum_{\tau \geq 0}^{\mu_2} D_{\tau tc}\delta_\tau^S + \sum_{\tau \geq 0}^{\mu_2} D_{\tau tc}Investigated_c\delta_\tau^I + \epsilon_{tc}. \tag{2}$$

---

the investigated city in terms of (i) total population, (ii) percent black, (iii) poverty rate, and (iv) the mean value of the outcome variable in the 27 months before the event. We take the intersection of these four lists to form the donor pool for each investigation and crime type. As a final step, we remove any city that was simultaneously treated as the investigated city. This donor pool is then used for the creation of a synthetic control city. See the Data Appendix for details on restrictions of donor pools and Appendix Table 3 for a description of what cities make up each synthetic control city.

[18]We tested other combinations of predictors in this sample and there was not any significant improvement in the root mean squared prediction error.

[19]If we do not de-mean crime or include month-away-from-investigation dummies, but instead add a post-investigation dummy and month-year and city fixed effects the results are strikingly similar. We do not have enough investigated cities to replicate the exact specification in Mas (2006).

For each post investigation period, we cumulatively add the difference-in-difference estimates δ_τ^I to obtain the total *unexplained* gap in the number of crimes between investigation and synthetic cities τ months after the investigation: $\Gamma_\tau = \sum_0^\tau \delta_T^I$. Γ_τ is the cumulative difference-in-difference estimate of the effect of having an investigation on crime τ months after an investigation is announced. A positive value of Γ_τ implies that the gap in the crime rate between investigation and synthetic control cities in the τth month after an investigation is wider than the average gap in the crime rate between these two groups in the pre-investigation period, *ceteris paribus*. Note: cumulative estimates can be calculated for one treated unit or aggregated across many – an important advantage of the synthetic control approach.

Given the autocorrelation in monthly crime levels within cities, standard errors are clustered at the city x investigation level. Moreover, as the cumulative difference estimates are linear combinations of monthly estimates, we calculate the standard error of the cumulative difference estimates accounting for correlation between different estimates.

C. Descriptive Statistics

Appendix Table 4 displays descriptive statistics. Columns (1)-(4) provide the mean for each variable in investigated cities, synthetic cities chosen to match homicide pre-trends, synthetic cities chosen to match total crime pre-trends, and cities that were neither investigated nor used to construct synthetic controls, respectively. Column (5) reports the p-value for a test of equality of observables between investigated and synthetic cities and column (6) reports a similar p-value for investigated cities and non-investigated, non-utilized cities. P-values have been adjusted for small sample size as the standard normality assumption is not likely to hold. Following Morgan and Winship (2007), we divide the mean of the variable by the square root of the average of the within investigated city and comparison city variances. We then regress the adjusted variable on an indicator variable for an investigated city or synthetic city (for column (5)) or an indicator for investigated city or non-investigated, non-utilized city (for column (6)). The p-value is calculated using bootstrapped standard errors with 1000 replications (Efron, 1979).

Generally, investigated cities and synthetic cities are well balanced. The p-value on an F test of joint significance is 0.170. Investigated cities (and their synthetic controls) are quite different, however, from cities that were neither investigated nor utilized to construct synthetic controls. Investigated cities have a significantly larger percentage of blacks, less whites, lower fraction of population between the ages of 15 and 24, higher poverty, higher unemployment, higher fraction

14

of single parent homes, and higher fraction of individuals on public assistance. They also have substantially larger populations.

4 The Effect of Investigations on Crime Rates

4.1 Raw Data

Figure 1 displays monthly homicide (panel A) and total crime (panel B), demeaned, for our $(-27, 24)$ event window, which has the advantage of allowing us to examine both the persistence of potential effects and pre-investigation trends over a relatively long time span. We consider the effects of investigations for the 27 investigations held in 25 principal cities only, to maintain consistency over different empirical specifications.[20] For homicides, there is neither a level nor a trend break in the month of the event. Total crime demonstrates a small change that is statistically significant but substantively small.

Digging deeper, Figure 2 plots the size of the level change for each investigation city by increasing size of the level break. Of the twenty seven investigations, one has a statistically significant decrease in homicide rates, five have a statistically signficant increase in homicide rates, and twenty-one have insignificant impacts. Four out of the five cities which have statistically positive breaks in homicide rates also have a remarkable level break in total crime rates in the month the investigation is announced. The only exception to this pattern – although displaying an insignificant level break in total crime – displays a striking trend break with total crimes increasing for the twenty four months post investigation.[21]

Figures 3A and 3B explore the heterogeneity of the impact of investigations to understand what could be driving five investigations to have starkly different impacts on crime compared to the rest. Figure 3A (resp. Figure 3B) plots level breaks in homicide rates (resp. total crime) against city demographics. The various city-level demographic outcomes considered are total population, percent poplation between 15 and 24 years of age, racial composition, median income, poverty rate, percent on public assistance, percent high school graduates, percent with a four year college

---

[20]We display results for the 15 investigations from non-principal cities in Appendix Figure 2.

[21]To ensure that the positive effects on crime from the five cities is not a random occurence, we permute and check if the results still hold against all possible combinations of five cities from the total sample of 27 investigations. To conduct the permutation test, we estimate the cumulative impact of investigations 24 months post investigation for all possible subsets of five investigated cities from the sample of all investigations. We, then, calculate the p-value of the cumulative impact of investigations for our original sample of five cities in the distribution of all possible cumulative impacts. The results remain statistically significant and robust to accounting for finite sample inference using a Fisher exact p-value (Appendix Figure 6).

degree, isolation index, percent households with single parents and unemployment rate. The only outcomes that are significantly correlated with level breaks in homicide rates are total population, percent on public assistance, and percent of the population with a four year college degree. After adjusting the p-values for multiple hypothesis testing, however, none of them are significant.

Next, we probe into the incident that sparked the investigation in various police departments. Five investigations were initiated due to "viral" incidents of deadly use of force. These incidents were namely the deaths of Freddie Gray in Baltimore, MD, Laquan McDonald in Chicago, IL, Timothy Thomas in Cincinnati, OH, Tyisha Miller in Riverside, CA, and Michael Brown in Ferguson, MO. Each of these incidents were followed by national media attention, protests and riots over several days. The other twenty-two investigations were mostly conducted after complaints from civilians or lawsuits from organizations like the American Civil Liberties Union. The five investigations that were preceded by viral incidents of deadly use of force perfectly correlate with the five that have statistically significant positive level breaks in homicide rates in the month of the investigation.

Figure 4 displays the raw data for homicide rates and total crime for our event window for all investigations that were not preceded by viral incidents of deadly force (Panel A) and for the five investigations that were (Panel B). For investigations without viral incidents, there is neither a level nor a trend break in either homicide or total crime rates. Crime rates are steadily decreasing before and after the month the investigation was announced. This data suggests that pattern-or-practice investigations are benign.

Investigations that were sparked by viral incidents of deadly force have a stunning pattern. There are large level breaks in both homicides and total crime, in each principal city. The change in homicide rates in the month an investigation is announced is roughly one homicide per 100,000 or 14.44, total, per month, for the five cities analyzed. For the twenty-seven months before the investigation, homicides seem to be slightly increasing. There is a large level break (p-value = .000) in homicide rates in the month the investigation is announced, and surprisingly, no apparent decline in homicide rates for the twenty-four months following the announcement of the investigation.

Total crime follows a similar, though distinct pattern. The change in the total crime rate in the month of the announcement of the investigation is 51.23 more felony crimes per 100,000 people – or 925.32 total – an 11.60 percent increase. For the months preceding the investigation, total crime is decreasing and, again, demonstrates a large level break (p-value = .000) in the month the investigation is announced. Twenty-four months after the investigation, total crime continues to be markedly higher.

The validity of any event study depends on an important assumption that no other events occurred at the time of the event that could lead to false appropriation of the causal effect. With only five observations it is plausible that, by random chance, something else that influenced crime occurred in those five cities in the months that coincide with the investigation. The more narrow the unit of analysis – from year to month to day – the more plausible our identification assumption becomes. Conversely, the more narrow the unit of analysis the more noise exists for low frequency events such as homicides. Throughout we have analyzed crime rates at the monthly level.

As a robustness check, we alter the units of μ_i and provide results for weekly and biweekly crime rates in Appendix Figure 1. We only have bi-weekly and weekly crime rates from incident reports for Batimore, Chicago and St. Louis. The other two investigations contain only monthly crime rates from FBI UCR. We check whether the level breaks that exist at the monthly level for these cities continue to exist when we disaggregate crime rates.

Using bi-weekly crime rates, the p-value on a positive level break at the bi-week of the investigation is 0.000 for both homicide and total crime. Using weekly crime rates, p-value on a positive level break at the week of the investigation is 0.020 for homicide and 0.000 total crime. And, given the high frequency of the total crime variable, the results are also robust to estimating the impact of investigations using daily crime rates (p-value of 0.027).

Taken together, the raw data are highly suggestive that investigations without viral incidents of deadly force had little effect on crime rates, whereas every investigation that was preceded by a viral incident of deadly use of force has had dramatic impacts on both homicide and total crime rates. In what follows, we explore the robustness of these findings across myriad dimensions. As we demonstrate below, the conclusions gleaned from the simple analysis of the raw data are remarkably robust.

4.2 Synthetic Control Estimates

An obvious concern with estimating level and trend breaks to analyze the impact of investigations on crime, is that other non-investigated cities may follow similar patterns. Indeed, many have commented on the so-called "Ferguson Effect" – a term coined by Doyle Sam Dotson III, the chief of the St. Louis police, to account for an increased murder rate in some U.S. cities as a result of the social unrest that followed the killing of Michael Brown in Ferguson, MO and the subsequent acceleration of the Black Lives Matter movement. In this scenario, analyzing the raw data as we

describe above would lead one to (falsely) attribute the change in crime that occurs in cities such as Baltimore, Chicago and St. Louis, to investigations.

We now compare the average number of crimes in the months prior to an investigation to the average number of crimes in the months after an investigation for both investigation cities and their synthetic controls. The resulting estimator measures the impact of investigations on crime rates in investigated cities relative to synthetic controls. An advantage of this approach over more standard propensity score matching is that we can provide the impact of investigations on crime for each treated unit.

Consistent with our analysis of the raw data, investigations that were not preceded by viral incidents of deadly force have a negative and statistically significant impact on homicide and total crime rates. This is demonstrated in Figure 5 Panel A by pooling the results across twenty-two investigations mainly investigated after civilian complaints or ACLU lawsuits, for which we can construct synthetic control cities.[22] Appendix Figures 3A (resp. Appendix Figures 3B) provide the impact of investigations on homicide (resp. total crime) for each of the twenty-two investigations.

The cumulative amount of crime that we estimate due to investigations in the twenty-four months after the announcement of an investigation, averaged across the twenty-two cities, is -6.39 (2.43) per 100,000 for homicides and -750.38 (431.05) per 100,000 for all felony crimes. That is, if anything, investigations that were not foreshadowed by viral incidents of deadly use of force marginally reduced homicides and total crime. Appendix Figures 3A and 3B provide city specific cumulative estimates over time.

In stark contrast, all five investigations sparked by viral incidents of deadly use of force exhibit a remarkable increase in both homicide and total crime. Figures 5A and 5B plot each city, separately, along with their synthetic control for homicide and total crime, respectively. The results for each city for both homicide and total crime are strikingly similar – a level break in the month of the event while their synthetic controls remain roughly constant throughout the event window. Riverside, CA is the only city which witnesses a reversion of the homicide rates to pre-investigation averages after eighteen months. For all other cities, homicide rates are consistently higher for the twenty four months after the investigation is announced. The cumulative amount of homicides that we

---

[22]Of the 37 investigations that were not sparked by viral incidents of deadly use of force, 15 are sheriffs' departments and cities that are not principal cities (i.e. small cities). This makes construction of their synthetic control cities difficult. We analyze level and trend breaks in these 15 sheriffs departments and non-principal cities in Appendix Figure 2. None of the 15 investigations have a statistically signficant impact on homicide rates. On average, these investigations have a statistically significant, albeit small, positive impact on total crime rates. However, total crime trends downwards in the post investigation months.

estimate due to investigations in the twenty four months after the announcement, per 100,000, is 34.38 (4.96) for Baltimore, 14.38 (1.60) for Chicago, 29.53 (2.78) for Cincinnati, 4.67 (3.75) for Riverside and 53.76 (5.60) for St. Louis, or 179 more homicides, in total, *per investigation.* Total crime follows similar patterns.[23]

In other words, the causal effect of the investigations in these five cities – sparked mainly by the deaths of five African-American civilians at the hand of police – has resulted in 893 more homicides than would have been expected with no investigation and more than 33,472 additional felony crimes.[24] This result is not obtained by chance.

Since asymptotic inference cannot be conducted on synthetic control estimates, Abadie et al. (2010) propose placebo tests to demonstrate the significance of their treatment impacts. In order to do so, they permute treatment assignment randomly across each unit in the donor pool and apply the synthetic control method to estimate placebo treatment effects. This provides a distribution of effects which is then compared to the real treatment effect. We conduct the same exercise for all cities investigated after a viral incident of deadly use of force to understand whether the large cumulative effects on homicide rates and total crime rates were obtained by chance. The results are provided in Appendix Figures 5A and 5B.

The first graph compares the cumulative amount of homicides estimated due to the investigation in Baltimore to placebo cumulative amounts if the investigation had been randomly assigned to any other city in Baltimore's donor pool. The distribution of t-statistic for cumulative effects are plotted. Other graphs show the same distribution for Chicago, Cincinnati, Riverside and St. Louis respectively. For all investigations excluding Riverside, the probability of obtaining results of the magnitude of those obtained in any city investigated after a viral incident are exceedingly small (exact p-values between 0.014 and 0.038). For Riverside, the exact p-value is 0.207.[25]

---

[23]To show that the increase in homicides and total crime were not completely driven by St. Louis in the St. Louis/Ferguson combination city, we plot crime trends for homicides per 100,000, person crime per 100,000 and total crime per 100,000 for the city of Ferguson in Appendix Figure 4. The change in homicide rate on the month of the annoucement of the investigation is 1.25 (p-value = 0.185) while the change in person crime is 20.94 (p-value = 0.063). Interestingly, homicides have a positive and significant trend in the 39 months after the announcement of the investigation.

[24]Recall, Mas (2006) estimates that arbitration rulings in favor of the employer cause 600 excess crimes per 100,000 people after 23 months. We estimate roughly 1157 excess crimes per 100,000 after 23 months – 92.83% larger than the estimates in Mas (2006).

[25]As noted before, Riverside is the only city that witnesses a decline in homicide rates 18 months post investigation. In fact, the exact p-value for the cumulative amount of homicides after 18 months in Riverside compared to its placebo treatment effects is 0.092.

4.3 Regression Estimates

A. ORDINARY LEAST SQUARES WITH SYNTHETIC CONTROLS

Table 1 reports a series of parametric regression estimates corresponding to a (-27,24) window, by estimating the following equation:

$$crime_{tc} = \alpha^S + \alpha^I Investigated_c + \boldsymbol{X}_c\boldsymbol{\beta} + Post_{tc}\gamma^S + Post_{tc}Investigated_c\gamma^I + \epsilon_{tc} \qquad (3)$$

where α^S denotes the average felony crimes in the pre-investigation months for synthetic cities. Conversely, $\alpha^S + \alpha^I$ denotes the average felony crimes in the pre-investigation months for investigated cities. The impact of an investigation is captured by γ^I.

Panel A estimates equation (3) using all investigations in principal cities for which we have non-missing data. Panel B restricts the sample to the investigations that were not preceded by viral incidents of deadly use of force and panel C restricts the sample to the five investigations that were preceded by viral incidents of deadly use of force. Column (1) reports the change in crime rates from the pre- to the post-investigation period for investigated and synthetic control cities. The estimates – given there are no controls – can be interpreted as simple differences in means. The main coefficient of interest is Post-Event x Investigated (γ^I from (3) above). Column (2) weights by population in the city using decennial census data closest to the investigation year. Column (3) includes population weights as well as controls aggregated into three indices: an economic index, a race index, and an education index. As with the construction of synthetic controls, the data is taken from the previous decennial census (in 1990 or 2000) or the 2011 ACS 5-year estimates closest to each investigation date. Adding these variables leaves the magnitude of the coefficients virtually unchanged.

Consistent with the raw data, when the sample is all investigations, estimated effects are small and statistically zero (-0.08 (0.19) homicides per 100,000 and -15.76 (18.89) per 100,000 for total crime). When we partition the data to separate investigations by the incident that sparked the investigation, interesting results emerge. If the investigation was preceded mainly by civilian complaints, allegations or lawsuits, pattern-or-practice investigations *reduce* homicides by 0.26 (0.10) per 100,000 a month. The impact on total crime is statistically insignificant at the 5% level.

For investigations that were sparked by a viral incident of deadly use of force, the estimate of γ^I reverses sign and substantially increases in overall magnitude. The coefficient on γ^I is close to

1 homicide per 100,000 per month. This is a 48.28 percent increase. For total crime, the increase is 47.65 (16.36) crimes per month, per 100,000 – an 10.79 percent increase.

B. Propensity Score Matching

Equation (3) is a simple way of obtaining estimates of the effect of investigations on homicide and total crime, but it relies on the construction of synthetic controls for a comparison group. This may be unappealing to some given the lack of transparency.

As a potential solution to this concern, we now match investigated cities and other cities with similar predicted probabilities or propensity scores (p-scores) of being investigated. The estimated p-scores compress a multi-dimensional vector of covariates – in our case city-specific demographics – into an index (Rosenbaum and Rubin, 1983). An advantage of the propensity score approach is that it provides a more transparent way to focus our comparisons on investigated and non-investigated cities among those with similar distributions of city-specific demographics. And, there is an active debate on the inference procedure for synthetic control methods (Firpo and Possebom, 2017); errors using propensity score matching are consistently estimated.

We implement p-score matching in three steps.[26] First, the estimated propensity scores are obtained by estimated probit regressions for whether or not a city is investigated, using percent population between the ages of 15 and 24, median income, percent hispanic, percent white, and unemployment rate.[27] The pool of cities used as potential comparison cities is the intersection of four samples. Each sample comprises 400 cities that are closest to the investigated city in terms of (i) total population, (ii) poverty rate, (iii) fraction black, and (iv) the mean of the outcome variable in the 27 months before the investigation. Second, the "treatment effect" for a given outcome is calculated by comparing the difference in outcomes between investigated cities and cities with "matched" values of the p-score. We do this in two ways. The first calculates a treatment effect for each city relative to its "nearest neighbor." In the case that a city is matched to multiple neighbors with the same propensity score, we take the average of all cities to construct a comparison city. Further, this matching is done with replacement so that individual cities can be used as controls for multiple investigation cities.

Third, a single treatment effect is estimated by averaging the treatment effect across all investigated cities for which there was at least one suitable match.

---

[26]See Dehejia and Wahba (2002) and Heckman, Ichimura, and Todd (1998) on propensity score algorithms.

[27]For a few cities, matching using these set of variables does not produce a match as the treated city gets identified perfectly. For these cities, probit regressions are estimated for whether or not a city is investigated, using percent population between the ages of 15 and 24, median income, and percent white.

The results of implementing the propensity score approach – displayed in Table 2 – are similar to those described throughout. For investigations not preceded by viral incidents of deadly use of force, the impact of investigations was relatively benign. If anything, they lowered crime. For investigations that were preceded by viral incidents of deadly use of force, however, the impact of investigations on both homicide and total crime are large and statistically significant – despite the limited number of observations. Taking the coefficients at face value, our p-score estimates imply 1099 excess homicides over twenty-four months and 31,293 excess felony crimes.[28]

5 Alternative Explanations

In this section we partially test the validity of our identification assumption by exploring two potential violations of equation (1). Perhaps the most obvious is that it is a controversial use of force that results in a fatality that causes crime to increase and this is being confounded with investigations preceded by viral incidents of deadly force. A second potential violation is that there is something unique about the five cities and any controversial use of deadly force in these cities in a sensitive time period leads to an increase in crime. We discuss each in turn.

A. 'Viral' Police Shootings and Crime

It is plausible that it is controversial uses of excessive force in a sensitive period that led to an investigation – not the investigation itself – that caused an increase in homicide and total crime. That is, shootings such as Tamir Rice, a 12-year old boy shot twice by police officers who mistook his toy gun as a real firearm, Christian Taylor in Arlington, TX, who was unarmed and shot four times after he broke into a car-dealership, or Walter Scott in North Charleston, who was shot eight times from behind while he was fleeing, will lead to increases in crime.

A partial test of this hypothesis is to investigate the impact of viral shootings that resulted in a fatality but for which the federal or state government did not launch an investigation. Finding an exhaustive list of all fatal shootings that caused national outcry and riots is difficult. The New York Times assembled eleven videos that were viewed at least two million times (New York Times, 2017). We adopt this as our definition of "viral" video, realizing that it is ad hoc.

Of the eleven videos documented by New York Times (2017), three of them are dropped from our analysis due to missing data. For instance, New York Police department provided tri-monthly

---

[28]We also estimated the impact of investigations using the same matching process but with kernel weighted averages of all cities in a donor pool instead of nearest neighbor matching. We used a Gaussian kernel with the default bandwidth of 0.06. Results were qualitatively unchanged.

– rather than monthly – data to the FBI until 2013.[29] Our final sample consists of viral shootings in eight police departments. To estimate the impact of these shootings on homicide and total crime, we estimate the model described in equation (1) with a (-30,12) month window, replacing the "Investigated" variable with a "Viral Shooting" variable.

Figure 7 display the results comparing crime rates over a (-30,12) month event window for viral shooting cities and their synthetic controls. The trends between viral shooting cities and their synthetic controls before and after the viral shooting look similar, suggesting no impact of viral shootings on crime. Appendix Figures 7A and 7B provide the impact of viral shootings on homicide and total crime, respectively, for each of the eight viral shootings considered.

The cumulative amount of crime that we estimate due to viral shootings in the twelve months after the shooting averaged across the eight cities, is 2.94 (1.09) for homicide and 136.32 (218.91) for total crime. The equivalent number for investigated cities that were preceded by viral incidents of deadly use of force is 12.88 (2.99) for homicide and 639.85 (207.68) for total crime.

B. The Effect of Police Shootings of Unarmed Black Men in Baltimore, Chicago and Cincinnati on Crime Rates

Investigations are not randomly assigned. Thus, one may suspect that there is something special about cities that the Department of Justice decides to investigate. This, coupled with the increased scrutiny on policing after a controversial incident of deadly use of force, may violate our exogeneity condition. If there is something special about Baltimore, Chicago, Cincinnati, Riverside and Ferguson, then comparing them to cities that have had viral shootings but no investigations is inadequate because there is a reason that those cities were not investigated.

To make this more concrete, consider the following thought experiment. Imagine that what leads to an investigation is that the Civil Rights Division is sufficiently convinced that a police department is racist. In the case of Baltimore, Chicago, Cincinnati, Riverside and Ferguson there were enough signals of racism that the death of Freddie Gray and the shootings of Laquan McDonald, Timothy Thomas, Tyisha Miller and Michael Brown were enough to tip the scales in favor of an investigation. Yet, for Baton Rouge, St. Paul, or North Charleston, there were *not* enough signals of racism that the shootings of Alton Sterling, Philando Castille, or Walter Scott were enough to lead to an

---

[29]We also submitted a FOIA request for monthly data, but they communicated that they only store tri-monthly data.

investigation.[30] In this case, our estimates are not simply the causal effect of investigations but are about a controversial shooting in a city with a racist police department in a sensitive time period.

A partial test of this alternative explanation is to examine police shootings of unarmed blacks in Baltimore, Chicago and Cincinnati in the pre-investigation months but before the shootings that led to their investigation.[31] There are two police shootings in Baltimore, five in Chicago and two in Cincinnati that satisfy this criterion. We cannot implement the synthetic control method for this set of shootings, as the event windows, especially for Baltimore and Chicago, are too small. In Baltimore, the length between the two shootings is 47 weeks. In Chicago, the average length between shootings in the sample period is 14.5 weeks.

Thus, we analyze first differences and, again, check for level or trend breaks in the weeks surrounding the event. Figure 8 displays results. For both cities, we test whether jointly there is a break at the week (or, month for Cincinnati) of any shooting. In Baltimore, the p-values on a joint test of trend and level break are 0.005 and 0.130 respectively. In Chicago, the p-values are 0.854 and 0.001 respectively. For Cincinnati, the corresponding p-values are 0.535 and 0.310 for homicides and 0.059 and 0.191 for total crime. If anything, it seems that crime decreases in the weeks after such shootings take place. Taken together, this evidence suggest that it is the investigation, not an unarmed shooting in a sensitive period and in a city that will eventually be investigated, that caused crime to increase in these five cities.

It is important to remember, however, these shootings were not "viral", so it leaves open the possibility that our investigation treatment effect is confused with a viral shooting in a city with significant racism. We cannot directly test this potential violation. Casting doubt on its relevance, however, is the fact that four of the five deaths in the five cities were not viral videos.

6   Potential Mechanisms

Our analysis has generated a rich set of new facts. In general, the impact of pattern-or-practice investigations is benign. But this masks significant and potentially important heterogeneity. For investigations that were not sparked by viral events of deadly use of force against black civilians,

---

[30]There are media reports that the Trump Administration refused to start pattern-or-practice investigations that were imminent in the final days of the Obama administration. See "Sweeping Federal Review Could Affect Consent Decrees Nationwide" published by the New York Times on April 3, 2017. Despite repeated attempts to obtain information from the Justice department on the set of investigations that were imminent but cancelled, we were unsuccessful.

[31]There are no such shootings in Ferguson or St. Louis and we, currently, do not have data on officer involved shootings from the investigation year for Riverside.

investigations lowered homicide rates and total crime. For investigations that were sparked by a controversial incident of deadly use of force against a black civilian, investigations caused a large increase in both homicide and total crime rates. These impacts do not occur in cities that had viral shootings but no investigation. And, shootings of unarmed black men in the cities that were investigated – in months close to but before the investigation – had no impact on crime.

In this section, we take the point estimates at face value and provide a (necessarily) more speculative discussion of potential mechanisms that may partially explain our set of facts – concentrating on the five cities that have demonstrated marked increases in crime. Unfortunately, our ability to test mechanisms is severely limited by data, as the types of data collected by individual police department varies greatly and so too does their willingness to release the data to researchers even after multiple FOIA requests. For instance, despite the fact that we used exactly the same process to obtain data from each city, very few fulfilled requests for sensitive data such as 911 response times or police activity. We view the forthcoming exercise as more of a rigorous case-study than definitive proof.

With these important caveats in mind, we consider three potential mechanisms: (1) changes in policing activity, (2) decreased community trust or engagement, and (3) changes in the nature or content of consent decrees.

6.1 Police Activity

"As she was preparing to leave, I asked this officer if she was willing to admit that the conflict with [State Attorney] Mosby, and the mutual mistrust, has led police officers to pull back from their essential duties. One of the most discouraging statistics in Baltimore has been a 63 percent increase in homicides last year…"Absolutely," she said. "You should get used to 300 murders a year." Baltimore vs. Marilyn Mosby, New York Times, September 28, 2016.

Systematic data on police activity is perhaps the most difficult to obtain from local police departments – in part, because it relies on police departments to release highly sensitive data about the whereabouts of officers at given points in time or manually input significant numbers of hand written notes (more than 1,000 police-civilian interactions per day in Chicago), and in part, because the types of interactions with civilians that are recorded varies significantly by police department.[32]

---

[32]Response time involving calls for service is another margin that police can potentially alter their effort. We attempted, and were largely unsuccessful, to obtain data on 911 response times. The only city that responded positively was St. Louis. Appendix Figure 8 plots average response time for a (-50,50) day window around the investigation. If anything, 911 response times gradually increase after the day of announcement of the investigation.

We were able to assemble data from ten police departments – though it was requested from twenty five – on police activity as measured by interactions with civilians.

No Viral Shooting and Investigation

Albuquerque, NM

We have data on traffic stops for one city that was investigated without a viral incident of deadly force – Albuquerque, NM.[33] Panel A of Figure 9 plots demeaned monthly traffic stops per 100,000 for a (-30,24) window around the investigation into Albequerque, NM – which was investigated for use of excessive force. In general, stops are trending downward but there is neither a level break in the month of the investigation (p-value is 0.607) nor a break in trend (p-value = 0.674).[34]

Viral Shooting and No Investigation

We have data on traffic stops for five cities that experienced a viral shooting but no investigation – Arlington, TX, Cincinnati, OH, St. Paul, MN, Charlotte, NC and Tulsa, OK. Panel B of Figure 9 plots monthly (and daily where possible) traffic stops per 100,000 for the available data window around viral shootings. In general, stops might decrease after the viral shooting but they trend back up in the following months. For instance, stops seem to decrease in the first two months in Arlington, TX but increase thereafter. Similarly in Cincinnati, OH, St. Paul, MN and Charlotte, NC, stops decrease in the month of the viral shooting but revert back to pre-shooting means in the later months.

Viral Shooting and Investigation

Chicago, IL

The Chicago Police Department has one of the most comprehensive datasets on contacts with civilians. Data on all police-civilian interactions were collected on a three by five inch index card.

---

[33]Repeated attempts to obtain more data from the other twenty-one cities were unsuccessful (see Appendix Table 2 for a complete description of our attempts).

[34]Davis et al. (2005) provide an interesting potential counter example to the results found for Albuquerque, NM. They show that 61% of surveyed Pittsburgh police officers reported major changes in job performance as a result of their investigation. And, of those, 56% indicated that officers became less active while 41 % reported officers became less efficient. The authors also report that 79% of surveyed police officers mention officers taking a less proactive approach due to the decree. Importantly, this reduction in police activity corresponds to an increase in crime rates in Pittsburgh.

The data include information on why the stop was initiated – mostly traffic stops, suspicious behavior, gang/narcotics related, crime victim, and dispersals. Other information collected consists of location and time of contact, civilian demographics and certain physical attributes of the civilian contacted. Prior to the investigation, there were approximately 50,000 police-civilian interactions recorded each month.

Figure 10A displays the number of police-civilian contacts in Chicago by month. The line at 0 is the month of the investigation. The month before the investigation there were 39,071 police-civilian contacts recorded. In the month after the investigation there were 4,371 police-civilian interactions recorded – an 89% decrease.

This decrease could be due to three factors. The first is the investigation. Second, as part of an ACLU lawsuit against Chicago Police Department's practice of unconstitutional stops, the Chicago Police Department changed its three by five index card to a two-page form on January 1, 2016. The additional information collected includes detailed information on the location of contact, reason behind contact, whether civilian was searched, and if searched, whether any contraband was found. There are three weeks between the announcement of the investigation and the implementation of the new contact cards which allow us to disentangle the impact of the investigation on police-civilian contacts.

Figure 10B displays the number of police-civilian contacts in Chicago by day. The first line – November 24, 2015 – is the day the Laquan McDonald video was released to the public revealing that he was shot sixteen times while walking away. The second line – December 7, 2015 – is the day the investigation was started. The third line is the day the new two-page contact cards were implemented. There is a discrete decrease in stops the day after the Laquan McDonald video was released, but contacts with civilians gradually increase returning to pre-release levels in a week. Once the investigation is announced, stops again decrease dramatically – but do not return to pre-investigation levels. One day before the investigation was announced there were 1,132 police-civilian contacts; 1,079 a day after; 838 a week after; and 359 police-civilian contacts three weeks after. We estimate that 80% of the decrease in monthly stops above is due to the investigation and the remaining 20% due to different contact cards.

A third potential explanation is that police stopped recording their interactions with civilians, but actual policing remained the same. This is difficult to test. But, in this scenario, one would expect there to be no correlation between police districts that decrease stops the most and reductions in crime. Panel A of Figure 11 correlates the average change in monthly de-meaned police activity

27

with the average change in de-meaned homicides and total person crime – omitting property crime – over each of the 22 districts in Chicago, weighting each geographic division by its average monthly pre-investigation crime rate.

The relationship between police activity and homicide is striking. Taking the slope at face value, a decline of 1000 police-civilian contacts increases homicides by 0.6. For person crimes i.e. homicides, rapes and aggravated assaults, a decline of 1000 contacts increases person crimes by 4.65.[35] A final piece of evidence that argues against the hypothesis that police simply stopped filling out contact cards but otherwise kept policing is that the correlation between stops and crime is statistically identical before and after the investigation. In fact, using the correlation between police contacts and crime, we estimate that 98% of the increase in homicides can be accounted for by reductions in police-civilian interactions in Chicago.

Riverside, CA

One of our most important pieces of data was obtained from Riverside, CA, who experienced a viral shooting and investigation in 1999 – before celluar phones had camera or video capabilities or social media existed, 9 years before Obama took office, and almost 15 years before the rise of the Black Lives Matter movement. We have data on police-civilian interactions for both pedestrian and traffic stops from 1997-2001. Figure 10C plot monthly police-civilian interactions from these data.

There were approximately 2,572 monthly traffic and pedestrian police stops in Riverside between January 1997 and February 1999. The month before the investigation was announced there were 2,996 stops. The month after there were 2,331 – a 22.2% decrease, followed by 1,181 the month after that. After 16 months, police-civilian interactions were, again, at pre-investigation levels.

St. Louis, MO

For St. Louis, we conduct a parallel analysis using Self Initiated Activity or SIA. The St. Louis Metropolitan Police Department includes SIA as a part of its Calls for Service dataset. It includes all activities that were initiated by police officers and primarily consist of traffic violations, occupied and unoccupied car checks, pedestrian checks, building checks, directed and foot patrol, investigations, problem solving, truck inspections and business interviews. The data includes information

---

[35]We also present the same graphs for total crimes in Appendix Figure 9. The correlation between police activity and total crime seems to be positive and statistically insignificant. This is completely driven by property crimes which seem to have increased unambiguously over all geographic sub-divisions in a city regardless of the decline in police activity, during the investigation.

on the time at which the call was placed, the dispatch, arrival and closing times of the call. It also includes the radio code assigned to the call which helps us identify self-initiated activity from other 911 calls placed. Finally, it includes detailed information on the location of the call. The total number of recorded SIA per month averaged at 19,936 prior to the investigation.

Figure 10D displays monthly de-meaned SIA trends. The number of recorded SIA a month before the shooting of Michael Brown and unrest was 17,153. This declined to 11,532 in the month of the shooting and large-scale unrest. A month after the investigation, or two months after the unrest, recorded SIA was 10,449. Once we de-mean the data, this decline translates into 1,334 less SIA per 100,000 civilians in the month following the investigation. Similar to Chicago, this decrease can be explained by three factors. The first is the investigation. The second is that police officers stopped recording SIA to prevent these activities from being scrutinized. The third is that police officers reduced SIA due to the large scale violent protests that occurred after Michael Brown's shooting. One may argue that the decline in SIA we observe was entirely due to the protests and policing activity never improved after that event.

Figure 10E displays self-initiated activity in St. Louis by day. The first vertical black line – August 9, 2014 – denotes the day on which Michael Brown was shot. Violent protests erupted that night and continued for 11 days after. Missouri Governor Jay Nixon declared a state of emergency on August 16, 2014 and issued an executive order calling in the National Guard on August 18, 2014. He withdrew the National Guard from Ferguson on August 21, 2014 after witnessing improvements in the city.

There is a stark decline in SIA on the day of Michael Brown's shooting. A day before the shooting, the total number of recorded SIA was 674. This declined to 274 a day after the shooting and declined further to only 64 on August 20, 2014 while the protests were happening in full force. However, once the situation improved and the National Guard was withdrawn, SIA increased rapidly. In fact by September 3, 2014, a day before the investigation was announced, SIA was recorded at 658. Since the day of investigation, SIA steadily declined to 296 a month after and 183 two months after the date of announcement.[36]

Panel B of Figure 11 correlates the average change in monthly de-meaned SIA with the average change in de-meaned homicides and total person crime over each of the 89 neighborhoods in St.

---

[36]Although we could not obtain daily or monthly counts of traffic stops for Ferguson, the annual count of traffic stops in Ferguson as reported by the State Attorney General's Office was 5384 in 2013 (pre-investigation year), 4588 in 2014 (investigation year where the investigation happened in the last quarter), 1204 in 2015 and 1243 in 2016 (post-investigation years) respectively. This implies a decline of 75% in traffic stops between the year of the investigation and the year after.

Louis, weighting each geographic division by its average monthly pre-investigation crime rate. As before, the relationship between police activity and homicide is strong. Taking the slope at face value, a decline of 1000 self-intiated activity increases homicides by 1.55. For person crimes, a decline of 1000 contacts increases person crimes by 3.80.

BALTIMORE, MD

Data from Baltimore is more difficult to interpret. Police-civilian interactions in Baltimore are partitioned into traffic stops and pedestrian stops. Unfortunately, pedestrian stops, to date, are unavailable precisely during the critical period of the investigation window.[37] Importantly, the findings report issued by the Justice Department at the conclusion of the investigation describes systematic inconsistencies in their stops reporting. The Civil Rights Division audited records of pedestrian stops – comparing officers' completed reports to records on the departments computer aided dispatch system – and reported that Baltimore Police Department reported roughly 30% of stops audited. In other words, when officers were dispatched to a stop, they only recorded data on less than one-third of those stops. Using a similar strategy, comparing data from gun arrests in court records and stops data reporting finding a gun, the investigation found that Baltimore Police Department did not report a single instance out of 335 stops audited where a handgun was found. All results gleaned from Baltimore data should be interpreted with these caveats in mind.[38]

We obtained data on Baltimore traffic stops from 2012 to 2016. Figures 10F and 10G plot monthly and daily data from these data. There were approximately 364 daily traffic stops in Baltimore between March 28, 2015 and April 11, 2015. The day before the arrest of Freddie Gray, April 11, 2015, there were 335 recorded traffic stops. The death of Freddie Gray is associated with a phenomenal decrease in the number of traffic stops – reducing to only 8 stops on April 28, 2015. This time period in Baltimore had riots, looting, and general social unrest.[39]

The announcement of the federal investigation could not have decreased stops any more; traffic stops are bounded below by zero! After the investigation, there is a gradual increase in traffic stops. Despite the increase however, stops never return to pre-event levels during our study window.

<p style="text-align:center">***</p>

---

[37]Data from 2014, the first quarter of 2015, and 2016 are available. The investigation began in April 2015.

[38]A high-ranking Baltimore police officer familiar with the data also highlighted that they began keeping better records immediately following the announcement of the investigation, further complicating any event-study design.

[39]Some have posited that the theft of 175,000 doses of opioids may have driven the rise in crime in Baltimore. It is impossible to disentangle that from all other events that happen in a relatively short time frame, but if we calculate the average monthly homicides caused by drug overdoses as reported by the Centre for Disease Control and Prevention, in Baltimore city, we see no discernible differences after the investigation.

Taken together, there is significant evidence that police activity changed considerably after pattern-or-practice investigations were announced post viral incidents of deadly use of force. This pattern is not visible for any of the cities where the investigation was not preceded by viral incidents of deadly force or in cities that experienced viral shootings that resulted in fatalities, though we only have sparse data for earlier investigations so one cannot make definitive conclusions. [40]

6.2 Community Trust

Another potential mechanism we consider to explain the results is that community-police relations may change as a function of pattern-or-practice investigations and this leads community members to be less involved, to trust the police less, or to feel less obliged to cooperate when asked. Recall, at the beginning of an investigation, staff from the Civil Rights Division discuss the nature of the investigation with members of the community. If potential criminals believe that community engagement with police is a vital input to solving crimes and that engagement has decreased then the expected price for committing a crime will decrease and crime will increase (Becker 1968).

Systematic data on attitudes towards police is not available for our particular cities during the relevant time windows.[41] In lieu of this, we obtained daily data on 911 calls for service for a (-50,50) day window for three cities investigated after viral incidents of deadly force (Baltimore, Chicago

---

[40]An often-suggested reason for police reallocating or reducing their effort is fear of national media (Nix and Pickett, 2017), though many other explanations are possible. In almost every focus group we conducted with current police officers across a variety of departments we heard a familiar phrase: "I don't want to be in the next viral video." The potential of this hypothesis to explain the decline in police-civilian contacts, is perhaps best illustrated by the tragedy of the Chicago police officer who was beaten at a car accident scene and who did not draw her weapon – even though she feared for her life – because she was afraid of the media attention that would come if she shot him (Post 2016). Chief of Chicago Police, Eddie Johnson, remarked "She looked at me and said she thought she was going to die, and she knew that she should shoot this guy. But she chose not to because she didn't want her family or the department to have to go through the scrutiny the next day on national news." Appendix Figure 10 provides a time-series graph corresponding to a (-27,24) window for the quantity of articles published on police misconduct (See Data Appendix for details). In the 27 months before the event, cities that were investigated without viral incidents, those that were investigated after viral incidents, and those that experienced viral shootings (but were not investigated) have remarkably similar trends in media attention. Yet, the amount of attention that occurs for these different samples of cities differs in the month of the event. Cities investigated without viral incidents experience 7.47 articles in the month of the investigation and this returns to pre-investigation levels quickly thereafter. Similarly, cities that experience a viral shooting have media attention increase from 0.03 in the months before the event to 11.26 on the month of the shooting. The media attention quickly dissipates and resurrects roughly eight months after the event – typically to announce that the officer in the shooting has been acquitted. Cities investigated after viral incidents follows a similar pattern of media – though stunningly more intense. Media attention in these cities increase from 1.87 articles per month, in the months before the event, to 45.86 articles on the month the investigation is announced. Media attention decreases but it does not return to pre-investigation levels and spikes again 12 months later when the findings of the investigations are typically released. Its plausible that a negative article involving police every three days, and almost 2 per day in the month of the investigation, may be enough to make police officers feel like they are under siege.

[41]The Gallup poll shows that only 52% of Americans surveyed have confidence in the police – the lowest percentage seen in 22 years.

31

and St. Louis), one city investigated without viral incidents of deadly force (Seattle, WA), and two viral shooting cities (Arlington, TX and Cincinnati, OH). For the cities in which we obtained data, we do not have systematic information on the type of call placed or demographics of civilians who placed the calls. We have daily counts of the number of calls for service.

Appendix Figure 11 plots 911 call volume before and after the event, where the event is either an investigation or a viral shooting. We test whether the data exhibits either a level or trend break in the month of the event. For Baltimore, Chicago and St. Louis, the average number of 911 calls per 100,000 the day before the investigation was 459.77 and 434.07 the day after the investigation. Adjusting for total crime rates, the average call rate was 519.11 a day before the investigation and 455.10 the day after. Notably, the average call rate (adjusted for crime) slowly increases and is at the pre-investigation daily average three weeks after the investigation is announced.[42] Given 911 call volume is back to pre-event levels in less than a month, it is unlikely it explains increases in crime that lasts twenty four months.

6.3 Changes in the Nature of Consent Decrees

A final potential mechanism is that the nature of investigations and the resulting consent decrees (when a pattern-or-practice of unconstitutional policing has been found) have become more aggressive or targeted toward the type of policing that is highly correlated with crime. In this scenario, crime increases because police feel under appreciated or micromanaged and this lowers their effectiveness.

Some police departments report that the investigation and consent decree process left them a significantly better police department (Stone et al., 2009), while others report that the process "has made police less active and aggressive in fighting crime" (Davis et al., 2005). In the latter case, 79% of officers reported reducing proactive policing.[43]

The "evidence" on changes in consent decrees not acting as a potential mechanism to explain our set of facts is clear. We read and coded variables on number of major total sections, number of sections concerning officer use of force, number of sections concerning civilian stops, and total pages from seventeen consent decrees – all those that had an agreement and make it to our final analysis sample of principal cities. There are two important limitations to this approach. First, the

---

[42]Comparing our results to Desmond et al. (2016), they find that average weekly calls adjusted for crimes decline by -0.0088 post their event and also slowly increase after. However, it takes approximately thirty weeks for weekly calls in their analysis to reach pre-event levels.

[43]Consistent with our general hypothesis, crime in Pittsburgh increased by 306.82 as a result of the investigation.

extensiveness of a consent decree may be a poor proxy of the intensity of an investigation. Second, using this approach, we cannot measure the extensiveness of investigations that were closed without finding a pattern-or-practice.

With these caveats in mind, Appendix Table 5 displays the data collected from each consent decree. Column (1) includes thirteen consent decrees for investigations not preceded by viral incidents of deadly force, column (2) includes the four consent decrees that have been agreed upon for investigations succeeding viral incidents of deadly force (Baltimore, Cincinnati, Riverside and Ferguson – Chicago is still in negotiations), and column (3) reports the p-values on the difference between columns (1) and (2). The rows of the table represent different measures extracted from the consent decrees.

The data suggest that findings reports for the two groups of investigations look statistically similar. Thus, the nature of consent decrees do not necessarily impact crime.

7   Conclusion

Rooting out bias and ensuring constitutional policing is one of the most important issues of our time. Pattern-or-practice investigations – a way to police the police – are a key tool to accomplish this. On average these investigations have negligible impacts on subsequent homicide or total crime rates. But, as we have illustrated throughout, this aggregate number masks important heterogeneity.

The incident that sparks an investigation into a police department is an important determinant of how the investigation will impact policing and crime. For investigations that are sparked by mostly civilian complaints, allegations, lawsuits or media reports of excessive force, investigations caused a statistically significant decline in homicide and total crime rates. These investigations saved lives – 61 per investigation, in the 24 months following investigations.

For the five investigations that were sparked by nationally visible incidents of deadly use of force – Baltimore, Chicago, Cincinnati, Ferguson and Riverside– investigations cause statistically significant increases in both homicide and total crime. Contrary to other investigations, investigations during this time lost lives – 179 of them, per investigation, in the 24 months following the start of the investigation. That's 893 total. Almost 900 individuals whose potential may not have been realized. And, we are still counting. A back-of-the-envelope calculation suggests that these five cities converge to pre-investigation levels 51 months after the investigation and, by that time, almost 1214 excess homicides will have occurred.

The leading theory for why some investigations have led to an increase in crimes is a striking decrease in the quantity of police activity – which is evident in all cities we were able to collect data. All other theories considered contradict the data in important ways, though lack of complete data makes definitive conclusions illusive.

It is important to emphasize, however, that neither our data nor our analysis makes any claim regarding the net social cost or benefit of pattern-or-practice investigations. That's well beyond the scope of this paper. Indeed, many argue that the federal government has been an important catalyst for greater equality in housing, labor markets, marriage, and voting rights.

Despite the lack of welfare analysis, we hope these results encourage introspection on the trade-offs involved when we increase scrutiny on police departments, particularly in the midst of civil unrest. The social objective is to eliminate bias without causing police to retreat from activities that suppress crime, and save lives. A troubling possibility is that the types of police activities that keep crime rates low are inherently unconstitutional and hence we face a tradeoff between allowing uncomfortable amounts of police bias and reducing crime in the very communities which are most impacted by that bias.

One way forward is to design a set of incentives such that we increase the penalties of unconstitutional policing and, simultaneously, lower the probability of being wrongfully accused when controversial interactions occur. In this sense, we might keep the expected price of policing constant for officers. There is no free lunch. If the price of policing increases, officers are rational to retreat. And, retreating disproportionately costs black lives.

References

[1] Abadie, A., Diamond, A., & Hainmueller, J. 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.

[2] Ayres, I. and Levitt, S.D., 1998. "Measuring positive externalities from unobservable victim precaution: an empirical analysis of Lojack." *The Quarterly Journal of Economics*, 113(1), pp.43-77.

[3] Baron, D.P. and Myerson, R.B., 1982. "Regulating a monopolist with unknown costs." *Econometrica: Journal of the Econometric Society*, pp.911-930.
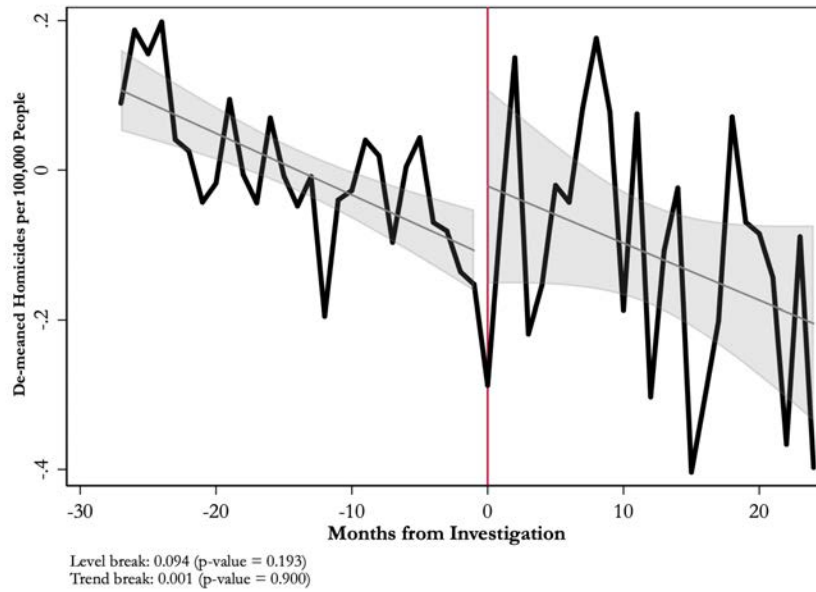
[4] Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach". *Journal of Political Economy 76:2*, 169-217.

[5] Cannon, Lou; Lee, Gary (May 2, 1992). "Much Of Blame Is Laid On Chief Gates". The Washington Post.

[6] Cheng, C. and Long, W. 2018. "The Spillover Effects of Highly Publicized Police Use-of-Force on Policing and Crime: Evidence from St. Louis".

[7] California Department of Justice. 2016. "Attorney General Kamala D. Harris Opens Investigations into the Kern County Sheriff's Office and the Bakersfield Police Department for Potential Civil Rights Violations".

[8] CNN. 2019. 10 years. 180 school shootings. 356 victims.

[9] Congressional Research Service. 2019. Recent Trends in Active-Duty Military Deaths.

[10] Davis, R.C., Henderson, N.J. and Ortiz, C.W., 2005. *Can Federal Intervention Bring Lasting Improvement in Local Policing?: The Pittsburgh Consent Decree.* New York, NY: Vera Institute of Justice.

[11] Dehejia, R. H., and Wahba, S. 2002. "Propensity score-matching methods for nonexperimental causal studies." *Review of Economics and Statistics, 84(1)*, pp. 151-161.

[12] Desmond, M., Papachristos, A.V. and Kirk, D.S., 2016. "Police violence and citizen crime reporting in the black community". *American Sociological Review, 81(5)*, pp.857-876.

[13] Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics, 7(1),* 1-26.

[14] Firpo, S., and Possebom, V. 2017. "Synthetic control method: Inference, sensitivity analysis and confidence sets".

[15] Fryer, Roland G. 2018." Police Brutality and Crime: Evidence from the First Viral Video".

[16] Fisher, Ronald A. 1935. *The Design of Experiments.* Edinburgh: Oliver and Boyd, Ltd, 1951 (6e).

[17] Ford, M. 2017. "America's Uneven Crime Spike". *The Atlantic.*

[18] Hawkins, D. 2016. "'Ferguson effect'? Savagely beaten cop didnt draw gun for fear of media uproar, says Chicago police chief". *The Washington Post*.

[19] Heckman, J. J., Ichimura, H., and Todd, P. 1998. "Matching as an econometric evaluation estimator." *The Review of Economic Studies, 65(2)*, pp. 261-294.

[20] Holm, S. 1979. A simple sequential rejective multiple test procedure. *Scandinavian Journal of Statistics 6:65-70.*

[21] The Independent Commission on the Los Angeles Police Department. "Report of the Independent Commission on the Los Angeles Police Department" (Christopher Commission Report). 1991.

[22] Jacobson, L.S., LaLonde, R.J. and Sullivan, D.G., 1993. "Earnings losses of displaced workers." *The American Economic Eeview*, pp.685-709.

[23] Jencks, C. and Peterson, P.E. eds., 2001. *The Urban Underclass*. Brookings Institution Press.

[24] Kindy, K., Fisher M., Tate J., and Jenkins J.. 2015. "A Year of Reckoning: Police FatallyShoot Nearly 1,000." *The Washington Post.*

[25] Laffont, J.J. and Tirole, J., 1986. "Using cost observation to regulate firms." *Journal of Political Economy, 94* (3, Part 1), pp.614-641.

[26] Mas, Alexander. 2006. "Pay, Reference Points, and Police Performance." *Quarterly Journal of Economics*, CXXI (3), 783-821.

[27] Maxwell, C. and Solomon D. 2018. "Expanding the Authority of State Attorneys General to Combat Police Misconduct".

[28] Morgan, S. L., and Winship, C. *Counterfactuals and Causal Inference: Methods and Principles for Social Research. 2nd ed.* Cambridge University Press; 2015.

[29] Nix, J. and Pickett, J.T., 2017. "Third-person perceptions, hostile media effects, and policing: Developing a theoretical framework for assessing the Ferguson effect". *Journal of criminal justice, 51*, pp.24-33.

[30] Park, H., and Lee, J. 2017. "Looking for Accountability in Police-Involved Deaths of Blacks". *The New York Times.*

[31] Prendergast, C., 2002. "The tenuous trade-off between risk and incentives." *Journal of political Economy, 110(5)*, pp.1071-1102.

[32] Pyrooz, D.C., Decker, S.H., Wolfe, S.E. and Shjarback, J.A., 2016. "Was there a Ferguson Effect on crime rates in large US cities?." *Journal of criminal justice, 46*, pp.1-8.

[33] Rosenbaum, P. R., and Rubin, D. B. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika, 70(1)*, pp. 41-55.

[34] Shi, L., 2009. "The limit of oversight in policing: Evidence from the 2001 Cincinnati riot." *Journal of Public Economics, 93(1-2)*, pp.99-113.

[35] Shjarback, J.A., Pyrooz, D.C., Wolfe, S.E. and Decker, S.H., 2017. "De-policing and crime in the wake of Ferguson: Racialized changes in the quantity and quality of policing among Missouri police departments." *Journal of criminal justice, 50*, pp.42-52.

[36] Stone, C., Foglesong, T.S. and Cole, C.M., 2009. *Policing Los Angeles under a consent degree: The dynamics of change at the LAPD.* Cambridge, MA: Program in Criminal Justice Policy and Management, Harvard Kennedy School.

[37] United States Department of Justice Civil Rights Division. 2017. *The Civil Rights Division's Pattern and Practice Police Reform Work: 1994-Present.* United States Department of Justice Civil Rights Division.

Figure 1: Estimating Level and Trend Breaks, All Investigations

Panel A: Homicides per 100,000 people

Level break: 0.094 (p-value = 0.193)
Trend break: 0.001 (p-value = 0.900)

Panel B: Total Crime per 100,000 people

Level break: 9.488 (p-value = 0.008)
Trend break: 0.002 (p-value = 0.993)

Notes: The black line plots average outcome in investigated cities, weighted by population, in a (-27,24) month window around the investigation for all cities that were investigated between 1994 and 2016 and have available data in the relevant window. This includes 27 investigations in 25 principal cities. The vertical axis plots the outcome variable of interest in each month, de-meaned by the average outcome in that calendar month in the pre-event period to account for seasonality in criminal behavior. The horizontal axis plots the month from the announcement of the investigation. Linear trends, along with a 95% confidence band, are plotted along with the outcome.

Figure 2: Heterogeneity in Level Breaks at Month of Investigation

Panel A: Homicides per 100,000 people
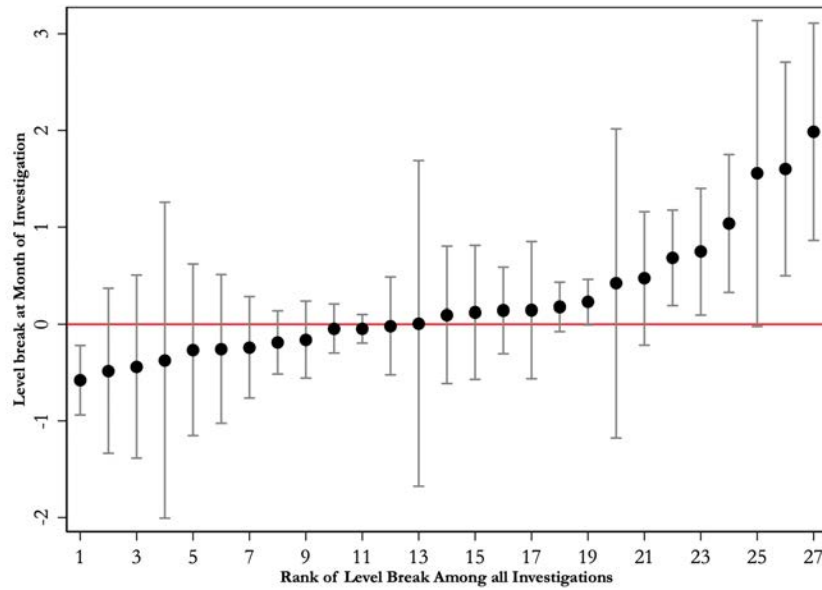
Panel B: Total Crime per 100,000 people
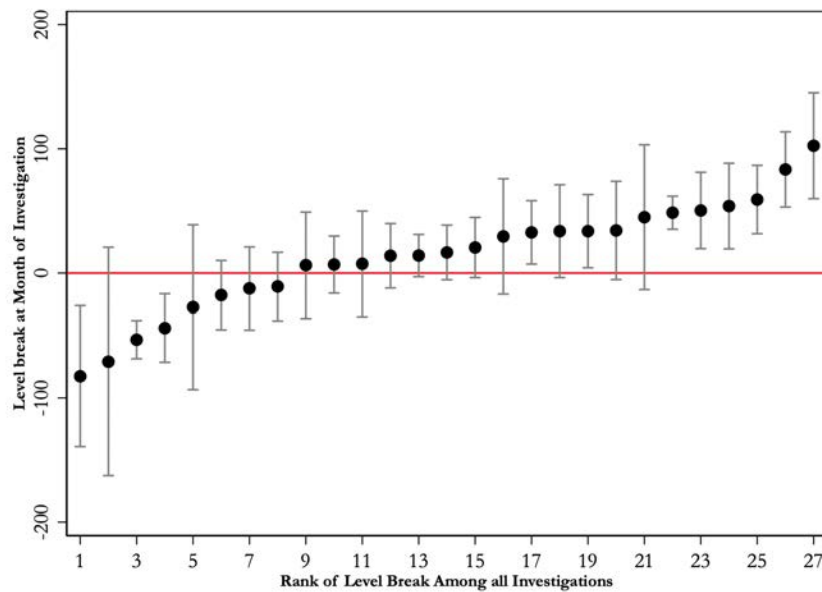
Notes: The vertical axis plots the size of the level break at the month of the investigation in each city for 27 separate investigations in 25 principal cities from 1994-2016 along with their 95% confidence intervals. The level breaks are plotted according to increasing size. The estimate for each investigation comes from a regression of the outcome of interest (de-meaned over the pre-period to account for seasonality in criminal behavior) on a linear month variable, an indicator for whether or not a given month is after the investigation, and the interaction of those two variables. The coefficient on the post indicator is plotted here. Confidence intervals are calculated using robust standard errors.
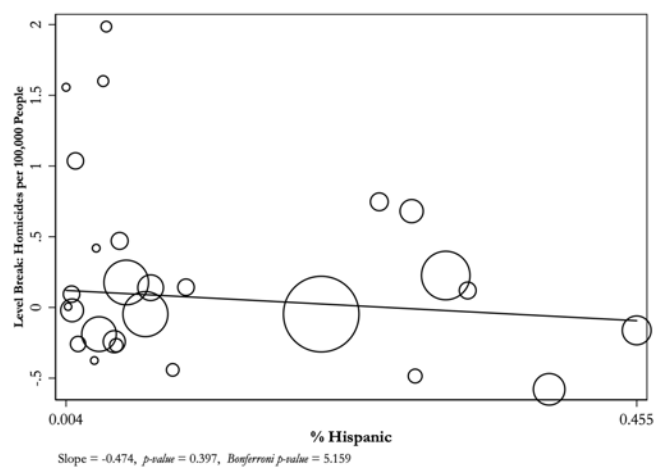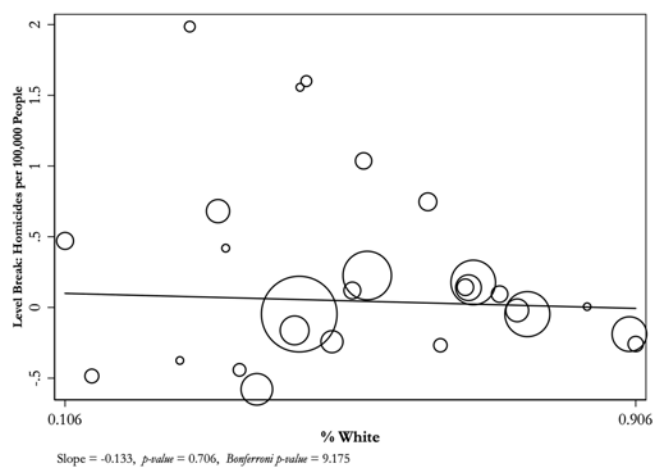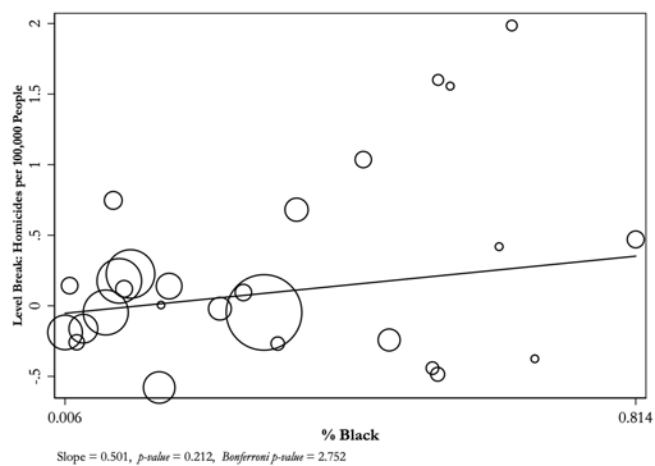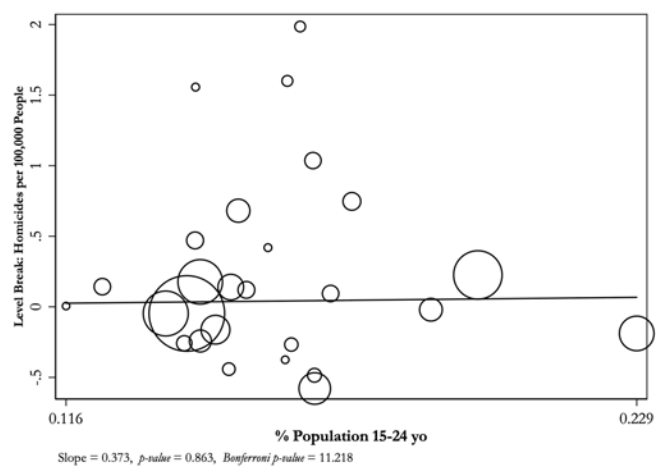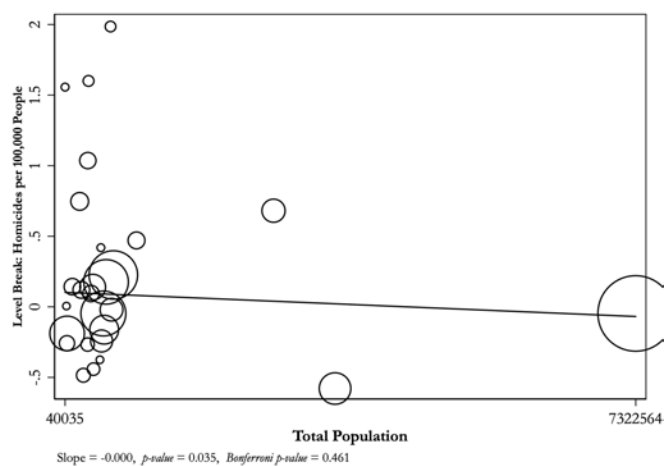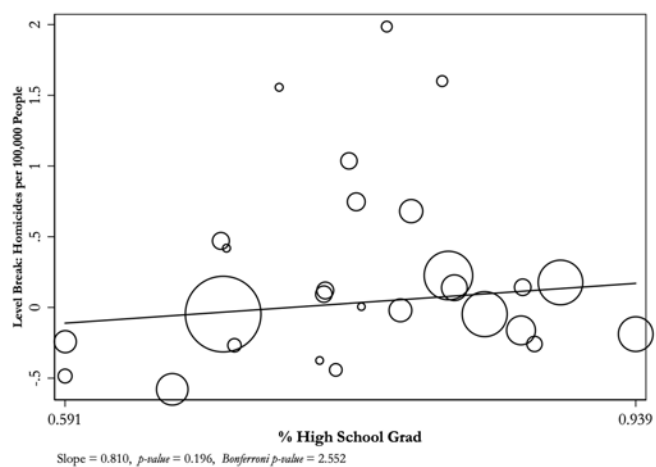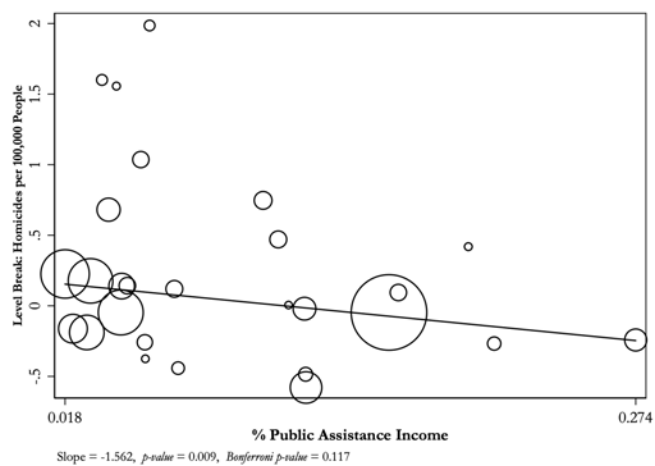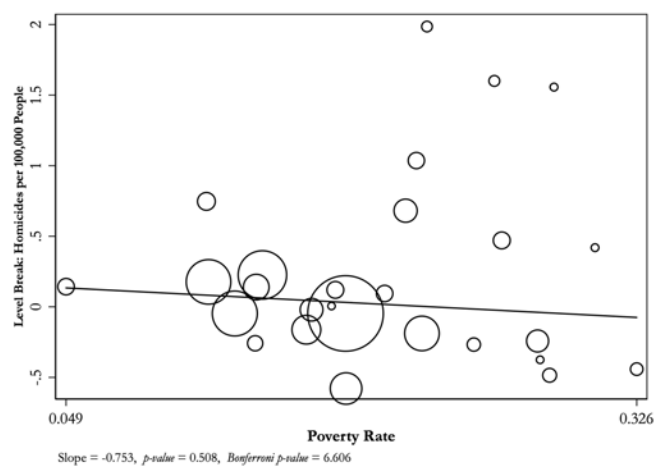
Figure 3A: Heterogeneity in Level Breaks, Homicides

Figure 3A: Heterogeneity in Level Breaks, Homicides

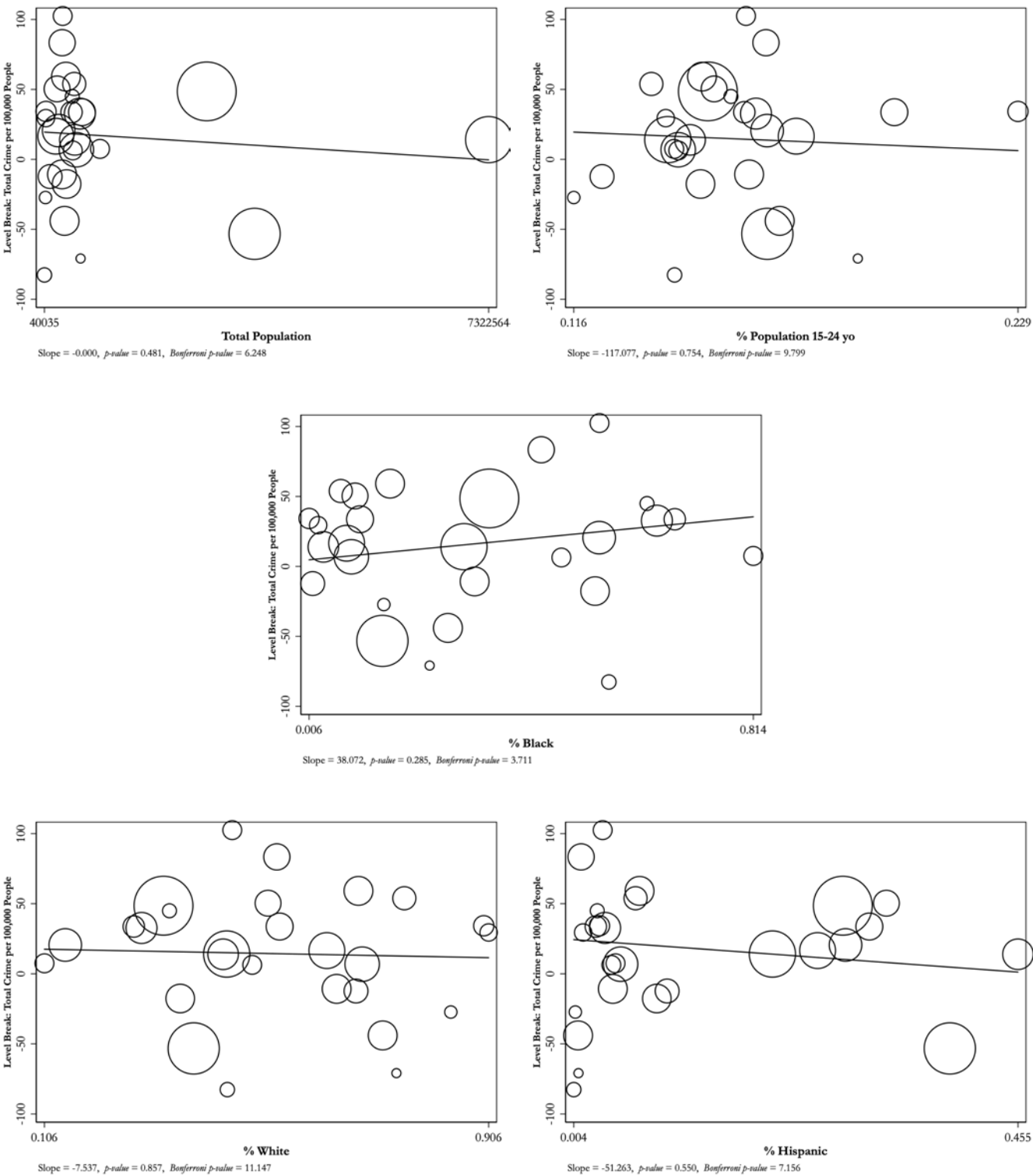Slope = 0.000, *p-value* = 0.576, *Bonferroni p-value* = 7.482

Slope = -0.753, *p-value* = 0.508, *Bonferroni p-value* = 6.606

Slope = -1.562, *p-value* = 0.009, *Bonferroni p-value* = 0.117

Slope = 0.810, *p-value* = 0.196, *Bonferroni p-value* = 2.552

Slope = 1.205, *p-value* = 0.025, *Bonferroni p-value* = 0.319

Figure 3A: Heterogeneity in Level Breaks, Homicides

Slope = 0.001, *p-value* = 0.595, *Bonferroni p-value* = 7.734

Slope = 1.003, *p-value* = 0.310, *Bonferroni p-value* = 4.036

Slope = -0.957, *p-value* = 0.661, *Bonferroni p-value* = 8.588

Figure 3B: Heterogeneity in Level Breaks, Total Crime

Slope = -0.000, *p-value* = 0.481, *Bonferroni p-value* = 6.248

Slope = -117.077, *p-value* = 0.754, *Bonferroni p-value* = 9.799

Slope = 38.072, *p-value* = 0.285, *Bonferroni p-value* = 3.711

Slope = -7.537, *p-value* = 0.857, *Bonferroni p-value* = 11.147

Slope = -51.263, *p-value* = 0.550, *Bonferroni p-value* = 7.156

Figure 3B: Heterogeneity in Level Breaks, Total Crime

Slope = -0.000, *p-value* = 0.635, *Bonferroni p-value* = 8.252

Slope = 14.057, *p-value* = 0.878, *Bonferroni p-value* = 11.413

Slope = -317.885, *p-value* = 0.015, *Bonferroni p-value* = 0.195

Slope = 196.527, *p-value* = 0.087, *Bonferroni p-value* = 1.135

Slope = 147.614, *p-value* = 0.008, *Bonferroni p-value* = 0.103

Figure 3B: Heterogeneity in Level Breaks, Total Crime

Slope = 0.311, *p-value* = 0.206, *Bonferroni p-value* = 2.680

Slope = 86.562, *p-value* = 0.237, *Bonferroni p-value* = 3.081

Slope = 65.825, *p-value* = 0.742, *Bonferroni p-value* = 9.646

Figure 4: Estimating Level and Trend Breaks, Investigations With and Without Viral Incidents of Deadly Force

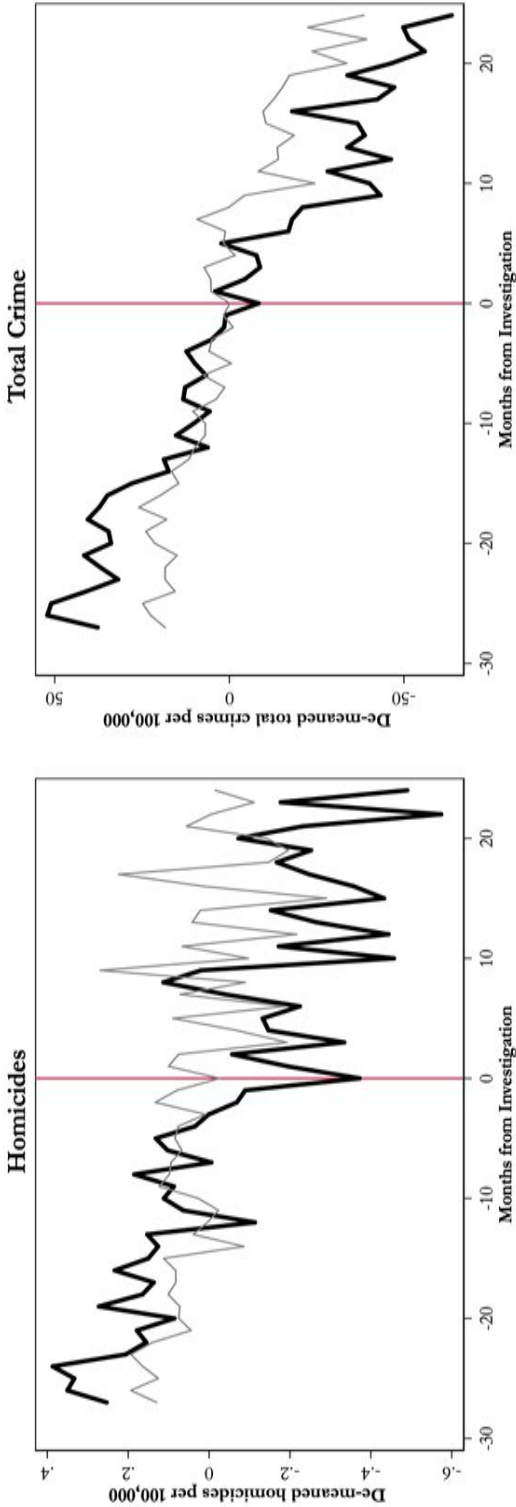*Panel A: Investigations Without Viral Incident of Deadly Force*
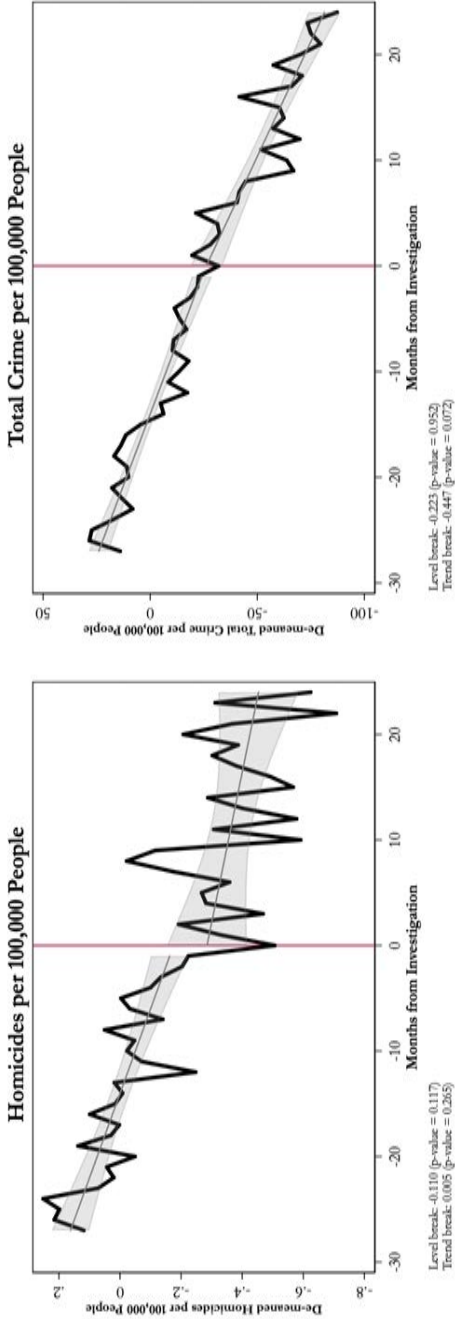
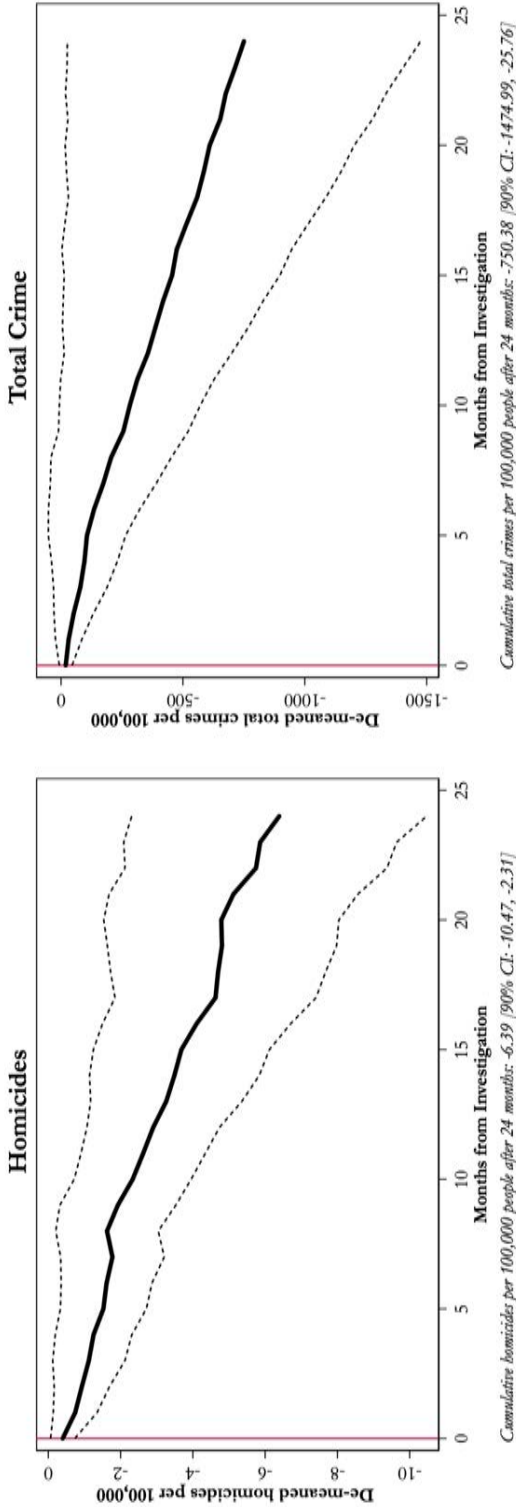*Panel B: Investigations With Viral Incident of Deadly Force*
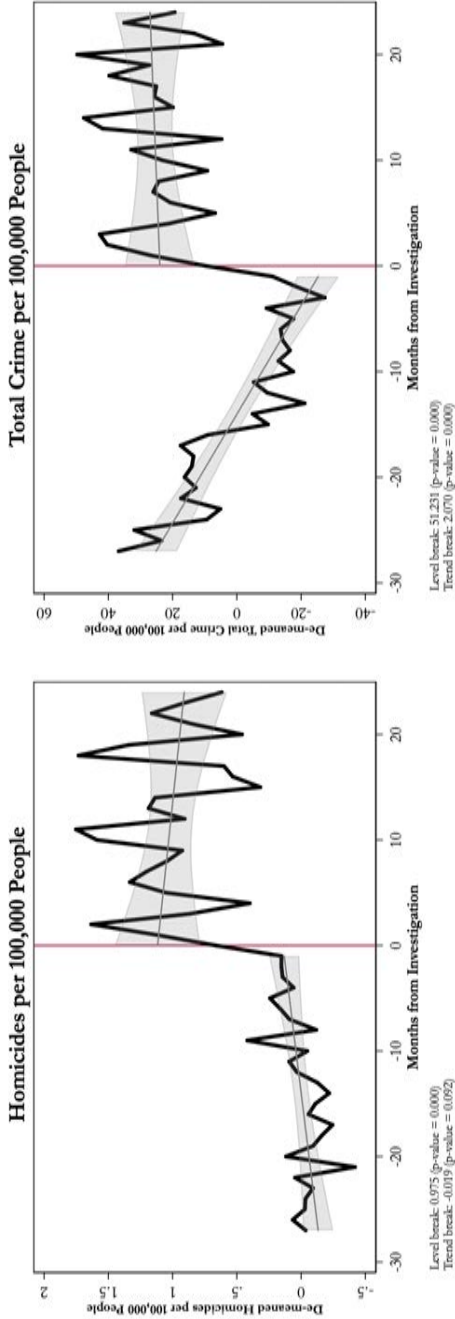
Notes: The black line plots average outcome in investigated cities, weighted by population, in a (−27,24) month window around the investigation for all cities that have available data and did not have a viral incident of deadly force preceding the investigation in panel A and cities that had a viral incident of deadly force preceding the investigation for panel B. This includes 22 investigations in 20 principal cities for panel A and 5 investigations in 5 principal cities for panel B. The vertical axis plots the outcome variable of interest in each month, de-meaned by the average outcome in that calendar month in the pre-event period to account for seasonality in criminal behavior. The horizontal axis plots the month from the announcement of the investigation. Linear trends, along with a 95% confidence band, are plotted along with the outcome.

Figure 5: The Impact of Investigations on Homicide and Total Crime, Investigations Without Viral Incident of Deadly Force, Pooled

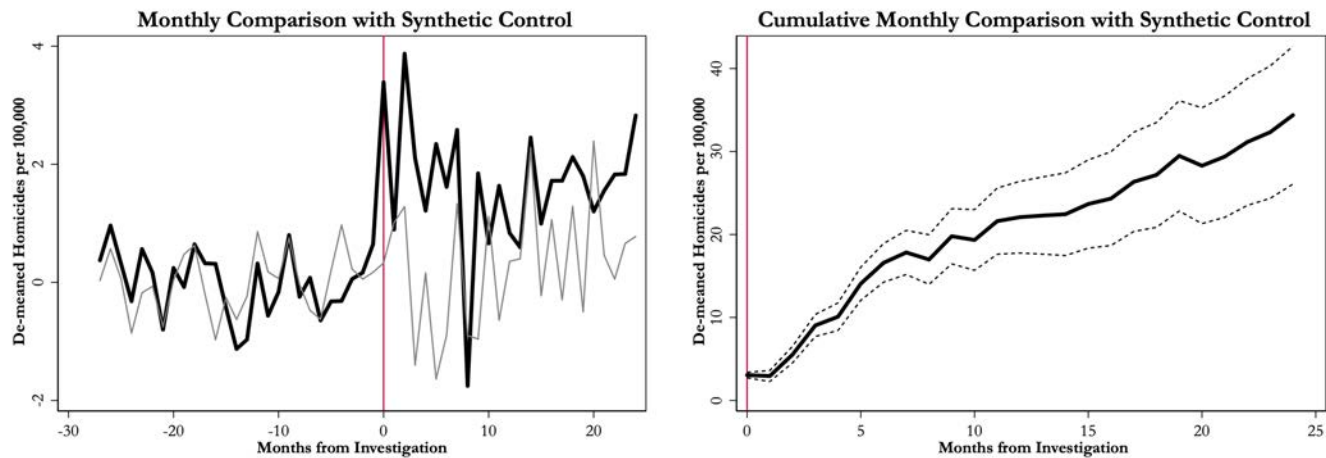*Panel A: Monthly Comparison with Synthetic Control, Controlled*

Panel B: Cumulative Monthly Comparison with Synthetic Control, Controlled

Cumulative homicides per 100,000 people after 24 months: -6.39 [90% CI: -10.47, -2.31]

*Cumulative total crimes per 100,000 people after 24 months: -750.38 [90% CI: -1474.99, -25.76]*

Notes: Panel A plots regression adjusted average monthly outcomes in investigated cities and their synthetic controls for all cities that were investigated but were not preceded by a viral incident of deadly force and have available data in the relevant window. This includes 22 investigations in 20 cities. The black line plots average outcome in investigated cities while the gray line plots average outcome from the corresponding synthetic control cities for each investigation, constructed using the procedure described in Abadie, Diamond, and Hainmueller (2010) using the companion *Synth* package for Stata. Panel B plots regression-adjusted estimates of the average cumulative differences in outcomes between investigated cities and their relevant synthetic control for 24 months following the investigation for all cities that were investigated and were not preceded by a viral shooting and have available data in the relevant window. All regressions control for education, economic, and race indices – for details on variable construction see the data appendix.

Figure 6A: The Impact of Investigations on Homicide, Investigations With Viral Incident of Deadly Force

Panel A: Baltimore, MD

Panel B: Chicago, IL
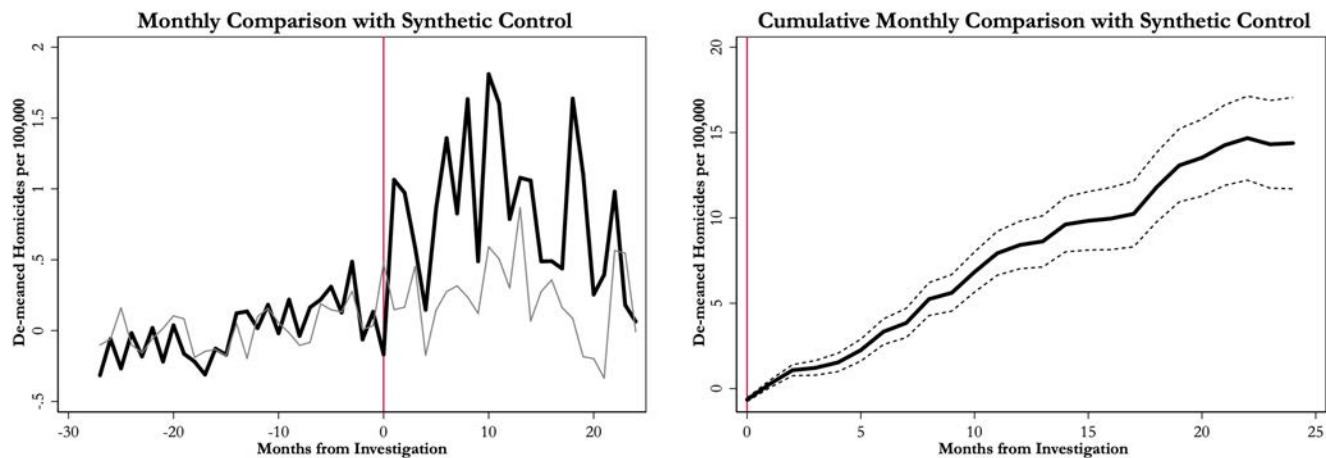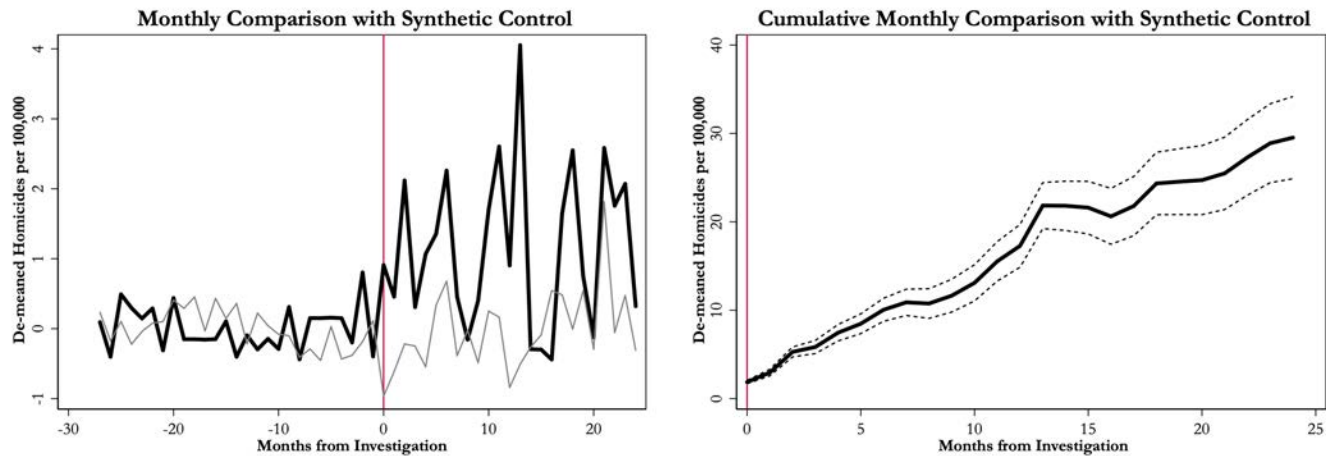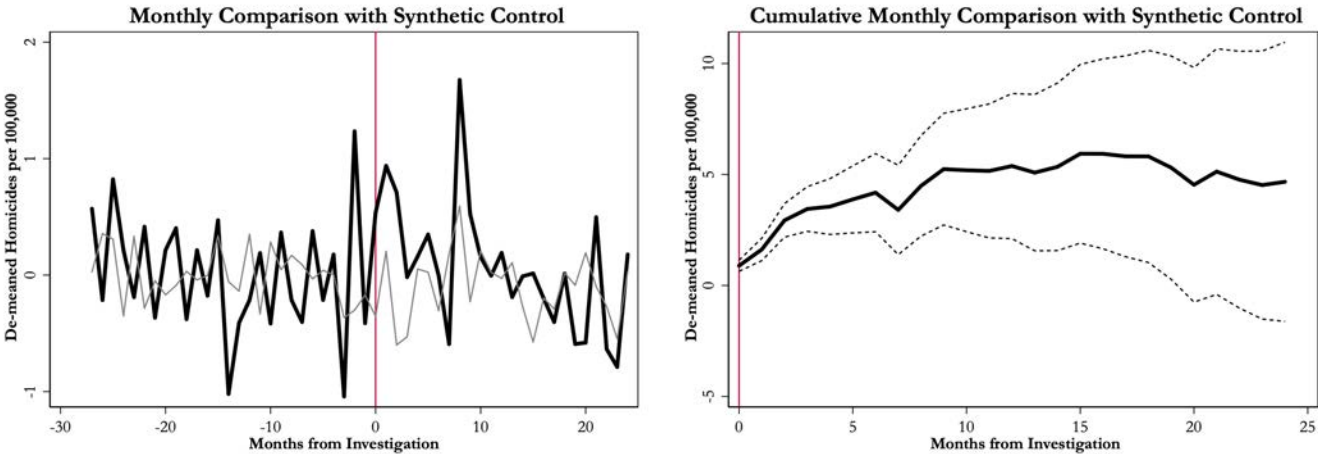

Panel C: Cincinnati, OH

Figure 6A: The Impact of Investigations on Homicide, Investigations With Viral Incident of Deadly Force

Panel D: Riverside, CA
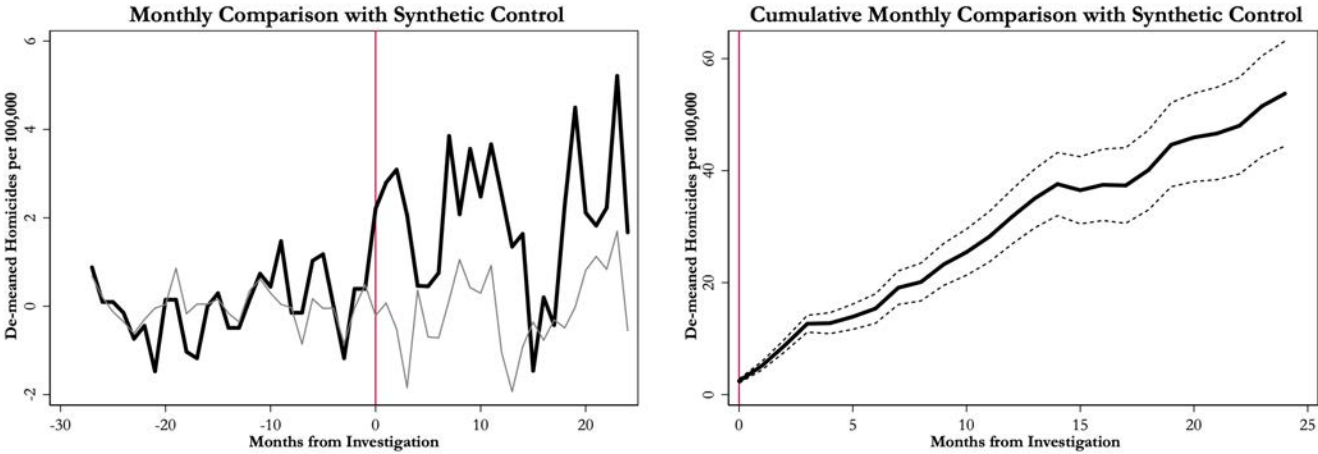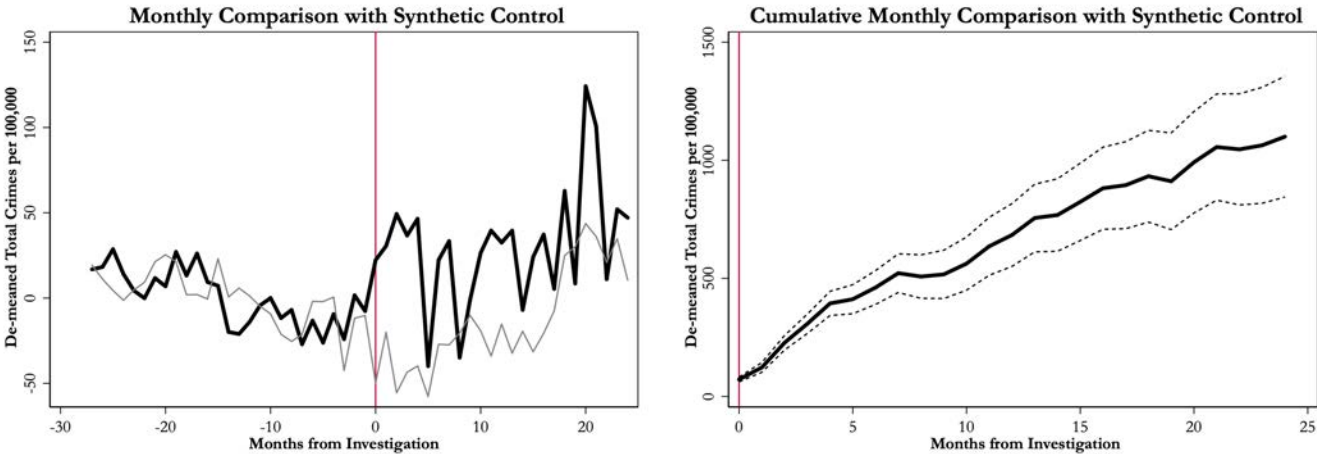

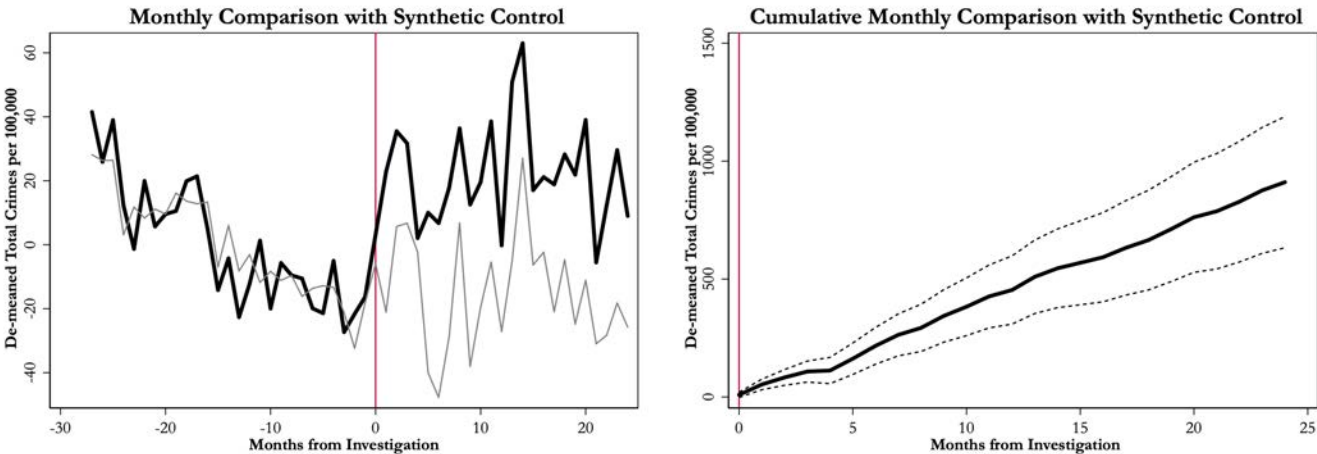*Panel E: St. Louis, MO*

Figure 6B: The Impact of Investigations on Total Crime, Investigations With Viral Incident of Deadly Force

Panel A: Baltimore, MD

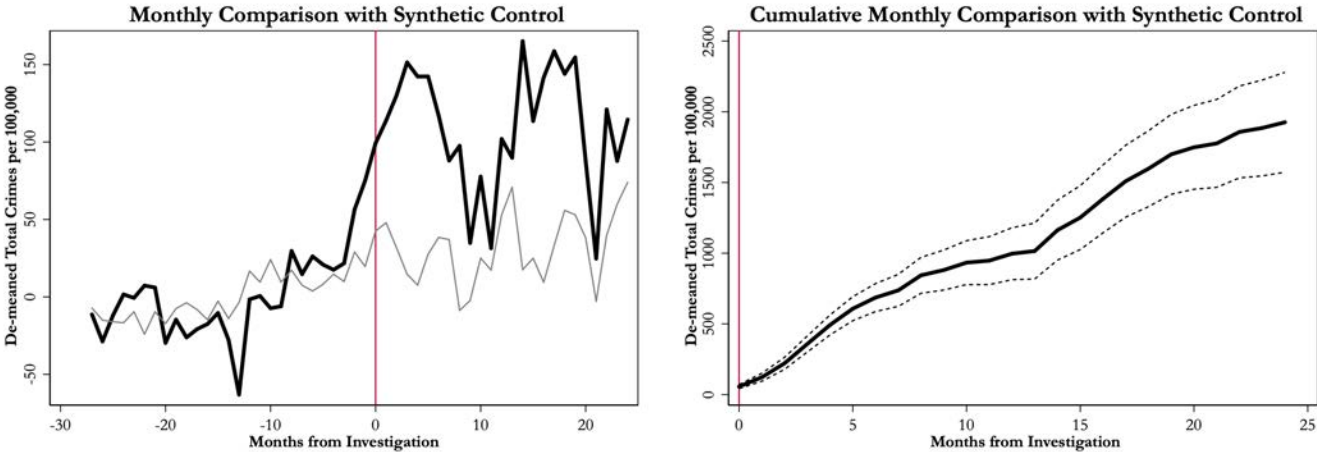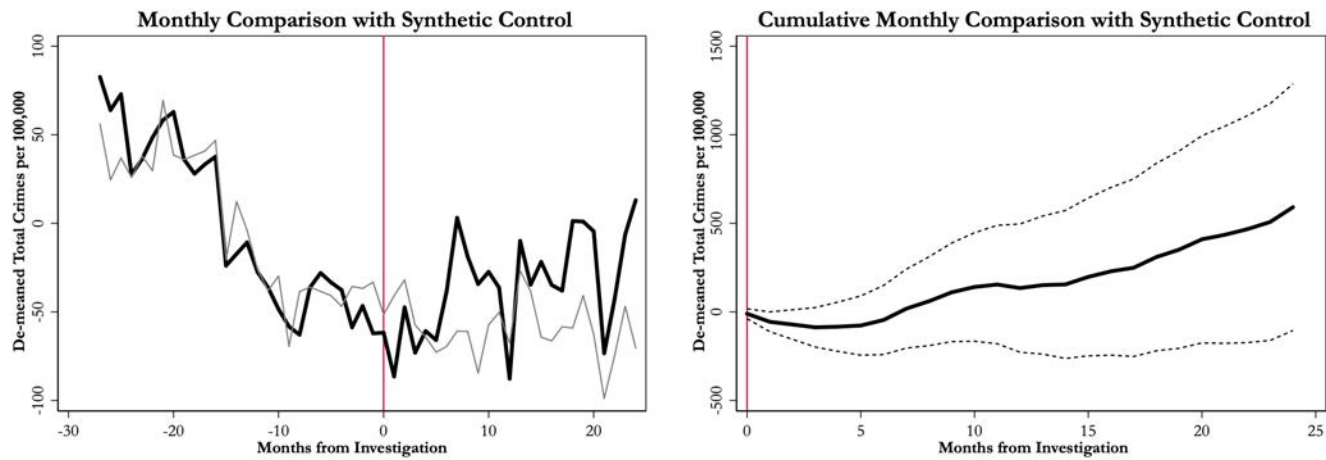*Panel B: Chicago, IL*

Panel C: Cincinnati, OH

Figure 6B: The Impact of Investigations on Total Crime, Investigations With Viral Incident of Deadly Force

Panel D: Riverside, CA
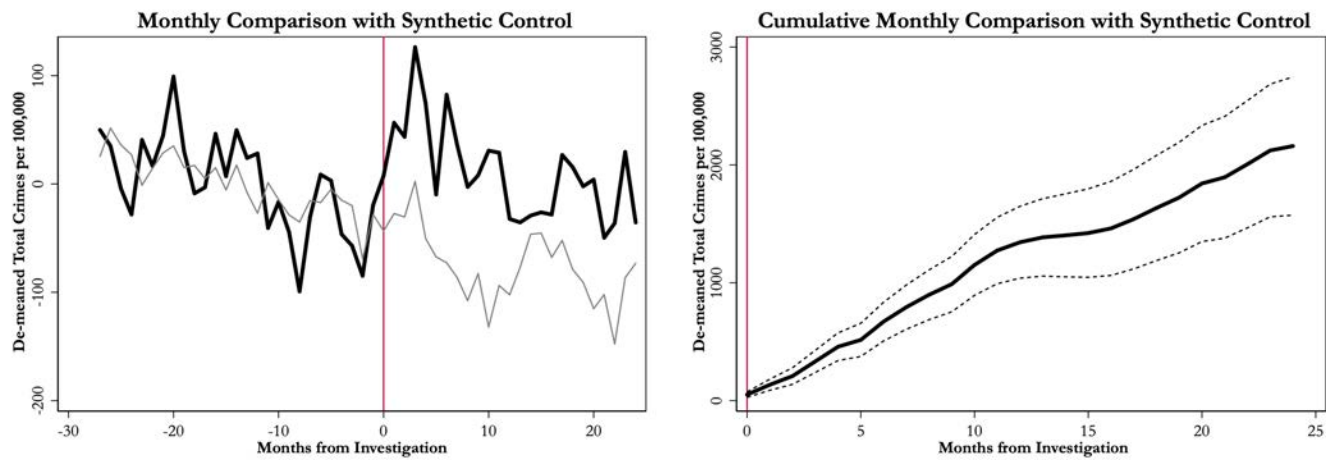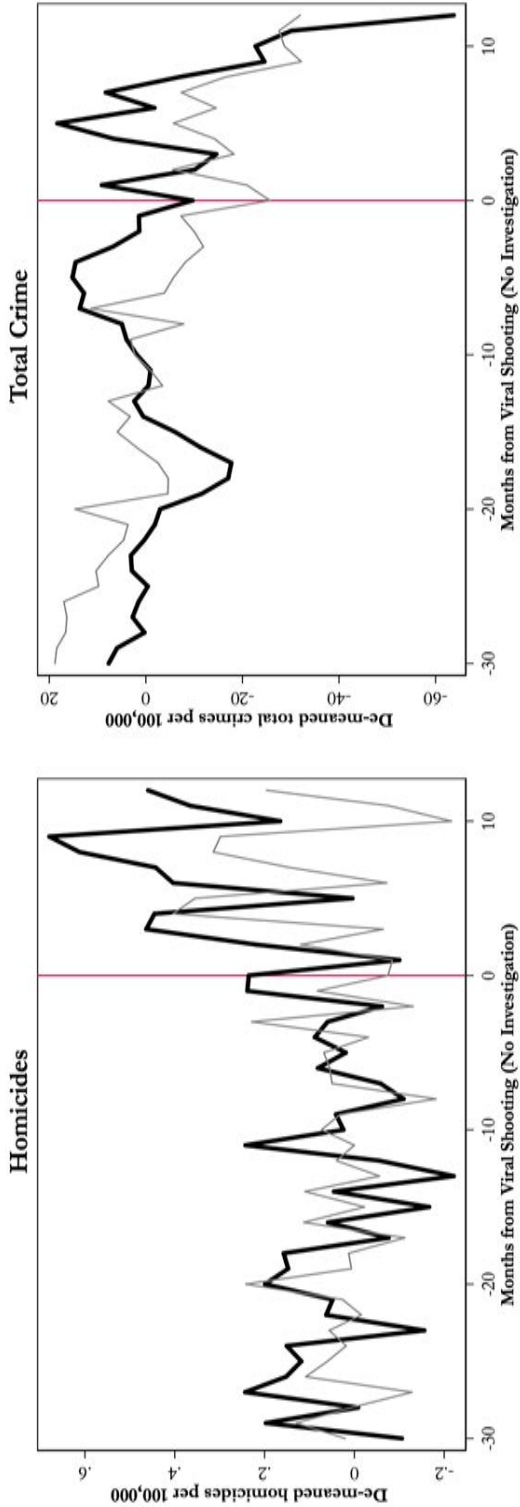

Panel E: St. Louis, MO

Figure 7: The Impact of Viral Shootings on Homicide and Total Crime

Panel A: *Monthly Comparison with Synthetic Control, Controlled*

Panel B: *Cumulative Monthly Comparison with Synthetic Control, Controlled*
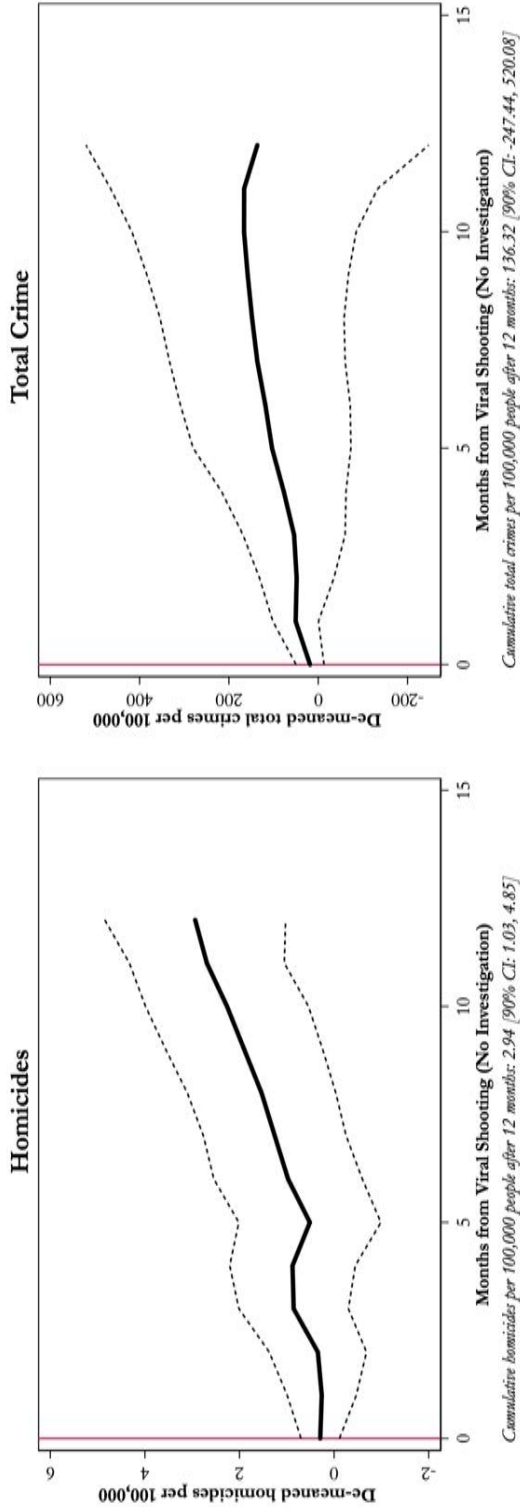
Notes: Panel A plots regression adjusted average monthly outcomes in cities that had viral shootings but no investigations and their synthetic controls, provided they had data in the relevant window. This includes 8 viral shootings in 8 cities. The black line plots average outcome in cities with shootings while the gray line plots average outcome from the corresponding synthetic control cities for each shooting, constructed using the procedure described in Abadie, Diamond, and Hainmueller (2010) using the companion *Synth* package for Stata. Panel B plots regression-adjusted estimates of the average cumulative differences in outcomes between cities with viral shootings and their relevant synthetic control for 12 months following the shooting. All regressions control for education, economic, and race indices – for details on variable construction see the data appendix.

Figure 8: The Impact of Shooting Unarmed Blacks on Crime in Investigated Cities Preceded by a Viral Incident of Deadly Force
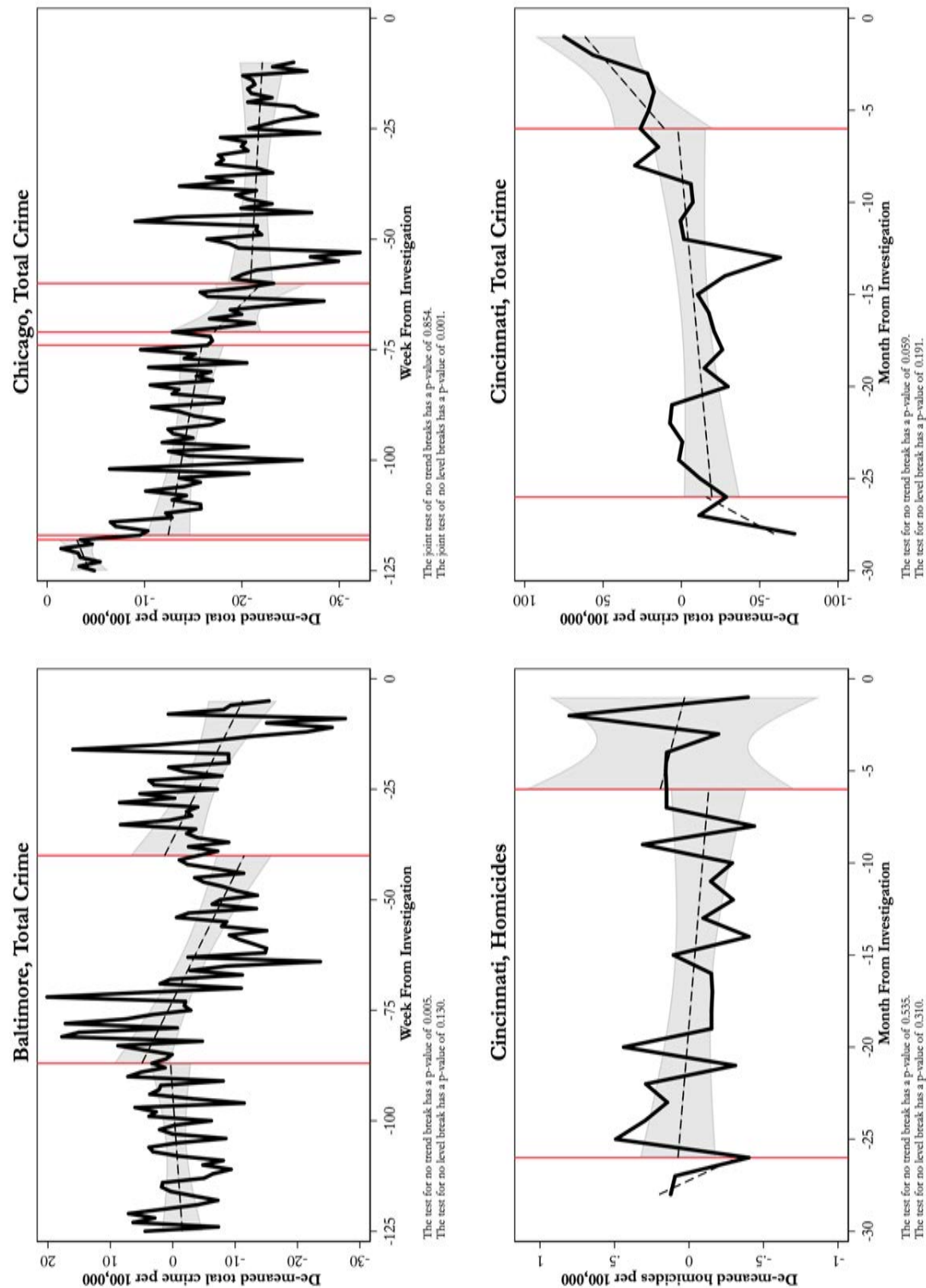
Notes: Each plot displays changes in average weekly total crime rates in Baltimore and Chicago over a period after the #BLM movement began and before the investigation in each city: relative to the investigation, the windows are (-125, -5) in Baltimore and (-125, -10) in Chicago. Each window ends before the events that sparked the investigation in each city (i.e. the death of Freddie Gray in Baltimore and the release of the video of the shooting of Laquan McDonald in Chicago). Crime rates, plotted on the vertical axis, are de-meaned by police department and month to account for seasonality in criminal behavior. The vertical lines indicate weeks in which there was a police shooting of a black civilian. Linear trends, along with a 95% confidence band, are plotted along with the raw crime rate.

Figure 9: Police Activity, Investigation Without Viral Incident of Deadly Force and Viral Shooting Cities

Panel A: Albuquerque, NM (Investigation Without Viral Incident of Deadly Force)

Note: The test for no trend break has a (two-sided) p-value of 0.607.
The test for no level break has a (two-sided) p-value of 0.674.
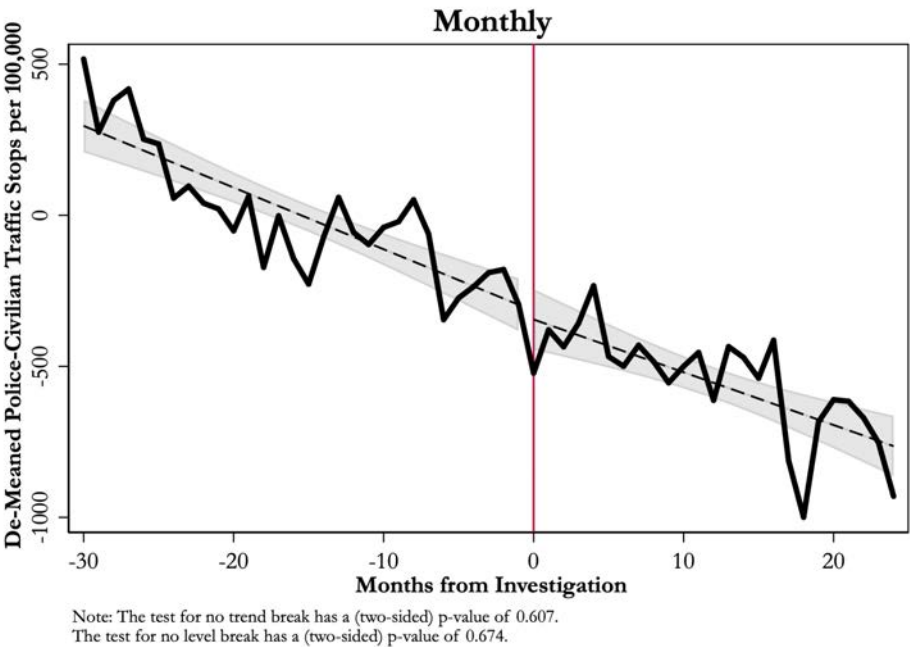
Panel B: Arlington, TX (Viral Shooting)
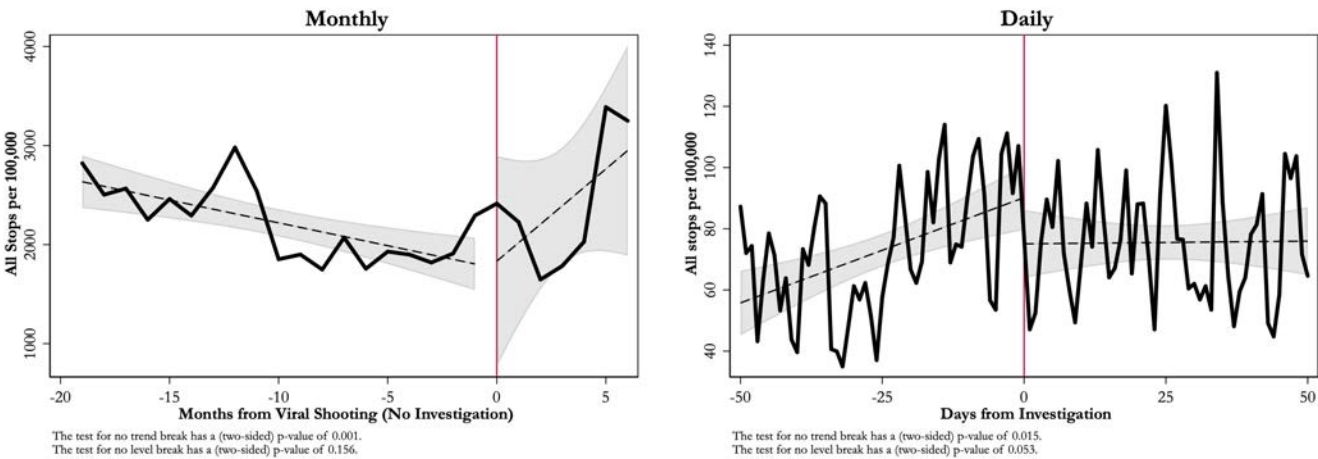
The test for no trend break has a (two-sided) p-value of 0.001.
The test for no level break has a (two-sided) p-value of 0.156.

The test for no trend break has a (two-sided) p-value of 0.015.
The test for no level break has a (two-sided) p-value of 0.053.

Figure 9: Police Activity, Investigation Without Viral Incident of Deadly Force and Viral Shooting Cities

Cincinnati, OH (Viral Shooting)

The test for no trend break has a (two-sided) p-value of 0.048.
The test for no level break has a (two-sided) p-value of 0.003.

The test for no trend break has a (two-sided) p-value of 0.000.
The test for no level break has a (two-sided) p-value of 0.224.
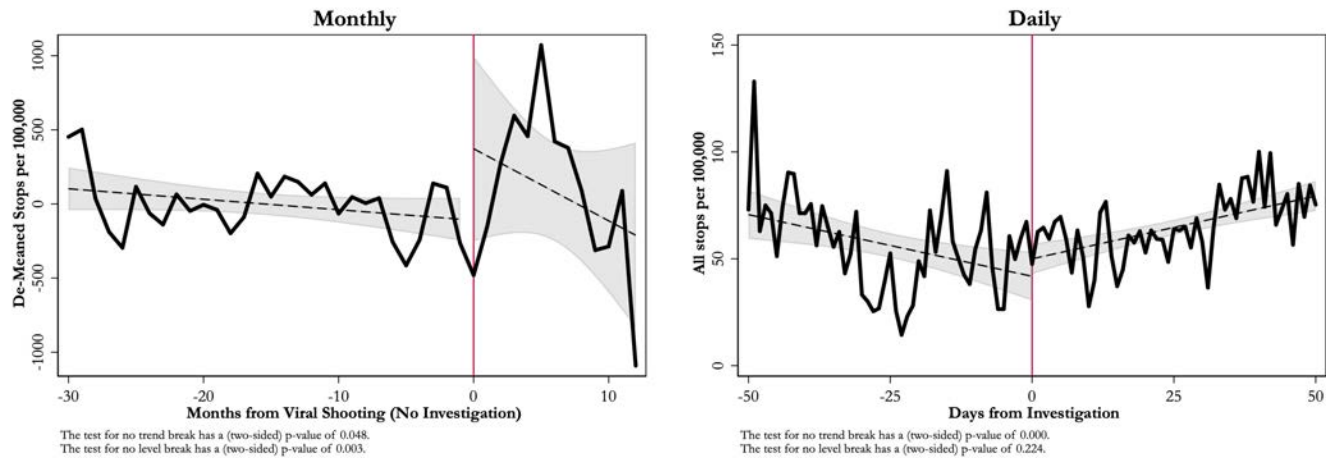
St. Paul, MN (Viral Shooting)

The test for no trend break has a (two-sided) p-value of 0.006.
The test for no level break has a (two-sided) p-value of 0.090.

The test for no trend break has a (two-sided) p-value of 0.000.
The test for no level break has a (two-sided) p-value of 0.000.
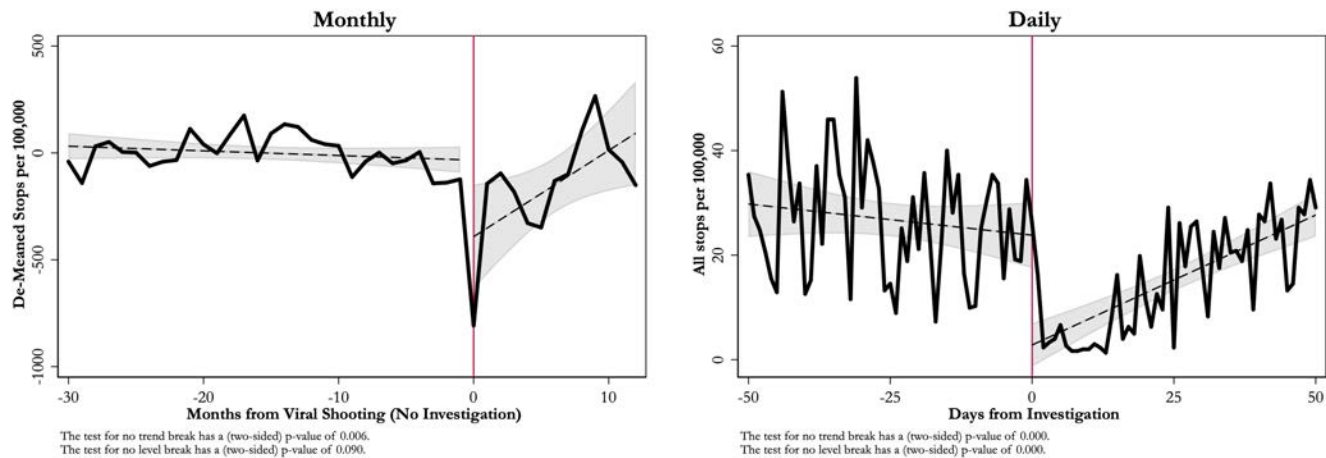
Charlotte, NC and Tulsa, OK (Viral Shooting)

The test for no trend break has a (two-sided) p-value of 0.129.
The test for no level break has a (two-sided) p-value of 0.257.

The test for no trend break has a (two-sided) p-value of 0.183.
The test for no level break has a (two-sided) p-value of 0.577.
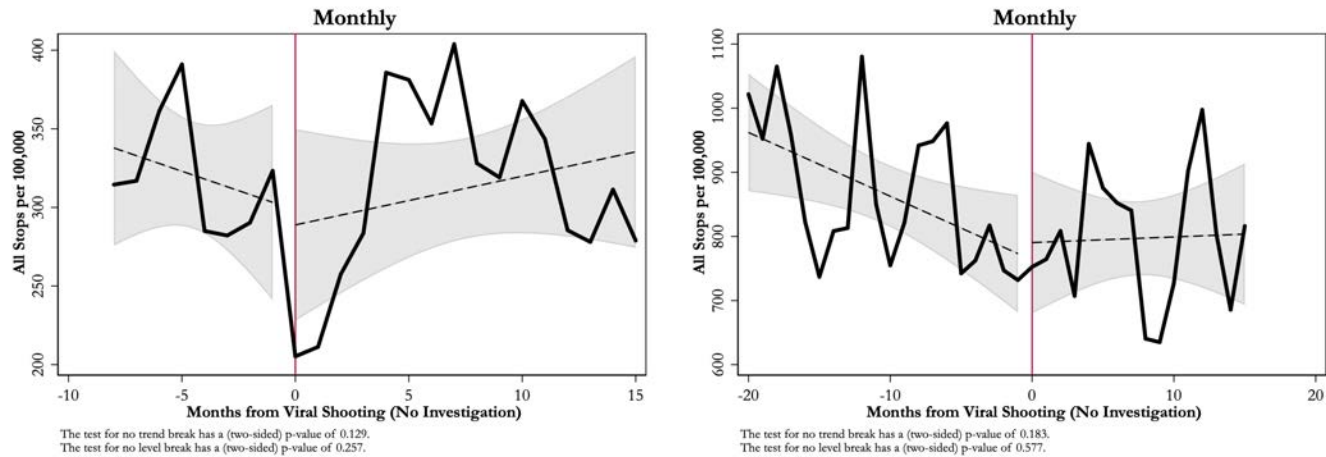
Figure 10: Police Activity, Investigation With Viral Incident of Deadly Force

Chicago, IL
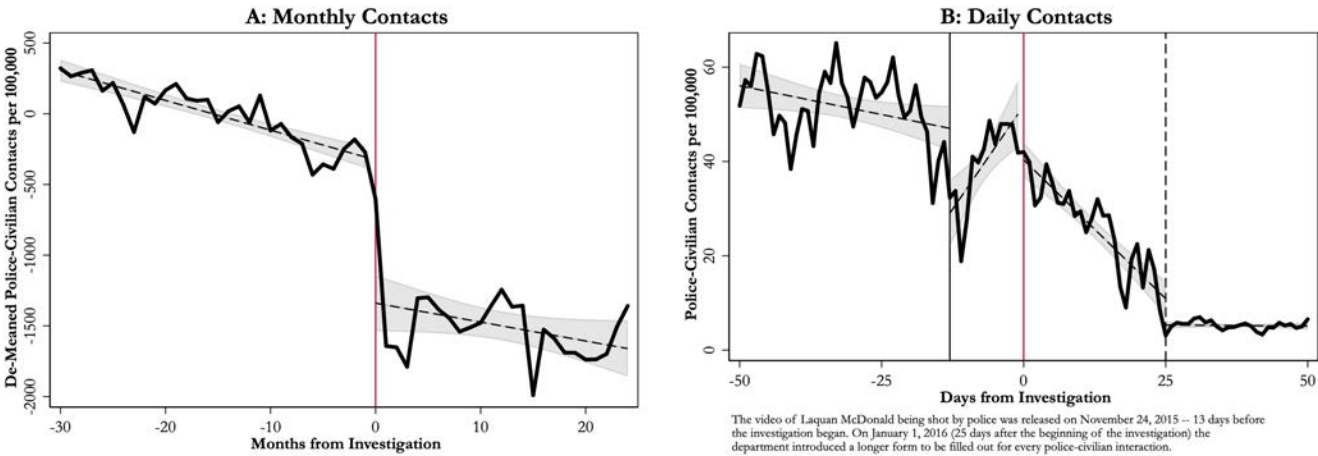
A: Monthly Contacts

B: Daily Contacts

The video of Laquan McDonald being shot by police was released on November 24, 2015 -- 13 days before the investigation began. On January 1, 2016 (25 days after the beginning of the investigation) the department introduced a longer form to be filled out for every police-civilian interaction.

Riverside, CA

C: Monthly Contacts
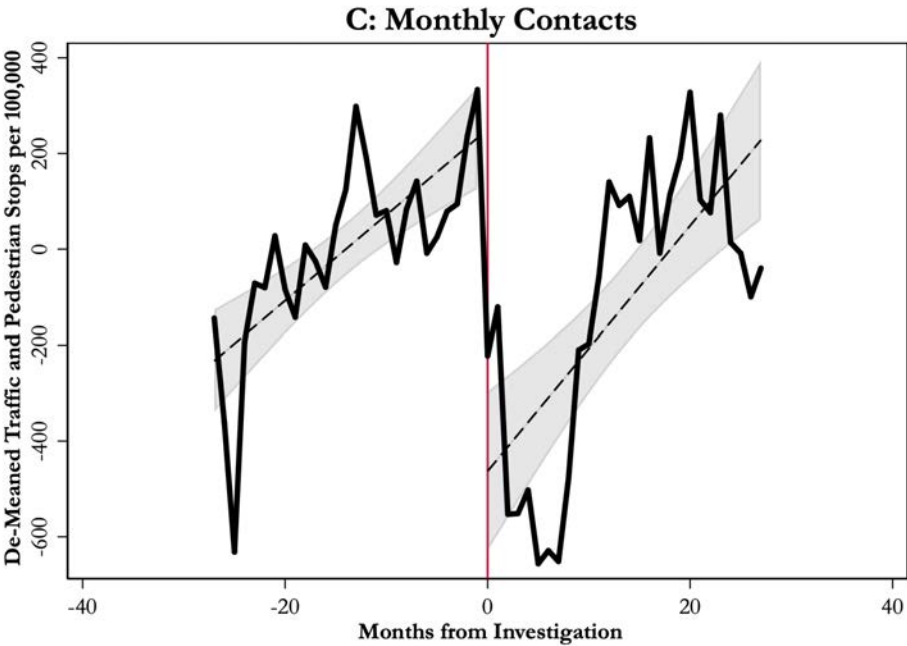
Figure 10: Police Activity, Investigation With Viral Incident of Deadly Force

St. Louis, MO
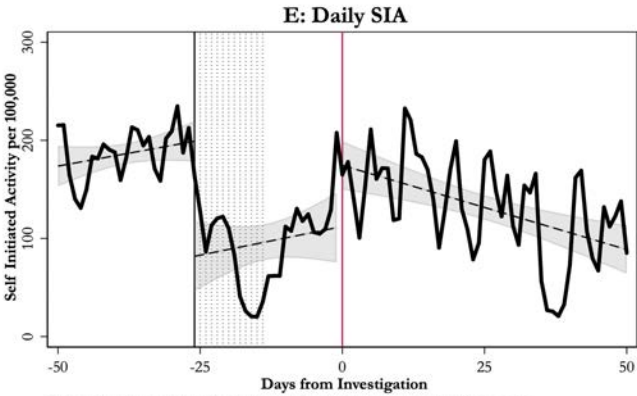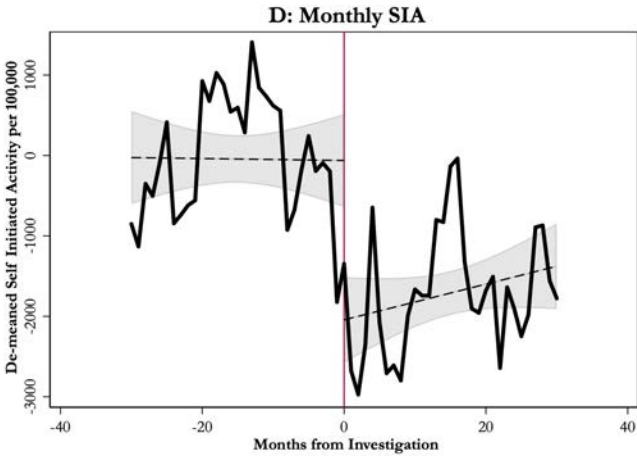
Michael Brown was shot by police on August 9, 2014 -- 26 days before the investigation began. Violent protests erupted that night, the Governor of Missouri declared a state of emergency on August 16th, and the National Guard was called in on the 18th. The National Guard was withdrawn on August 21st. Peaceful protests continued through early September, though several arrests were made.

Baltimore, MD
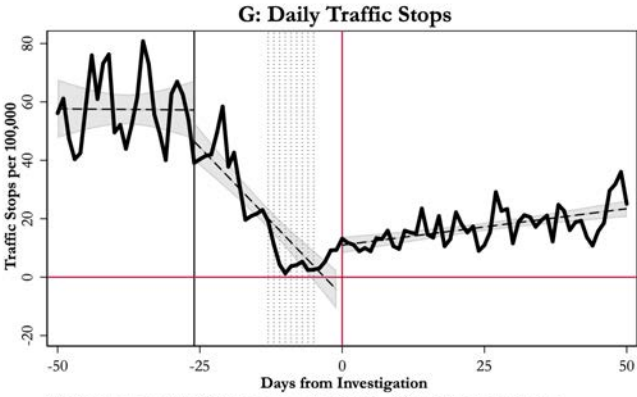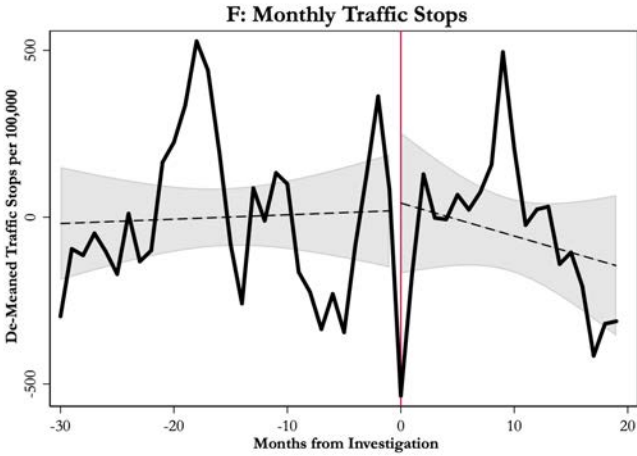
Freddie Gray was arrested and fell into a coma on April 12, 2015 -- 26 days before the investigation began. Gray died from his injuries on April 19th. Violent riots erupted on April 25th, the city of Baltimore declared a state of emergency on April 28, and a mandatory curfew remained in place until May 3rd (5 days before the investigation began).

Figure 11: The Correlation Between Police Activity and Crime

Panel A: Chicago, IL

Panel B: St. Louis, MO

Notes: For Panel A, the vertical axis plots the difference in monthly average outcome for 14 months pre investigation and monthly average outcome for 28 months pre investigation, for each of the 22 districts in Chicago, IL. The horizontal axis plots the difference in monthly average police civilian contacts for 14 months post investigation and monthly average police civilian contacts for 28 months pre investigation, for eac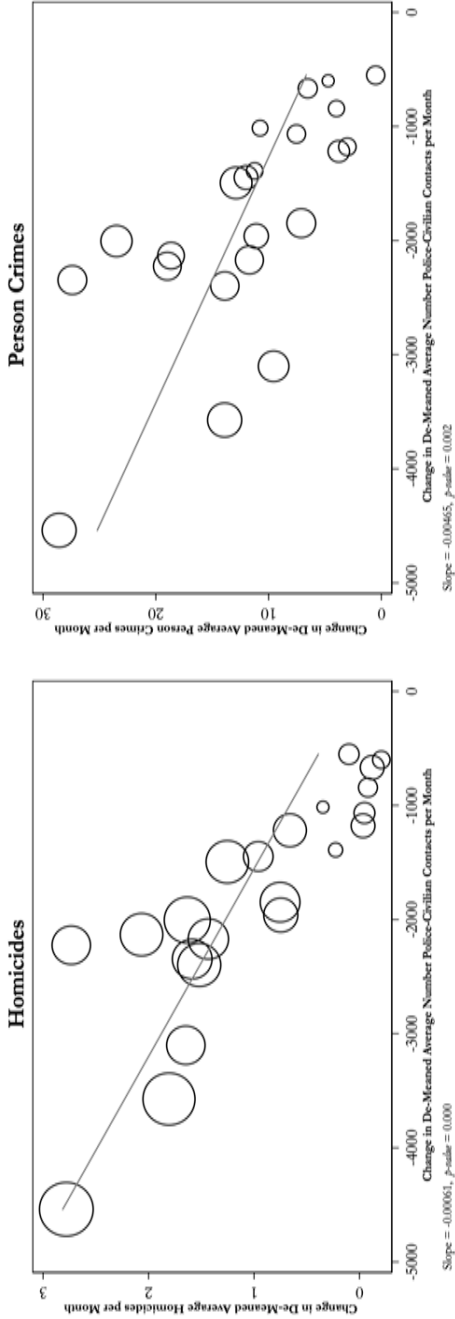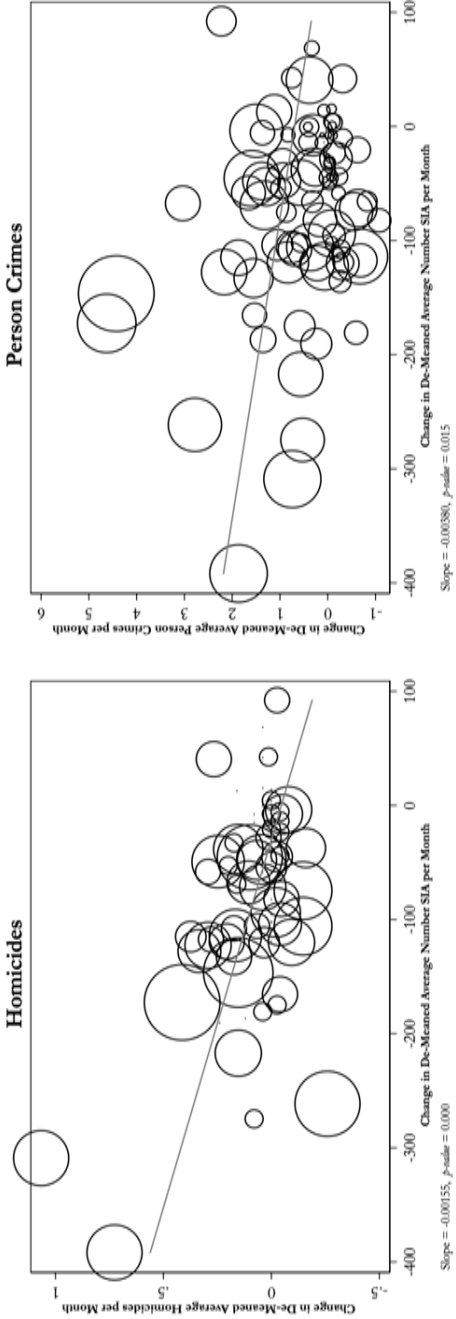h of the 22 districts in Chicago, IL. Outcomes and police civilian contacts are de-meaned by the average value in that calendar month in the pre-event period for each district to account for seasonality in criminal behavior. Each point in the graph is weighted by the average monthly outcome in the pre-event period for that district. For Panel B, the vertical axis plots the difference in monthly average outcome for 19 months post investigation and monthly average outcome for 28 months pre investigation, for each of the 89 neighborhoods in St. Louis, MO. The horizontal axis plots the difference in monthly average self-initiated activities for 19 months post investigation and monthly average self-initiated activities for 28 months pre investigation, for each of the 89 neighborhoods in St. Louis, MO. Outcomes and self-initiated activities are de-meaned by the average value in that calendar month in the pre-event period for each neighborhood to account for seasonality in criminal behavior. Each point in the graph is weighted by the average monthly outcome in the pre-event period for that neighborhood.

Table 1: Event-Study Estimates of the Effect of Investigations on Crime

	De-meaned Homicides per 100,000			De-meaned Total Crime per 100,000		
	(1)	(2)	(3)	(4)	(5)	6)
Panel A: All Investigations						
Constant	-0.000***	-0.001	-0.017	0.000	-0.125**	12.671***
	(0.000)	(0.001)	(0.059)	(0.000)	(0.052)	(4.060)
Investigated	0.000***	0.000	-0.021	-0.000	0.060	9.147
	(0.000)	(0.001)	(0.044)	(0.000)	(0.089)	(6.301)
Post-Event	-0.039	-0.036	-0.036	-19.862***	-23.433***	-23.403***
	(0.045)	(0.065)	(0.065)	(6.637)	(6.450)	(6.439)
Post-Event x Investigated	0.077	-0.077	-0.077	8.849	-15.707	-15.758
	(0.140)	(0.194)	(0.194)	(13.102)	(18.884)	(18.886)
R-squared	0.001	0.007	0.032	0.029	0.127	0.186
Panel B: Investigations Without Viral Incident of Deadly Force						
Constant	-0.000**	-0.001	0.083*	0.000	-0.135**	12.602***
	(0.000)	(0.001)	(0.048)	(0.000)	(0.059)	(4.483)
Investigated	0.000**	0.001	0.051	-0.000	0.078	10.965
	(0.000)	(0.001)	(0.034)	(0.000)	(0.104)	(6.609)
Post-Event	-0.061	-0.113**	-0.113**	-18.473**	-23.822***	-23.793***
	(0.051)	(0.051)	(0.051)	(7.139)	(7.353)	(7.338)
Post-Event x Investigated	-0.154*	-0.256**	-0.256**	-1.300	-30.029*	-30.085*
	(0.090)	(0.096)	(0.096)	(14.197)	(17.065)	(17.077)
R-squared	0.015	0.095	0.108	0.037	0.222	0.285
Panel C: Investigations With Viral Incident of Deadly Force						
Constant	0.000	0.000	-0.326**	-0.000	-0.092	4.765
	(0.000)	(0.000)	(0.109)	(0.000)	(0.108)	(24.347)
Investigated	-0.000	-0.000	-0.058	0.000	-0.009	7.524
	(0.000)	(0.000)	(0.070)	(0.000)	(0.149)	(10.455)
Post-Event	0.059	0.169**	0.170**	-25.973	-22.114	-22.038
	(0.082)	(0.057)	(0.056)	(17.919)	(13.998)	(14.017)
Post-Event x Investigated	1.094***	0.845***	0.844***	53.506*	47.714**	47.654**
	(0.334)	(0.210)	(0.211)	(28.471)	(16.344)	(16.364)
R-squared	0.292	0.346	0.389	0.167	0.206	0.239
Population Weights	N	Y	Y	N	Y	Y
Additional Controls	N	N	Y	N	N	Y

Notes: This table presents difference-in-difference results for our main event study. Standard errors, reported in parentheses, are clustered by city x investigation (54 clusters). The dependent variables are reported total crimes and homicides, respectively, per 100,000 people, which are de-meaned over the pre-investigation period in each city by calendar month to account for seasonal trends in criminality. Observations are investigation x month-from-investigation cells. The event window is from 28 months before the start of the investigation to 19 months after the investigation. The sample includes investigated cities and their corresponding synthetic cities, constructed using the procedure described in Abadie, Diamond, and Hainmueller (2010) using the companion *Synth* package for Stata. The synthetic cities are the omitted category in all specifications. In panel A, the sample includes all 27 investigations since 1994 with available data and their synthetic cities. In panel B, this is restricted to the 22 investigations that were not preceded by viral incidents of deadly force and in panel C, to the 5 investigations that were preceded by viral incidents of deadly force. In panel A, N = 2808, in panel B, N = 2288, and in panel C, N = 520 in all specifications. Columns (2), (3), (5), and (6) weight the estimates by the population of the city in the given year. Columns (3) and (6) include the following additional controls: education, economic, and race indices taken from the US Census in the most recent year before the investigation. For details on variable construction see the data apendix. Crime data comes from FBI UCR crime reports collated by ICPSR and from www.openicpsr.org.

Table 2: Event-Study Estimates of the Effect of Investigations on Crime Using Propensity Scores

	De-meaned Homicides per 100,000			De-meaned Total Crime per 100,000		
	(1)	(2)	(3)	(4)	(5)	6)

Panel A: All Investigations

Constant	0.000	0.000***	0.018	0.000	0.000	6.094**
	(0.000)	(0.000)	(0.062)	(0.000)	(0.000)	(2.628)
Investigated	0.000	-0.000	-0.023	-0.000	-0.000***	19.710**
	(0.000)	(0.000)	(0.067)	(0.000)	(0.000)	(8.916)
Post-Event	-0.106*	-0.058	-0.058	-27.108***	-24.062***	-24.062***
	(0.054)	(0.044)	(0.044)	(8.085)	(6.329)	(6.332)
Post-Event x Investigated	0.145	-0.057	-0.057	16.095	-15.535	-15.535
	(0.143)	(0.189)	(0.189)	(13.892)	(18.632)	(18.642)
R-squared	0.003	0.007	0.035	0.037	0.125	0.186

Panel B: Investigations Without Viral Incident of Deadly Force

Constant	0.000	0.000	0.069	0.000**	0.000	6.415
	(0.000)	(0.000)	(0.048)	(0.000)	(0.000)	(3.883)
Investigated	-0.000	-0.000	0.055	-0.000*	-0.000	22.567***
	(0.000)	(0.000)	(0.046)	(0.000)	(0.000)	(7.714)
Post-Event	-0.140**	-0.053	-0.053	-27.124***	-25.374***	-25.374***
	(0.058)	(0.052)	(0.052)	(9.532)	(7.988)	(7.993)
Post-Event x Investigated	-0.075	-0.321***	-0.321***	7.351	-29.143*	-29.143*
	(0.094)	(0.096)	(0.096)	(15.539)	(16.967)	(16.978)
R-squared	0.010	0.099	0.110	0.043	0.224	0.286

Panel C: Investigations With Viral Incident of Deadly Force

Constant	-0.000	-0.000	0.006	-0.000	-0.000	-5.890
	(0.000)	(0.000)	(0.099)	(0.000)	(0.000)	(13.340)
Investigated	0.000	0.000	-0.256	0.000	0.000	17.579
	(0.000)	(0.000)	(0.147)	(0.000)	(0.000)	(13.716)
Post-Event	0.044	-0.071	-0.071	-27.037*	-19.889**	-19.889**
	(0.134)	(0.076)	(0.076)	(13.033)	(7.699)	(7.722)
Post-Event x Investigated	1.109**	1.093***	1.093***	54.570*	45.886***	45.886***
	(0.351)	(0.219)	(0.220)	(25.679)	(11.386)	(11.420)
R-squared	0.180	0.352	0.392	0.124	0.178	0.243

Population Weights	N	Y	Y	N	Y	Y
Additional Controls	N	N	Y	N	N	Y

Notes: This table presents difference-in-difference results for our main event study. Standard errors, reported in parentheses, are clustered by city x investigation (54 clusters). The dependent variables are reported total crimes and homicides, respectively, per 100,000 people, which are de-meaned over the pre-investigation period in each city by calendar month to account for seasonal trends in criminality. Observations are investigation x month-from-investigation cells. The event window is from 27 months before the start of the investigation to 19 months after the investigation. The sample includes all investigated cities and their corresponding matched cities where matched cities are calculated using one-to-one exact matching with propensity scores. Propensity scores are calculated using percentage black, poverty rate and Percent population within 15 and 24 years of age. In case of non-convergence, propensity scores are calculated using percentage white, poverty rate and Percent population within 15 and 24 years of age. The matched cities are the omitted category in all specifications. In panel A, the sample includes all 27 investigations since 1994 with available data and their synthetic cities. In panel B, this is restricted to the 22 investigations that were not preceded by viral incidents of deadly force and in panel C, to the 5 investigations that were preceded by viral incidents of deadly force. In panel A, N = 2808, in panel B, N = 2288, and in panel C, N = 520 in all specifications. Columns (2), (3), (5), and (6) weight the estimates by the population of the city in the given year. Columns (3) and (6) include the following additional controls: education, economic, and race indices taken from the US Census in the most recent year before the investigation. For details on variable construction see the data apendix. Crime data from 1993-2017 comes from FBI UCR crime reports collated by ICPSR.