DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Essays on the Economic and Social Consequences of Policing

Citation
Tebes, Jonathan Kraemer. 2022. Essays on the Economic and Social Consequences of Policing. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37372130

Terms of Use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story .

Accessibility

HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Economics

have examined a dissertation entitled

"Essays on the Economic and Social Consequences of Policing"

presented by Jonathan Kraemer Tebes

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature ______________________

Typed name: Prof. Desmond Ang

Signature ______________________

Typed name: Prof. Edward Glaeser

Signature ______________________

Typed name: Prof. Andrei Shleifer

Signature Lawrence Katz

Typed name: Prof. Lawrence Katz

Date: April 25, 2022

Essays on the Economic and Social Consequences of Policing

A dissertation presented

by

Jonathan Kraemer Tebes

to

The Department of Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Economics

Harvard University

Cambridge, Massachusetts

April 2022

© 2022 Jonathan Kraemer Tebes

All rights reserved.

Dissertation Advisor:
Professor Lawrence Katz

Author:
Jonathan Kraemer Tebes

Essays on the Economic and Social Consequences of Policing

Abstract

This dissertation contains three essays on the economic and social consequences of policing. Chapters 1 and 2 examine the effectiveness and equity of the police's use of pedestrian stops as a crime deterrence tool. Over 3.5 million pedestrians are stopped by police in the United States every year. Using administrative data from New York City, Chapter 1 estimates the impact of pedestrian stops on neighborhood crime and high school dropout rates of neighborhood residents. Exploiting a 2012 reform that reduced stops by 95%, we compare neighborhoods that have similar crime rates but substantially different stop rates prior to the reform. Treated neighborhoods that experienced twice the reduction in stop rates do not display differential increases in felonies and violent misdemeanors, shootings, or killings over the five years following the reform. Comparing students across schools that are differentially exposed to changes in stop rates, we estimate that the reform reduced the probability of high school dropout by $0.36 - 1.66$ percentage points per academic year.

Chapter 2 explores whether racial disparities in stop rates reflect true racial differences in criminal behavior or are the result of unfair targeting by patrol officers. By instrumenting for neighborhood stop rates with the reform, we trace out the marginal return curve of stops by race and find that Black and Hispanic residents were stopped at substantially higher rates than would be optimal for crime detection.

Chapter 3 explores spillovers of fatal police shootings on civic engagement. Roughly a thousand people are killed by American law enforcement officers each year, accounting for more than 5% of all homicides. We estimate the causal impact of these events on voter registration and voting behavior. Exploiting hyper-local variation in how close residents live to a killing, we find that exposure to police violence leads to significant increases in registrations and votes. These effects are driven entirely by Black and Hispanic citizens and are largest for killings of unarmed individuals. We find corresponding increases in support for criminal justice reforms, suggesting that police violence may cause voters to politically mobilize against perceived injustice.

Contents

List of Tables

List of Figures

Acknowledgments

I am grateful to my dissertation committee for their unrelenting advice, encouragement, and support. Larry Katz provided wise comments and guidance throughout graduate school, always ready with a helpful reference or to provide thoughtful yet expeditious feedback. In addition to his immense warmth, Ed Glaeser provided a wonderful mix of price theory and historical knowledge about policing institutions. Throughout graduate school, Andrei Shleifer provided unrelenting support of my ideas, always encouraging me to stay the course with my research agenda. This dissertation would not have come to fruition without his guidance and friendship. Desmond Ang provided much appreciated advice on navigating graduate school, the job market, and beyond. I am extremely grateful to have witnessed and learned firsthand from the creativity and precision with which he studies important social problems.

Joscha Legewie was quick to take me under his wing, provide insight on New York City policing and educational institutions as well as data access. Without his support, encouragement, and friendship, this dissertation would not be possible. I am also indebted to my undergraduate professors – Jon Gruber and Esther Duflo – whose enthusiasm for teaching and mentorship inspired me to pursue a PhD in economics, and to Noam Angrist, who not only illustrated how economic research could generate positive social impact, but also mentored and encouraged a naive college junior to pursue a doctorate in economics. My time as an research assistant to Amy Finkelstein confirmed my interest in pursuing a graduate degree. I am immensely grateful for her support and mentorship, and am lucky to have witnessed her team's tour-de-force research process firsthand. I am also thankful for my classmates – Francesca Bastianello, Lydia Cox, Natalia Emanuel, Paul Fontanier, Benny Goldman, Thomas Graeber, Emma Harrington, Samuel Lite, Peter Maxted, Armando Miano, William Murdock III, Hannah Shaffer – who have become an extended family.

Finally, I am grateful to my family for their love and support, while in graduate school and always. My parents are my role models, as they each have forged a life full of purpose, committed to community-building and service to others. I am blessed with a wonderful brother Dan, sister-in-law Yui, and little sister-in-law Huong, who provided much-needed friendship and levity while writing my job market paper during a pandemic. And lastly, special thanks to my partner Caroline, who stood by me through the ups and downs of a PhD, always being my biggest champion and my home – “you are whatever a moon has always meant and whatever a sun will always sing is you".

To my parents – your unconditional love and unyielding support of my curiosity from a young age made an academic life possible.

To my youth soccer teammates – your friendships cut across societal divides and taught me the immense importance of a supportive community for young people.

To the countless children who have suffered at the hands of failed public policy and were systematically excluded from equal opportunity.

Introduction

This dissertation contains three essays on the economic and social consequences of policing. A common theme is quantifying the indirect social consequences of policing, such as serious crime responses to stop and frisk policies, the spillover of stop and frisk policies on high school graduation rates, or the civic responses of civilians who witness fatal police shootings.

Chapters 1 and 2, which are joint work with Jeffrey Fagan, examine the effectiveness and equity of the police's use of pedestrian stops as a crime deterrence tool. Over 3.5 million pedestrians are stopped by police in the United States every year. Using administrative data from New York City, Chapter 1 estimates the impact of pedestrian stops on neighborhood crime and high school dropout rates of neighborhood residents. Exploiting a 2012 reform that reduced stops by 95%, we compare neighborhoods that have similar crime rates but substantially different stop rates prior to the reform. Treated neighborhoods that experienced twice the reduction in stop rates do not display differential increases in felonies and violent misdemeanors, shootings, or killings over the five years following the reform. Analysis of police surges reveals that when increases in stops are accompanied by increases in police officers, serious crime significantly declines. But alone, heightened stop rates have no measurable impact on serious crime. Comparing students across schools that are differentially exposed to changes in stop rates, we estimate that the reform reduced the probability of high school dropout by $0.36 - 1.66$ percentage points per academic year, carrying an annual social benefit of over $200 million.

Chapter 2 explores whether racial disparities in stop rates reflect true racial differences in criminal behavior or are the result of unfair targeting by patrol officers. To answer this question, we compare the *accuracy* of stops that were deterred by the 2012 reform across racial groups as a test for racial bias. Our analysis is motivated by a model of officer stopping decisions where statistically-rational officers face diminishing marginal returns to stops and equate marginal returns across racial groups so as to maximize overall returns to stops. Racial bias is then identified by racial differences in outcomes of marginal stops (Becker, 1957). This approach allows us to explore whether closing racial disparities in stop rates poses an equity-efficiency trade-off for patrol officers (Feigenberg and Miller, 2021). By instrumenting for neighborhood stop rates with the timing of the reform, we mitigate common concerns of omitted variables bias and infra-marginality bias and are able to trace out the marginal return curve to stops for each race. Pairwise comparisons of marginally-deterred stops in the first year of the reform suggest that both Black and Hispanic pedestrians were over-stopped relative to white pedestrians prior to the reform. Furthermore, by finding the point at which marginal returns are equated across races, we estimate that the police continued to over-stop Black and Hispanic pedestrians by a combined *298,000 stops per year* in the first year of the reform, after aggregate stop rates had already fallen by 35%. Since our approach allows officers to statistically discriminate or "rationally" racial profile, we interpret these results as a *conservative estimate* of racial discrimination (Arnold and Hull, 2020).

In Chapter 3, which is joint work with Desmond Ang, we examine spillovers of fatal police shootings on civic engagement. Roughly a thousand people are killed by American law enforcement officers each year, accounting for more than 5% of all homicides. We estimate the causal impact of these events on voter registration and voting behavior. Exploiting hyper-local variation in how close residents live to a killing, we find that exposure to police violence leads to significant increases in registrations and votes. These effects are driven entirely by Black and Hispanic citizens and are largest for killings of unarmed individuals. We find corresponding increases in support for criminal justice reforms, suggesting that police violence may cause voters to politically mobilize against perceived injustice.

Chapter 1

The Effects of Police Stops on Crime and High School Dropouts[1]

1.1 Introduction

In the aftermath of George Floyd's murder, 20 million Americans protested police brutality and systemic racism (Buchanan *et al.*, 2020a). Many of them called for an overhaul of existing police institutions. Yet city-wide increases in officers per capita have consistently been found to causally reduce serious crime, leading some scholars to even claim U.S. cities are currently under-policed (Evans and Owens, 2007; Chalfin and McCrary, 2018; Mello, 2019). Much less is known about the effectiveness and equity of pedestrian stops, although over 3.5 million pedestrians are stopped by police in the United States every year (Harrell and Davis, 2020). While the primary goal of pedestrian stops is to protect civilians from imminent harm, police departments often concentrate stops in higher-crime neighborhoods

[1]Co-authored with Jeffrey Fagan. Dr. Fagan was instrumental in securing and understanding NYPD policing data. We also thank Edward Glaeser, Lawrence Katz, Joscha Legewie, and Andrei Shleifer for unrelenting advice and encouragement on this project. We are grateful for helpful comments from Desmond Ang, Noam Angrist, Caroline Chin, Natalia Emanuel, Paul Fontanier, Benny Goldman, Dev Patel, Robert Fluegge, and seminar participants at Harvard Public/Labor lunches and seminars. The NYU Research Alliance of NYC Schools provided the education data for this project. Tebes also benefited from generous financial support from the NSF Graduate Research Fellowship, the Stone Scholar Fellowship for Social Policy and Inequality Research, and the Horowitz Foundation. All errors are our own.

with the aim of deterring serious crime.

The impact of pedestrian stops on neighborhood crime and community well-being is theoretically ambiguous. Increased stop rates could raise the probability of apprehension and thus reduce the expected payoff of criminal activity (Becker, 1968). Alternatively, stops may impose substantial costs on local communities, as stops can be traumatizing (Geller *et al.*, 2014; Boyd, 2018), breed institutional distrust (Kirk and Papachristos, 2011a), and disrupt educational investments (Legewie and Fagan, 2019a; Bacher-Hicks and de la Campa, 2020). Stops also mechanically increase the likelihood of more serious downstream police actions, such as uses of force or arrests for minor offenses (Knox and Mummolo, 2020).

This paper explores the effectiveness of using pedestrian stops as a crime deterrence tool. To this end, we analyze the impact of a 2012 federal lawsuit – *Floyd, et al. v. The City of New York, et al.* – that ruled NYPD's stopping practices were unconstitutional and led to a permanent, 95% reduction in city-wide stop rates. We assess *effectiveness* by estimating the impact of the reform on two key outcomes: neighborhood crime and high school dropout rates. We complement our crime analysis with an analysis of police surges that increased both officers and stops in higher-crime areas prior to the reform, allowing us to disentangle the effect of stopping protocols from patrol officer presence.

We begin the crime analysis by first showing that city-wide crime trends appear unaffected by the the reform. Aggregate crime responses, however, may mask heterogeneous neighborhood responses, especially since the magnitude of stop reductions varied greatly across neighborhoods. To estimate the causal effect of stops on crime, we identify neighborhoods that, prior to the reform, had similar crime rates but substantially different stop rates. Specifically, we split neighborhoods into treatment and control groups based on mean stop rate residuals from a regression of stop rates on neighborhood crime measures during an earlier training period. Using a flexible difference-in-differences framework, we explore how crime rates differentially evolve in treatment neighborhoods relative to control neighborhoods over the five years following the reform. For the four years prior to the reform, treatment and control neighborhoods display parallel trends for multiple crime and policing outcomes.

Difference-in-differences estimates reveal that stop rates, uses of force, and frisks all fall by twice as much in treatment neighborhoods relative to control neighborhoods. Treatment neighborhoods, however, do not display differential increases in felonies or violent misdemeanors, major felonies, shootings, or killings over the five years following the reform. The confidence interval on felonies and violent misdemeanors rules out an increase of 1.5% of the pre-period mean. To put this null effect in context, New York City experienced a 38% decline in felony crime from 2000 to 2010. Counter to the crime deterrence hypothesis, we find that treatment neighborhoods exhibit *reductions* in non-violent misdemeanors and violations, which can be entirely explained by reductions in stop-related arrests.[2]

Given the lack of crime responses to the reform, we investigate whether police patrols, more broadly, deter crime. To do so, we analyze the impact of neighborhood police surges prior to the reform. About every six months, the police commissioner selected "Impact Zones" to receive additional patrol officers and resources. Using a difference-in-differences framework that controls for linear pre-trend differences, we find that Impact Zone assignment increases both stop rates and the number of officers conducting stops by about 35%. By year three of assignment, we estimate that felonies and violent misdemeanors are reduced by 8% ($p = 0.011$) and major felonies are reduced by 11% ($p = 0.015$). These results are robust to allowing for reasonably large parallel trend violations a la Roth and Rambachan (2021), and to a broader geographical definition of Impact Zone exposure that takes into account potential crime spillovers on nearby neighborhoods. Given that officer assignments were relatively unaffected by the reform, our combined findings imply patrol officer presence but not stop rates matter for deterring serious crime.

Employing a similar differences-in-differences strategy, we next use the reform to estimate spillovers of neighborhood stop exposure on high school dropout rates.[3] We proxy

[2]Our results are robust to alternative treatment definitions, controls, and parallel trends violations a la Roth and Rambachan (2021).

[3]We focus on high school students since they are 21 times more likely to be stopped by police than middle school students and face critical educational investment decisions.

for stop exposure in a student's peer network by exploiting school-level variation in stop exposure. The advantage of this approach is that it is able to detect effects in the presence of within-school peer effects and accounts for the fact that high school students likely spend time in areas outside of their home Census tract. Specifically, we split schools into quartiles based on the number of stops per square mile observed in students' home Census tracts during the three school years prior to our sample period. We then estimate differential changes in yearly enrollment outcomes for students who attend fourth and third quartile schools relative to students attending "control" schools that rank in the bottom half of training-period stop exposure.

As a result of the reform, students attending fourth (third) quartile schools experience an annual decline in stops in their home neighborhoods that is more than three (two) times larger than that same decline for students in control schools. We estimate effects on two outcomes. The first outcome is an indicator for whether a student was discharged by a non-DOE agency directive ("institutional directive"), such as from the Department of Corrections. This provides a lower-bound estimate of criminal justice related school dismissals. The second outcome is an indicator for whether or not a student drops out of high school or is discharged by institutional directive. This outcome measures the holistic impact of the reform on the likelihood a student leaves high school.

We find that students in fourth-quartile schools display a 0.14 percentage point or 54% reduction in likelihood of being discharged by an institutional directive relative to students attending control schools ($p = 0.004$). That is, 84 students per year were *not* discharged by institutional directive following the *Floyd* decision. Depending on the specification, we estimate that the likelihood that a fourth-quartile student dropped out or was discharged by institutional directive fell by 0.36-1.66 percentage points per school year ($p < 0.02$), preventing about 660 students per year from leaving high school prematurely. IV estimates indicate that increasing the average stop rate in students' home neighborhoods by 100 stops per year across an entire school carries an annual social cost of $23 million. Moreover, our results likely underestimate true effect sizes since impacts on students in control schools are differenced out.[4] Both in terms of student-level effect sizes and the number of students affected, effect magnitudes are substantially larger than previous estimates for middle school students as well as estimated impacts of neighborhood exposure to police killings (Ang, 2021a; Bacher-Hicks and de la Campa, 2020). Heterogeneity analyses reveal that Black and Hispanic male students are the most likely to benefit from the reform.

This chapter makes three main contributions. First, we contribute to an understanding of crime deterrence.[5] Our analysis of the reform provides new causal evidence on the effectiveness of pedestrian stops as a crime deterrence tool. The precision of our null result suggests that, relative to alternative policing activities, pedestrian stops have very little impact on serious crime. These findings reinforce the notion that proactive and disorder policing tactics, which aim to deter crime by strictly enforcing low-level offenses and maintaining order in higher-crime areas, have only limited impacts on serious crime (Braga and Bond., 2008; Caetano and Maheshri, 2014; Chalfin and McCrary, 2018). On the other hand, analysis of Impact Zones accords with a large literature that finds hot-spot policing (Braga and Bond., 2008; Weisburd *et al.*, 2006; MacDonald *et al.*, 2016), officer deployments (Di Tella and Schargrodsky, 2004), and police employment (Chalfin and McCrary, 2018; Mello, 2019) to effectively deter neighborhood crime.

Second, this chapter adds to a growing body of evidence that finds policing to impose negative externalities on local communities of color (Mello, 2021; Ang, 2021a; Legewie and Fagan, 2019a; Bacher-Hicks and de la Campa, 2020). Over 90% of the 660 students who were prevented from dropping out of high school by the *Floyd* decision were Black or Hispanic. Previous work finds that exposure to aggressive policing during middle school can harm contemporaneous test scores, high school graduation rates, and college-going rates of Black boys (Legewie and Fagan, 2019a; Bacher-Hicks and de la Campa, 2020). This chapter studies stop exposure during high school – a time when stops are 21 times more likely than during

[4]We confirm that our results are not driven by other educational reforms over this period, such as changes to suspension policies.

[5]See Chalfin and McCrary (2017) for a thoughtful summary of criminology and economic literature on this subject.

middle school. By leveraging sharp changes in city-wide stop rates, we are better able to gauge the full magnitude of treatment effects, and have the advantage of studying a reform that has the potential to be more broadly adopted.

Lastly, in tandem with the second chapter, this chapter sheds light on the mechanisms underlying the impact of neighborhoods on social mobility (Ludwig *et al.*, 2013; Chetty and Katz, 2016; Chetty *et al.*, 2020). Chetty *et al.* (2020) show that, conditional on parental income, Black boys have lower incomes in adulthood in 99% of Census tracts. Chapter 2 of this dissertation reveals that Black male teenagers are substantially over-stopped by police, while this chapter shows that these interactions disproportionately translate into high school dropouts and interactions with the carceral state. Combined with previous studies showing the indirect harm of police killings (Ang, 2021a), ticketing (Mello, 2019; Goncalves and Mello, 2021), and prosecution of low-level arrests (Agan *et al.*, 2021), it becomes clear that the police play a pivotal role in the life trajectory of young Black men. This chapter thus provides actionable policy advice about one reform that could help close place-based racial opportunity gaps.

The remainder of this chapter proceeds as follows. Section 1.2 describes the institutional details and provides descriptive statistics. Section 1.3 describes data sources and outcomes. Section 1.4 estimates crime deterrence effects. Section 1.5 estimates impacts on high school dropout rates and Section 1.6 concludes.

1.2 Setting

1.2.1 Pedestrian stops

Officers are afforded considerable discretion over whether or not to conduct a stop. They may stop a pedestrian if there is a *reasonable suspicion* that he has committed or is about to commit a crime (*Terry v. Ohio*, 1968). That is, the officer must be able to articulate specific facts that, together, would lead a reasonable person to suspect criminal activity. These include fitting the description of a suspect near the location where a crime was reported, walking in a way that evades police contact, carrying items that have been reported stolen recently, or wearing bulky clothes in warm weather. If the officer believes the pedestrian to be armed and dangerous, they may “frisk" or pat-down the pedestrian's outer clothing to search for a weapon. During the course of a stop, if the officer observes evidence of criminal behavior, such as unlawful possession of drugs or a weapon, “probable cause" is established and the officer may arrest the pedestrian.

Investigative stops primarily serve two policing functions. First, they enable officers to protect civilians from imminent harm by investigating suspicious activity. Stops can be particularly useful for confiscating illegal firearms before they are used in violent crimes. Second, the concentration of stops in higher-crime areas may deter future crime by increasing the probability an offender is apprehended (Becker, 1968). Mayor Bloomberg often justified racial disparities in stop rates with this crime deterrence rationale:

> *They have argued that police stops are discriminatory because they do not reflect the city's overall census numbers. By that flawed logic, our police officers would stop women as often as men and senior citizens as often as young people... The absurd result of such a strategy would be far more crimes committed against black and Latino New Yorkers. When it comes to policing, political correctness is deadly.* (Bloomberg, 2013)

The use of pedestrian stops increased dramatically in the 1990s, as urban police departments transitioned from “reactive" to “proactive" policing strategies. The NYPD was at the forefront of this transition, adopting Compstat in 1995, which was the first crime tracking and management system in the U.S. CompStat enabled the police to allocate officers to neighborhoods known for crime during peak hours of criminal activity. Commanders attended monthly CompStat meetings designed to hold them accountable to reducing major felony crimes in their designated area. As an advocate of “broken windows" policing theory – the idea that visible signs of crime and civil disorder encourage more serious crimes – Commissioner Bratton instructed commanders and their line officers to conduct a high volume of stops and strictly enforce minor offenses, such as fare evasion, public drinking, and graffiti. During the 1990s, homicide rates fell in New York City by 74%, the largest decline of any large city over this period (Levitt, 2004). Interpreting this decline as evidence of efficacy, other departments quickly created their own data-driven management systems and adopted similar proactive policing strategies.[6] Data-driven, proactive policing would become a prominent fixture of urban policing strategies for the next two decades (National Academies of Sciences Engineering and Medicine, 2018).

From 2006 to 2011, the NYPD conducted over 545,000 stops per year, a rate that is 10,000 times greater than the annual number of police-involved shooting incidents.[7] As shown in Figure 1.1, stops were concentrated in higher-crime neighborhoods that were disproportionately Black and Hispanic – 84% of stops were of Black or Hispanic pedestrians compared to a residential population that is just 51% Black or Hispanic. Much of this racial disparity can be explained by where stops were conducted. Assigning the residential racial composition of a Census tract to stops conducted in that tract predicts that 69% of stops would be Black or Hispanic, suggesting that over half of observed racial differences in stop rates can be attributed to the concentration of stops in predominantly Black and Hispanic neighborhoods.

Pedestrians stopped by police were also disproportionately male (93%) and young – 54% were under the age of 25, and a third were under the age of 20.[8] Using residential population estimates from the 2010 Census, we estimate that in higher-crime neighborhoods, the police conducted 7.5 stops per Black male resident between the ages of 14 and 18 per year.[9] For Hispanic males, the equivalent stop rate was 2.8 stops per resident, while it was only 0.4 stops for male teenagers of other races. This means that during high school, the average Black (Hispanic) male student living in a higher-crime neighborhood was stopped by police 30.0 (11.2) times. Most of these stops did not result in the detection of criminal

[6] Five years after its creation, one-third of the nation's 515 largest police departments intended to deploy a CompStat-like program by 2001 (Eterno and Silverman, 2019).

[7] We divide the number of stop incidents by the number of officer-involved shooting incidents recorded by the NYPD from 2006-2011.

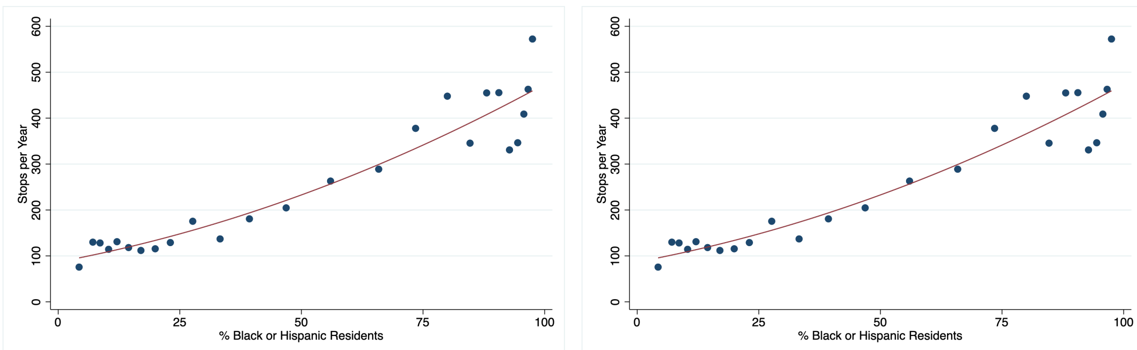
[8] To see the full distribution of stops by pedestrian age, see Figure A.2.

[9] Higher-crime neighborhoods are defined as tracts that rank in the top 25% of felony crimes per square mile during the training period. We estimate race-specific 14-18 year-old male populations for each Census tract by computing a weighted average of 10-14 and 15-19 year old bins and multiplying these estimates by the estimated racial composition of the entire Census tract.

activity. Of the 14% that did detect pedestrian wrong-doing, 5.8% resulted in an arrest, 0.9% led to the discovery of a weapon, 1.7% led to the discovery of drugs, and 6.4% led to a court summons being issued for a minor violation. Only 1 in 770 stops led to the discovery of an illegal firearm.

Figure 1.1: *Neighborhood Stop Rates by Felony Crime Rates & Racial Composition in NYC*

(a) *Neighborhood Felony Crime Rate* **(b)** *Neighborhood Racial Composition*



Notes: This figure displays the number of stops per year by neighborhood characteristics. The x-axis of Panel A refers to the percentile rank of each tract in terms of felony crimes per per 1 mile^2 recorded between Jan 1, 2006 and Apr 15, 2008. The x-axis of Panel B refers to the fraction of residents who identify as Black or Hispanic in the 2010 Census. The sample includes 2,096 of 2165 Census tracts, dropping tracts with less than 500 residents listed in the 2010 Census. Sample period spans Jan 1, 2006 to Apr 15, 2012.

Surprisingly, stop rates recorded during the height of stop and frisk are similar to stop rates in *other large urban police departments over the past five years*. Table 1.1 compares stops and stop outcomes across New York City, Chicago, and Philadelphia – the only departments of the ten largest police departments to publicly release information on all pedestrian stops.[10] Annual stop rates in Philadelphia and Chicago were, respectively, 135% and 90% the level conducted by the NYPD from 2006-2011. After the reform, stop rates in New York City fell to just a fraction of what stop rates are currently in Philadelphia and Chicago. The reform we study therefore is likely relevant for many other large U.S. cities.

[10]Most police departments do not release information on pedestrian stops. Of the ten largest police departments in the country, only four release any data on investigative stops. Only three provide information on all stops, including stops that do not result in an arrest or citation. All three departments released these data publicly in order to comply with a court mandate.

Table 1.1: *Pedestrian Stop Rates in New York City, Chicago, & Philadelphia*

	NYPD		Chicago PD	Philadelphia PD
	Pre-Floyd	Post-Floyd		
	(1)	(2)	(3)	(4)
Stops	546,745	19,001	161,755	141,506
Stops per 100k	6,625	223	5,959	9,008
Stops per Violent Indexed Crime	11.07	0.39	5.74	9.38
Stops per All Indexed Crime	2.81	0.11	1.41	2.22
Stops per Full-time Officers	15.55	0.54	12.78	21.96
Stops: % Black or Hispanic	83.56	85.35	90.22	78.79
$\frac{\text{\% of Stops Black or Hispanic}}{\text{\% of Population Black or Hispanic}}$	1.63	1.68	1.56	1.42
Stops: % with Arrest Made	5.79	24.05	13.13	11.43
Stops: % with Weapon Discovery	0.95	6.69	2.04	-
Stops: % with Gun Discovery	0.13	1.91	1.38	-
Stops: % with Drug Discovery	1.73	6.70	6.13	3.40
Stops: % with Court Summons Issued	6.44	2.80	7.97	-
Years	2006 - 2011	2014 - 2019	2016 - 2019	2014 - 2019

Notes: This table provides descriptive statistics on stops conducted by the New York City Police Department (NYPD), the Chicago Police Department (CPD), and the Philadelphia Police Department (PPD). Data for the CPD and PPD come from administrative records that are publicly available on each department's website. Crime data come from the FBI's Uniform Crime Reporting (UCR) program. All index crime refers to the sum of violent and property crime as reported in UCR data. The fraction of a city's population that is Black or Hispanic is measured using the 2010 Census for Column 1 and the 2019 ACS 5-year estimates for Columns 2 through 4.

1.2.2 *Floyd et al. v. The City of New York*

This paper primarily explores the consequences of a federal lawsuit – *Floyd, et al. vs. City of New York, et al.* (hereafter referred to as "*Floyd*") – that alleged the NYPD implemented and sanctioned "a policy, practice, and/or custom of unconstitutional stops and frisks of City residents" in violation of the 4th and 14th Amendments of the U.S. Constitution. This was one of the first rulings to severely limit the use of investigative pedestrian stops. Since this case, the ACLU and the Department of Justice have taken legal action to implement similar reforms in Chicago, Ferguson, and Philadelphia (ACLU of Illinois, 2015; Dept. of Justice, 2016; ACLU of Pennsylvania, 2011, 2021).

We display a timeline of important moments in the case and changes to the NYPD's stop and frisk policies in Figure A.1. Even though the case was first filed in January 2008, the NYPD continued to increase annual stop rates through the beginning of 2012. In fact, an internal memo issued by Commissioner Kelly in October 2011 recommended managers set performance goals for the number of monthly pedestrian stops conducted by patrol officers. On April 16, 2012, the Court allowed expert testimony from Dr. Fagan, which indicated that the vast majority of stops did not result in detection of any wrong-doing and that Black and Hispanic residents were stopped at disproportionate rates. A month later, the court granted "Class Action" status, allowing potentially thousands of stopped pedestrians to join the lawsuit. That same day, Commissioner Kelly issued a memo that formalized efforts to scale back stop rates, which had already begun to decline following Dr. Fagan's testimony. Around this time, Mayor Bloomberg released a brief statement supporting the reform, stating: "we need to mend, not end, the practice, and the reforms Commissioner Kelly has put into place ensure the focus is quality, not quantity." (Goldstein and Ruderman, 2012) [11] After a nine-week trial, on August 13, 2013, the court found the NYPD liable for a pattern and practice of racial profiling and unconstitutional stops, and ordered broad reforms that would be supervised through a court monitor.

The monitor implemented three key reforms. First, the police academy added training to help cadets distinguish between lawful "criminal profiling" and illegal "racial profiling". Second, the police revised its patrol guide to more clearly and narrowly define when officers have lawful authority to conduct a stop under *reasonable suspicion*. Third, the monitor worked with the NYPD to improve data collection and management processes to ensure officers were keeping an honest record of all stops (Zimroth, 2016).

Figure 1.2 plots city-wide trends in stops, major felonies, non-major felonies and violent misdemeanors, and non-violent misdemeanors and violations throughout the study period. The first dashed line is set to the week of April 16, 2012, since this is when the NYPD began reducing stop rates in response to the case. The second dashed line marks the court's decision. Panel A illustrates that weekly stops fell by 95% over this period. Panel B, C, and D illustrate that reported crime rates remained rather flat for the first few years following

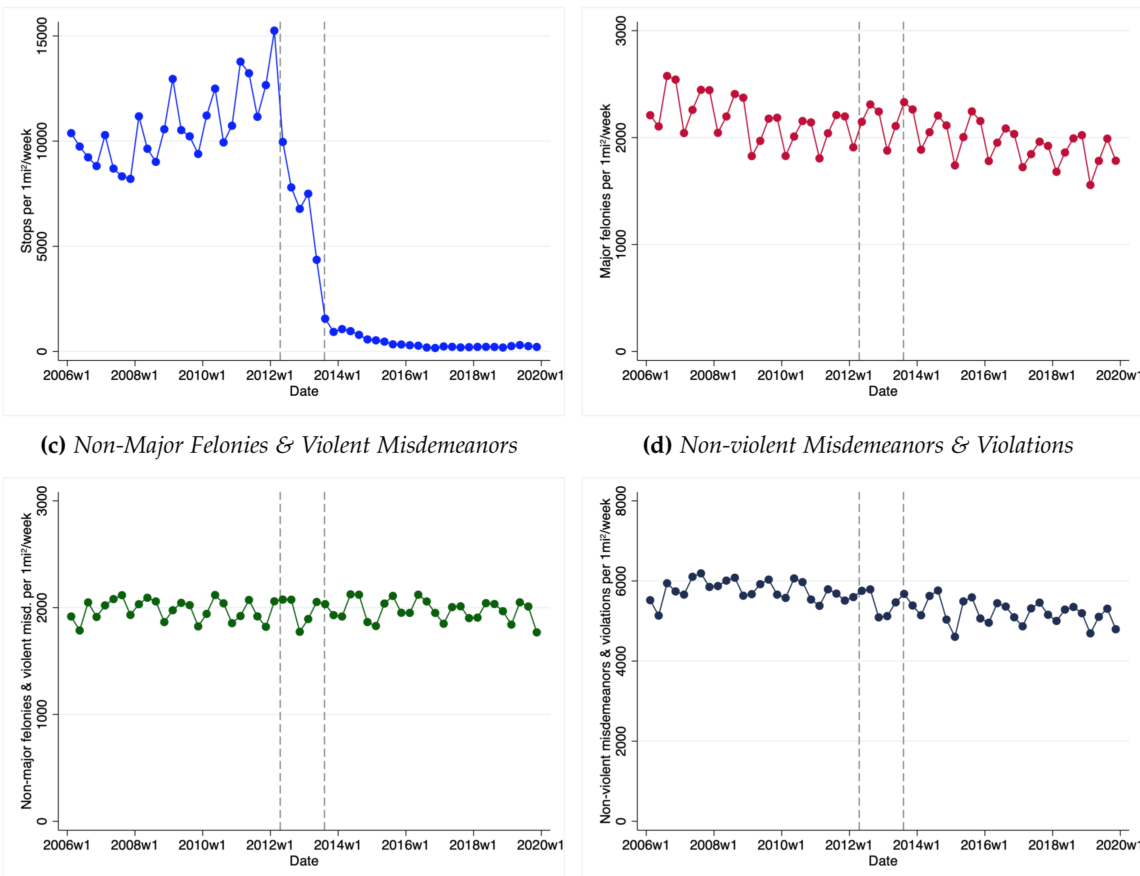
[11] In January 2013, the same court ruled in a separate lawsuit – *Ligon et al. v. The City of New York* – that the NYPD must immediately cease its practice of unlawful trespass stops outside public housing buildings in the Bronx, further reducing stop rates.

the reform before declining slightly. Importantly, there is no evidence that aggregate crime rates rose in response to *Floyd*.

Figure 1.2: *City-wide Trends during Study Period*

(a) *Stops* **(b)** *Major Felonies*

(c) *Non-Major Felonies & Violent Misdemeanors* **(d)** *Non-violent Misdemeanors & Violations*



Notes: This figure graphs city-wide means in stops, reported major felonies, non-major felonies and violent misdemeanors, and non-violent misdemeanors and violations by quarter from 2006 through 2019. All outcomes are measured as weekly rates per 1 mile^2. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling.

There is little evidence that changes to other aspects of policing coincided with the reform. The assignment of additional officers to high-crime areas under Operation Impact (e.g. see below) continued through July 2015, more than three years after the NYPD started to reduce stop rates. Similarly, the number of officers employed by the NYPD and city-wide clearance rates remained relatively flat for over two years after the reform (see Appendix

Figure B.2 and B.1). Finally, we do not observe a break in trend for total pay or overtime pay following the reform (see Appendix Figure B.3). Given this context, we interpret the reform as a shock to stopping protocol that held officer assignments relatively fixed.

With the reduction of stops, patrol officers were able to shift their attention to other patrol activities, such as pursuing reports of more serious crime or building community ties. Broad adoption of community policing tactics, however, would not be adopted for over three years after Dr. Fagan's testimony.[12]

1.2.3 Operation Impact

Operation Impact began in 2003 as a "natural outgrowth" of CompStat (Golden and Almo, 2004). About every six months, precinct and borough commanders nominated crime hotspots that the police commissioner would then select to be "Impact Zones" to receive additional officers and resources. In the first phase of Operation Impact, about 1,500 officers or two-thirds of all new academy graduates were deployed to Impact Zones. There were four general policing strategies employed in Impact Zones. The first was to increase foot patrols during times when crimes are most commonly committed. This included "vertical patrols", where officers would sweep the halls of public housing units, top-to-bottom. Second, officers were instructed to issue citations and court summonses for low-level "quality-of-life" crimes, such as loitering or drinking alcohol in public. Third, officers were encouraged to stop and question anyone they suspected of committing crimes in the area. Finally, the department allocated more specialized units, such as the Firearms Investigation Unit, to address the particular crime problems of the Impact Zone. (Golden and Almo, 2004)

We obtained maps of Impact Zone borders from the NYPD for thirteen phases that span January 2006 through July 2012. Overlaying Census tracts with these maps, we generate a panel dataset containing information on the fraction of a Census tract's area that is covered by an Impact Zone as well as the distance from an Impact Zone for tracts that are not

[12]In June of 2015, Commissioner Bratton established Neighborhood Coordinating Officers that were tasked with spending time each day engaging community members and nurturing relationships within their sector.

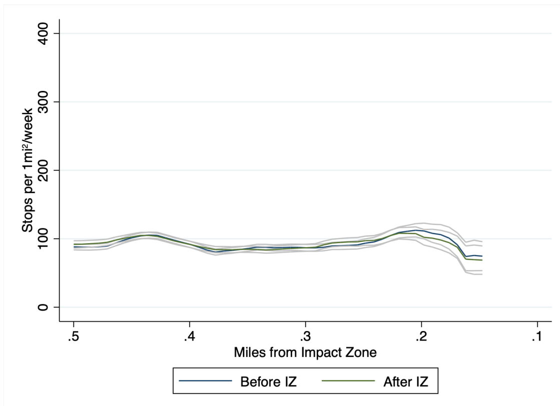
touched by an Impact Zone.[13] In a fashion similar to Ang (2020), we plot the distance gradient of changes in stops and major felonies for the five months before and after an assignment event.[14] Figure 1.3 shows changes in stops and major felonies for tracts that prior to an assignment event were greater than 0.5 miles from an Impact Zone but, following assignment, were partially covered by an Impact Zone (Panels B and D) or were within 0.5 miles (Panels A and C). Panels A and C demonstrate that both stops and major felonies were initially unaffected by the assignment of a neighboring tract to an Impact Zone. Panel B demonstrates that stop rates begin to diverge when 10% or more of a tract is covered by an Impact Zone, and reach about 35% of the pre-assignment mean at 20% coverage. To be conservative, we define Impact Zone treatment as a tract with over 5% of its area covered by an Impact Zone.

[13]To ease computation, we measure distance as the minimum distance from the centroid of a Census block within an uncovered Census tract to a Census block within a "covered" Census tract.

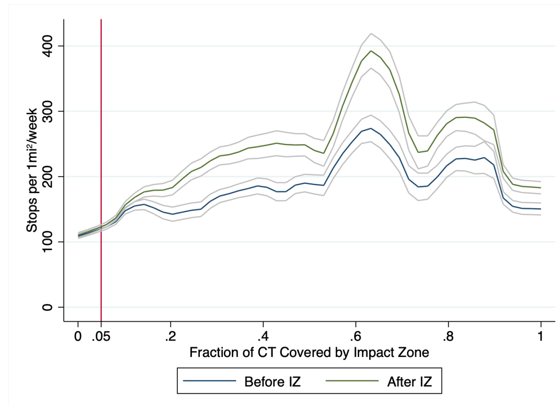
[14]Specifically, we include the 23 weeks before and after Impact Zone assignment since all assignment windows in our study period have a duration of at least 23 weeks.

Figure 1.3: *Impact Zone Assignment on Stops and Major Felonies*

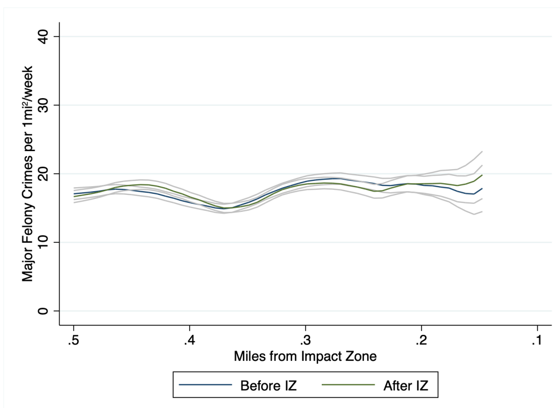
(a) *Stops: Neighboring Census Tracts*



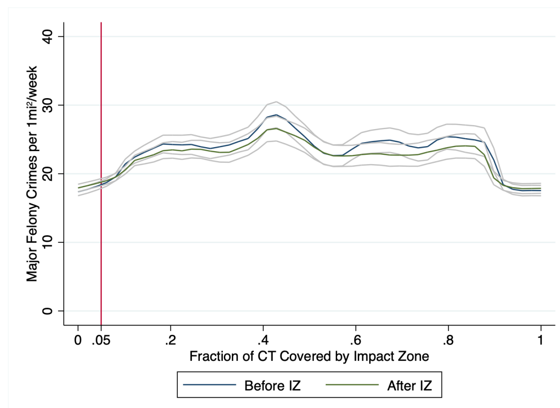
(b) *Stops: Impact Zones*



(c) *Major Felonies: Neighboring Census Tracts*



(d) *Major Felonies: Impact Zones*



Notes: This figure displays local polynomial regressions of stops and major felony crimes per square mile by Impact Zone proximity, separately for before and after an Impact Zone assignment event. All outcomes are residualized by week-year fixed effects and then added back to the constant term of the fixed effect regression. The estimation sample consists of tract-week observations for the 23 weeks surrounding an assignment event, since this is the minimum duration of all assignment periods. Panels B and D include neighborhoods that, following assignment, are partially covered by an Impact Zone, while Panels A and C include neighborhoods within 0.5 miles of a Impact Zone. Standard errors are estimated using pilot bandwidths equal to 1.5 times the kernel bandwidths and are depicted as grey bandwidth lines (Fan *et al.*, 1996).

1.3 Data

1.3.1 NYPD Data on Policing and Crime Measures

Policing and crime data for primarily come from administrative NYPD records. These data include information on pedestrian stops from the "Stop, Question, and Frisk" program, information on criminal activity compiled from crime complaints, arrest records, and shooting data, and boundary maps on the roll-out of Operation Impact.

Stops were recorded by NYPD officers on the "Stop, Question, and Frisk Report Worksheets", known as form UF-250, and contain geographic coordinates of where a stop occurred, date and time-of-day information, demographic information on the person stopped (i.e. race, sex, approximate age), information on police actions taken (e.g. did the officer frisk the pedestrian or use force?), and stop outcomes. Stop outcomes include whether the stop led to an arrest, the discovery of drugs or weapons, or the issuance of a court summons. We use proprietary stop data obtained by Dr. Fagan for years 2006-2016, since these data additionally contain de-identified officer IDs describing the lead officer who conducted the stop.[15] For years 2017-2019, we use publicly-available stop data.

Prior to 2012, administrative stop records were used internally to monitor whether officers were conducting a sufficiently high number of pedestrian stops, and thus under-reporting is of limited concern. While some of the observed reduction in stop rates due to *Floyd* may have resulted from under-reporting, it is likely this fraction is small. Following *Floyd*, the court-ordered monitor worked closely with NYPD brass to monitor and audit officer reporting of stops.[16] Second, if the majority of changes in stop rates were due to under-reporting, we would expect to see hit rates rapidly increasing after *Floyd*, since then the only recorded stops would be those that detected criminal activity. However, this is not the case; hit rates remain surprisingly flat throughout the first year post-*Floyd*, as stop rates

[15]These proprietary data also contain the exact location of where the stop occurred, while publicly-released stop data code stop locations at the nearest intersection.

[16]An audit of stops conducted from 2016-2019 found that officers did not report up to 30% of stops conducted. For this reason, we analyze stop outcomes data prior to 2016. The audit, however, indirectly confirmed that stop rates had indeed fallen by over 95%, since at the time of the audit stops were roughly 3% of pre-reform levels.

fell by over 35%.

Crime data capture all crime complaints reported to the NYPD since 2006, including incidents that were never closed by police. Data are at the incident-level and include information on the date the complaint was reported to police, geographic coordinates of where the incident took place, and offense categories that describe the most serious offense related to an incident. Offense information includes broad severity categories, such as whether the crime was a "felony", "misdemeanor", or "violation", as well as more granular offense types (e.g., murder, rape, felony assault, grand larceny, etc.). Crime reports can be initiated by civilians *or police officers* that observe criminal activity. To this end, police presence in a neighborhood can affect observed crime rates by altering civilian reporting and police detection of criminal activity.

We supplement crime reports with administrative records on all arrests and reported shooting incidents that occurred in NYC since 2006. Both datasets contain information on the date, time-of-day, and geographic coordinates of a given incident. Arrest data contain the same detailed offense categories recorded in crime complaints. Unlike crime reports, however, arrest records have complete race information, allowing us to explore how the racial composition of offenders evolves throughout the study period. Unfortunately, we cannot link crime reports to arrest records, and thus cannot observe how clearance rates vary across neighborhoods. We can, however, examine trends in approximate clearance rates at the city-level by dividing the total number of arrests by the total number of reported crimes over a given period (see Figure B.1). Shooting data include all reported shooting incidents, most of which do not involve police officers, and contain information on whether the shooting results in a civilian's death, which hereafter we refer to as "killings".[17] Finally, we supplement the above data with demographic information about the residential population of each tract using the 2010 Census, the 2010 American Community Survey five-year estimates, and NYC Housing Authority maps.

[17] The NYPD did not start using ShotSpotter technology until late 2015, and rolled out this technology broadly in higher-crime areas in 2016. Thus, most of the shootings that occur during our study period were reported through traditional channels.

Crime Outcomes

We first examine the impact of stop rates on two broad crime categories – felonies and violent misdemeanors, and, non-violent misdemeanors and violations. These categories are associated with substantially different penalties and social costs. In New York City, the most severe misdemeanor crimes carry a weight of up to 364 days in jail and a fine of up to $1000.[18] Violations are even less severe, carrying a weight of up to 15 days of jail time. Agan et al. (2021) show that non-prosecution of marginal non-violent misdemeanors leads to large reductions in the likelihood that a defendant is charged with a new criminal complaint over the subsequent two years. This suggests that charges for some low-level offenses, such as for graffiti or marijuana possession, may even impose negative costs on society if it raises the chance the perpetrator commits future crimes.

Felonies and violent misdemeanors, on the other hand, involve serious crimes that impose substantial costs on society. For example, Autor et al. (2017) estimate that the direct costs of the average violent felony crime is $67,794. We additionally report impacts on two mutually-exclusive sub-categories – "major felonies" and "non-major felonies and violent misdemeanors", since the police often use major felony crime to track progress in a given area.[19] Major felonies include murder and non-negligent manslaughter, rape, robbery, felony assault, burglary, grand larceny, and grand larceny of motor vehicles.[20]

Stop Outcomes

We define a pedestrian's criminal activity as whether the officer made an arrest or discovered illegal drugs or weapons on the pedestrian during the stop. We do not include whether an officer issues a court summons, since officers have considerable discretion over this

[18]In New York state, misdemeanor theft involves goods worth no more than $1,000.

[19]For example, commander performance was often judged at CompStat meetings by changes in major felonies within their designated area (e.g. precinct or borough).

[20]These seven offenses roughly track the sum of violent and property crime measures in the FBI's Uniform Crime Reports, making effects on major felonies readily comparable to other studies examining crime deterrence effects (e.g. Chalfin and McCrary (2018)).

decision, and their tendencies to issue summonses changed over the study period. In contrast, whether an officer discovers drugs or a weapon on the pedestrian is objective and does not change over the study period. In a similar vein, arrests impose a higher bar of criminal activity that involve more serious offenses, and often require the officer to take the pedestrian into the police station. Arrest tendencies may still vary over the study period, however. For this reason, we also report effects on each stop outcome separately in the Appendix.

1.3.2 High School Enrollment Data

We employ administrative enrollment records from the NYC Department of Education to examine the impact of changes in stop rates on high school dropout and discharge rates. We focus on high school students since stop rates diverge at age 14 and are the highest at ages 18-20.[21] Data were obtained through NYU's Research Alliance for New York City Schools and contain detailed enrollment information on the universe of high school students enrolled in NYC public schools from 2001 through 2019. Enrollment status is given for each student in October and June of each school year. As required by the city, school administrators must note the specific reason for why a student is not enrolled in a given semester. We use these enrollment categories to construct our two main outcomes.

The first outcome is an indicator for whether a student was discharged from high school based on an institutional directive by an non-DOE agency, court, or other authority (hereafter, "discharged by institutional directive"). This includes students who are in the custody of the Department of Corrections, or who attend special programming under the supervision of the NYS Department of Mental Health or Office of Children and Family Services. It is important to note that this outcome does not include situations where a student is arrested after they stop attending school and therefore likely provides a *lower-bound* estimate of the direct effect of how changes in stop rates affect student arrest rates.[22]

[21] See Figure A.2 for race-specific age profiles of stop rates.

[22] Students are listed as having "dropped out" after 20 days of consecutive absence.

The second outcome indicates whether the student dropped out of high school or was discharged by institutional directive. Across both outcomes, less than 7% (18%) of students go on to graduate high school within four (six) years. We set each outcome indicator to 1 if and only if the student switches from being listed as “enrolled" in October of a given school year to being marked as discharged by October of the subsequent school year. This definition allows us to isolate the exact timing of when students stop attending high school.[23] These data also contain detailed demographic information, including student race, sex, month of birth, languages spoken at home, and the Census tract of the student's home address.

We restrict the sample to high schools that were operational from school years 2010 through 2017, which includes 505 schools and 90.2% of all students enrolled in the public school system over this period.[24] We further restrict the sample to students who are enrolled in October of their initial 9th grade school year, and keep the first four school years of their high school tenure.[25] This leaves a final sample of 2,092,366 student-year observations that span eight school years.

1.4 Crime Deterrence Effects

1.4.1 Empirical Strategy

To estimate crime deterrence effects, we implement a flexible difference-in-differences (DD) framework that compares changes in crime rates induced by *Floyd* in neighborhoods that had similar crime rates but substantially different stop rates prior to the reform. This approach exploits the fact that a significant portion of the variation in neighborhood stop rates pre-*Floyd* cannot be explained by observable crime rates. Prior to the reform, “treatment" neighborhoods have high stop rates conditional on crime (H), while our “control"

[23]Note that students may drop out more than once if they re-enroll and dropout again.

[24]Throughout this section we refer to school years by spring-term year. For example, the 2010 school year refers to the 2009-2010 school year.

[25]This includes observations for students who have left the NYC school system. For example, a student who begins 9th grade in 2012 will be observed four times in our data, spanning 2012 through 2015 school years, even if they choose to leave school after their junior year.

neighborhoods exhibit medium stop rates conditional on crime (M). After *Floyd*, stop rates in both groups fall to very low levels (L). Therefore, we estimate the crime deterrence effect of conducting $H - M$ additional stops prior to the reform by subtracting changes observed in control neighborhoods from changes observed in treated neighborhoods

$$\left[\frac{d\text{Crime}}{ds_n}\right]_{H\to M} = \left[\frac{d\text{Crime}}{ds_n}\right]_{H\to L} - \left[\frac{d\text{Crime}}{ds_n}\right]_{M\to L} \tag{1.1}$$

where we expect $\left[\frac{d\text{Crime}}{ds_n}\right]_{H\to M} > 0$ if the additional stops have a crime deterrence effect.

To separate neighborhoods into treatment and control groups, we first split the sample into three distinct periods. The first period is the "training" period and spans January 1st, 2006 through April 15th, 2008. The second period is the "pre-period" and contains the four years leading up to the reform. The third and final period is the post-*Floyd* period and spans April 16th, 2012 through April 15th, 2017.

Restricting data to the training period, we empirically identify neighborhoods that experience disproportionately high stop rates conditional on observed crime rates by running the following least-squares regression for each tract n in week-year t:

$$Stops_{n,t} = \alpha_S Shootings_{n,t} + \alpha_F Felonies_{n,t} + \alpha_M Misdemeanors_{n,t} + \Gamma X'_n + \delta_t + \epsilon_{n,t} \tag{1.2}$$

$Shootings_{n,t}$, $Felonies_{n,t}$, and $Misdemeanors_{n,t}$ are vectors that include up to third-order polynomial terms as well as additional linear measures of average crime rates of Census tracts within 0.5 and 1.0 miles of a given neighborhood.[26] δ_t are week-year fixed effects and X'_n includes neighborhood deciles of total land area and total population, as recorded in the 2010 Census. For each n, we then calculate the mean stop residual during the training period – $\bar{\epsilon}_n^{Train} = \frac{\sum_{t\in Train}\epsilon_{nt}}{\sum_{t\in Train} I(t)}$ – and split neighborhoods into "treatment" and "control" groups at the median. Our findings are robust to alternative choices of covariates in Equation 1.2 and to using precinct-level variation in stop exposure.

Table 1.2 shows that, relative to control neighborhoods, treatment neighborhoods experi-

[26]Specifically, we calculate average crime rates in Census tracts with centroids within 0.5 and 1.0 miles from the centriod of a given Census tract n

ence almost twice as many stops, 80% more frisks, and 60% more uses of force during the pre-period. On average, stops are also 14% less likely to detect criminal behavior in treated neighborhoods. The allocation of officers across neighborhoods is likely the main reason for higher stop rates in treated neighborhoods since the number of officers conducting stops in a 30 day-span is 80% higher in treated neighborhoods whereas the average number of stops per officer is only 9% higher in treatment neighborhoods. Reported crime rates are comparable across study groups in the pre-period. While treatment neighborhoods experience slightly more weekly shootings per square mile, all other crime categories (major felonies, other felonies, and misdemeanors and violations per $1\text{mi}^2/\text{week}$) are slightly higher in the control group. A variety of neighborhood characteristics are also comparable across study groups, including median household income and the percent of residents who are Hispanic, live in single-parent households, or live below the federal poverty line. Treated neighborhoods differ in that residents are more likely to be Black, a larger fraction of the area contains NYC Housing Authority developments, and neighborhoods are less densely populated.

With treatment defined, we restrict the data to the four years before the reform through five years after the reform. We then compare how outcomes evolve for treatment neighborhoods relative to control neighborhoods by implementing the following difference-in-differences regression:

$$Y_{nt} = \sum_{\tau \neq -1} \beta_\tau TREAT_\tau + \delta_n + \delta_{p \times t} + \epsilon_{nt} \tag{1.3}$$

where τ denotes event-time in years relative to the onset of *Floyd*. δ_n are neighborhood fixed effects and $\delta_{p \times t}$ are precinct-by-week-year fixed effects. A neighborhood is defined as a Census tract, which covers an average area of 0.12 square miles or four-by-four blocks. $TREAT_\tau$ are relative time to treatment indicators that equal 1 for treatment neighborhoods in year τ. The coefficients of interest, $\{\beta_\tau\}$, then estimate the average change in Y between year τ and the reference year in treated neighborhoods relative to that same change over time in control neighborhoods in the same precinct p. We set the reference year as two years prior to *Floyd* (e.g. $\tau = -1$) since the mean stop rate in this year is close the four-year pre-period mean. All standard errors are clustered at the precinct-level, allowing for serial correlation within each of the 76 precincts (Bertrand et al., 2004). Our *identifying assumption* is therefore the standard common trends assumption: treatment and control neighborhoods would exhibit common trends in crime outcomes absent *Floyd*. With this assumption, any differential changes in crime rates in response to *Floyd* can then be attributed to differences in pre-*Floyd* stop rates.

1.4.2 Main Results

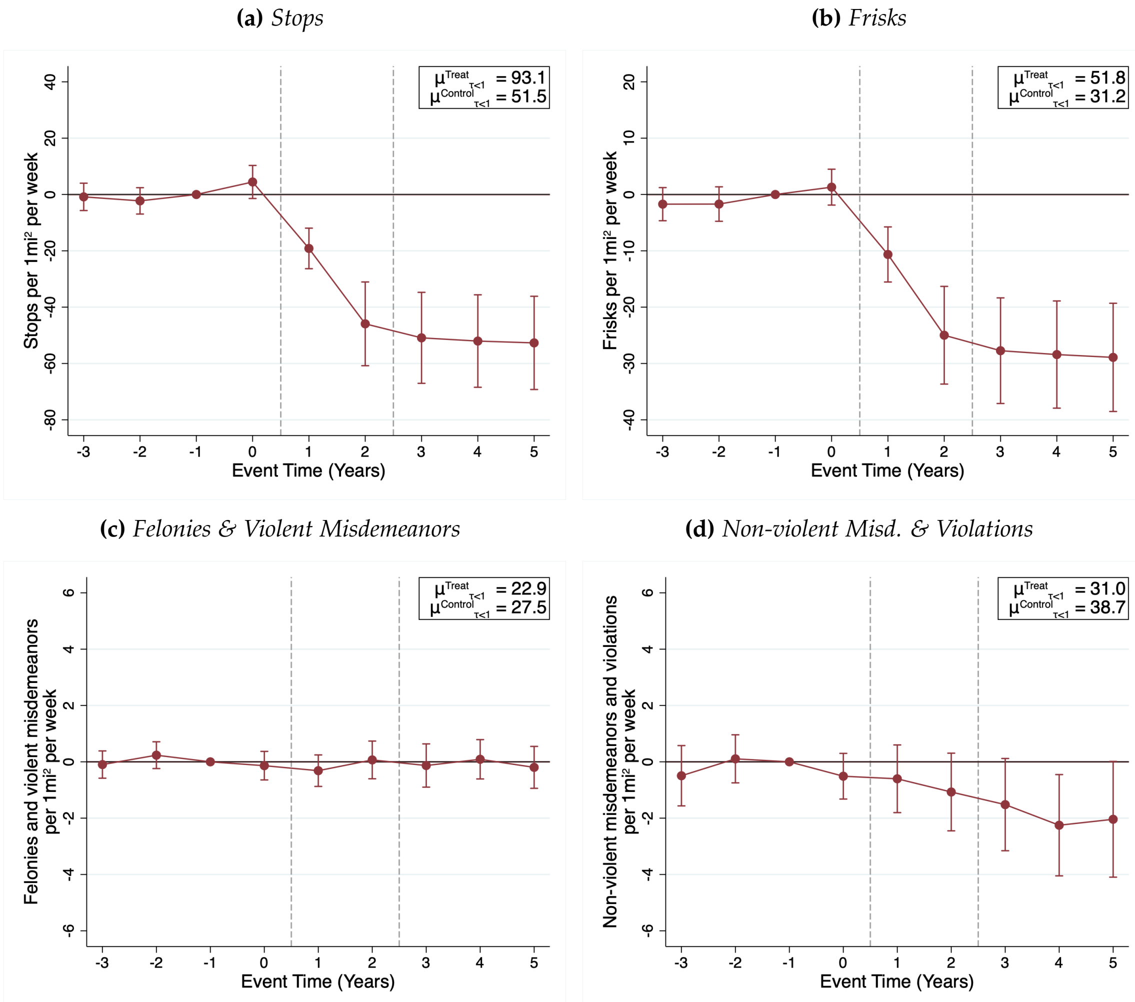
We first explore the causal effect of the *Floyd* by estimating Equation 1.3 on stops, frisks, felonies and violent misdemeanors, and non-violent misdemeanors and violations. All data are at the tract-by-week level and outcomes are normalized to weekly rates per square mile. The omitted year is two years prior to *Floyd*. Estimates are displayed in Figure 1.4.

Table 1.2: *Balance between Treatment and Control Neighborhoods prior to Floyd*

	Treatment	Control
Panel A: Stop, Question, & Frisk		
Stops	98.680	50.037
Frisks and Searches	55.389	30.151
Uses of Force	21.282	13.351
% Stops with Arrest or Weapons/Drugs Found	0.101	0.117
% Stops with Arrest	0.068	0.081
% Stops with Weapon Found	0.013	0.015
% Stops with Durgs Found	0.020	0.021
Mean Stops per Officer in 30 Days	1.518	1.394
Number of Officers with Stop Made in 30 Days	17.762	9.730
Panel B: Reported Crimes		
Shootings	0.253	0.217
Major Felonies	11.369	13.921
Other Felonies	4.837	5.133
Misdemeanors & Violations	38.921	45.254
All Black Arrests / All White Arrests	2.176	1.614
All Hispanic Arrests / All White Arrests	1.512	1.466
Major Felony Black Arrests / Major Felony White Arrests	0.615	0.629
Major Felony Hispanic Arrests / Major Felony White Arrests	0.400	0.517
Panel C: Neighborhood Characteristics		
Census Tract Area in Square Miles	0.169	0.106
Population in 2010 Census	3391.7	4172.9
% Black	0.268	0.223
% Hispanic	0.268	0.265
% Census Tract Area Belonging to NYC Housing Authority	0.037	0.005
Median Househould Income	57030.5	57110.8
% Residents Living Below Federal Poverty Line	0.161	0.158
% Residents with Bachelor's Degree or Higher	0.320	0.362
% Residents with Kess than HS Degree	0.182	0.172
% Residents with SSI	0.070	0.068
% Residents Living in Single-parent Household	0.315	0.291

Notes: This table reports means of outcomes and neighborhood characteristic for Treatment and Control neighborhoods during the three years prior to *Floyd*. *Panel A* reports SQF outcomes, *Panel B* provides means for various reported crimes and racial arrest ratios, while *Panel C* provides means of various neighborhood characteristics. N = 2,058 Census tracts split evenly into 1,079 Treatment and 1,079 Control neighborhoods.

Figure 1.4: *Floyd Difference-in-differences Estimates*



Notes: This figure graphs coefficients from Equation 1.3 on stops, frisks, felonies and violent misdemeanors, and non-violent misdemeanors and violations. Maroon dots denote point estimates and whiskers show 95% confidence intervals. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling. All standard errors are clustered at the precinct level. Pre-period outcome means are given for treatment and control neighborhoods in the northeast corner of each figure.

Prior to *Floyd*, Panels A and B display small positive trends in stops and frisks for treated neighborhoods relative to control neighborhoods, suggesting that our approach slightly understates treatment-control differences in stop rates. Due to our choice of $\tau = -1$ as the reference period, however, pre-treatment coefficients are jointly insignificant for both stops ($F = 0.06, p = 0.814$)) and frisks ($F = 0.31, p = 0.580$). The first dashed line represents the police's initial response to the court allowing expert testimony. In the following year, weekly stop (frisk) rates fell by an additional 20 stops (10 frisks) per square mile in treated neighborhoods relative to control neighborhoods. The second dashed line marks the post-ruling phase. By year three, weekly stop (frisk) rates have fallen by an additional 50 stops (28 frisks) in treated neighborhoods relative to control neighborhoods – a difference that is roughly the magnitude of the control group's pre-period mean.

Panel C displays impacts on our primary crime outcome – reported felonies and violent misdemeanors. For $\tau < 1$, there is no evidence of differential pre-trends. All treatment coefficients are less than 0.24 weekly crimes per square mile and each coefficient is not statistically significant, even at the 30-percent level.[27] Pre-treatment estimates are also jointly insignificant ($F = 0.00, p = 0.996$). Panel D shows a similar lack of pre-trend differences for reported non-violent misdemeanors and violations, as pre-treatment point estimates alternate signs and are jointly insignificant ($F = 0.78, p = 0.381$).

We find no evidence of increases in felonies or violent misdemeanors during the post-*Floyd* period. Post-treatment coefficients alternate signs, are statistically insignificant, and are close to zero. Table 1.3 reports coefficients from Equation 1.3 using a single post-*Floyd* treatment indicator. We estimate an effect of -0.097 crimes per 1mi^2 or -0.4% of the pre-period mean ($p = 0.672$). Put differently, treated neighborhoods that had twice as many stops in the four years prior to *Floyd* do not experience relative increases in serious crimes. In fact, our 95% confidence interval can rule out an increase of 1.5% of the pre-period mean. To put these effects in context, New York City saw a 38% decline in felony crime rates from 2000 to 2010. These estimates suggest that increasing stop rates did not play a significant

[27] The largest coefficient is just $\frac{1}{95}$th of the pre-treatment mean.

role in reducing serious crime.

Also contrary to the crime deterrence hypothesis, Panel D reveals that treated neighborhoods experience a relative reduction in non-violent misdemeanors and violations following the reform. We show in Table 1.3 that these reductions can be entirely explained by reductions in stop-related arrests, which predominantly involve non-violent misdemeanor offenses.[28] These findings suggest that the main result of heightened stop rates is increased detection and enforcement of minor offenses.

A full summary of our findings is provided in Table 1.3. In addition to our main outcomes, we report impacts on frisks, uses of force, shootings, and killings, and disaggregate reported crimes into major felonies, non-major felonies and violent misdemeanors, and non-violent misdemeanors and violations. Differential reductions in stops, frisks, and uses of force in treated neighborhoods were about double pre-period levels in the control group. Estimates of effects on shootings, killings, major felonies, and non-major felonies and violent misdemeanors confirm the notion that serious crime was unaffected by *Floyd*; they are close to zero and statistically insignificant.

One may be concerned that aggregate differences across treatment and control groups may mask crime increases among higher-crime treated neighborhoods. We investigate this by estimating effects separately for the highest felony-crime quartiles in Table 1.4, where quartiles are measured using training period data. The accompanying figure – Figure A.3 – displays coefficients from Equation 1.3 for neighborhoods with above-median felony crime rates during the training period. There is no evidence of differential increases in felonies and violent misdemeanors, shootings, killings, major felonies, or non-major felonies and violent misdemeanors in higher-crime treated neighborhoods. Additionally, aggregate reductions in non-violent misdemeanors and violations driven almost entirely by a 7% reduction in crimes in the highest-crime quartile ($p = 0.01$).

[28]Stop-related court summonses are not included in our crime reports data, while stop-related arrests are.

Table 1.3: *Effect of Floyd on Neighborhood Crime*

	Pre-period Mean		$\beta_{\tau\geq1}$
	Control	Treatment	
	(1)	(2)	(3)
Stops per 1mi^2/week	51.471	93.143	-44.475***
			(7.031)
Frisks per 1mi^2/week	31.196	51.793	-23.611***
			(3.964)
Uses of force per 1mi^2/week	13.837	20.302	-10.490***
			(2.215)
Stop-related arrests per 1mi^2/week	3.429	5.366	-2.096***
			(0.348)
Stop-related court summonses per 1mi^2/week	3.254	5.844	-2.660***
			(0.463)
Shootings per 1mi^2/week	0.235	0.240	0.007
			(0.014)
Killings per 1mi^2/week	0.045	0.043	0.000
			(0.004)
Felonies & violent misd per 1mi^2/week	27.457	22.938	-0.097
			(0.229)
Major felonies per 1mi^2/week	14.406	11.298	0.123
			(0.144)
Non-major fel & violent misd per 1mi^2/week	13.052	11.640	-0.220
			(0.161)
Non-violent misd & violations per 1mi^2/week	38.674	31.031	-1.272*
			(0.669)

Notes: This table reports estimates for various SQF and reported crime outcomes from Equation 1.3, except that we collapse the five-year post-*Floyd* period into a single indicator. Columns (1) and (2) report outcome means during the four years prior to *Floyd* for control and treatment neighborhoods, respectively. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N = 2,058 Census tracts with 50-50 % split of Treatment and Control neighborhoods.

Table 1.4: *Effect of Floyd on Neighborhood Crime in Highest Crime Neighborhoods*

	Felony Rate Q3		Felony Rate Q4	
	Pre-per. T Mean	$\beta_{\tau\geq1}$	Pre-per. T Mean	$\beta_{\tau\geq1}$
Stops per 1mi^2/week	113.960	-50.906***	263.931	-115.148***
		(4.865)		(11.276)
Frisks per 1mi^2/week	66.026	-28.266***	144.071	-60.456***
		(3.267)		(6.703)
Uses of force per 1mi^2/week	24.623	-11.443***	58.087	-28.319***
		(2.055)		(5.010)
Stop-related arrests per 1mi^2/week	6.394	-2.420***	15.291	-5.869***
		(0.376)		(0.820)
Stop-related court summonses per 1mi^2/week	7.141	-3.258***	17.106	-6.663***
		(0.437)		(0.857)
Shootings per 1mi^2/week	0.279	-0.023	0.717	0.031
		(0.024)		(0.047)
Killings per 1mi^2/week	0.053	-0.003	0.121	0.003
		(0.010)		(0.014)
Felonies & violent misd per 1mi^2/week	26.512	-0.143	61.645	-1.003
		(0.375)		(0.910)
Major felonies per 1mi^2/week	12.997	0.103	29.308	-0.097
		(0.177)		(0.557)
Non-major fel & violent misd per 1mi^2/week	13.515	-0.246	32.337	-0.906
		(0.272)		(0.652)
Non-violent misd & violations per 1mi^2/week	32.766	-1.051*	87.033	-6.027***
		(0.608)		(2.255)

Notes: This table reports difference-in-differences estimates for various SQF and reported crime outcomes using Equation 1.3 with a singular post-*Floyd* indicator for each felony crime quartile as measured during the training period. "Pre-per. T mean" refers to the mean outcome in the Treatment group during the four years prior to *Floyd*. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Crime Reporting

While it is possible that null effects reflect offsetting changes in civilian crime reporting and criminal behavior, this seems unlikely in our setting. Reduction in aggressive interactions with the police likely increase crime reporting (if at all) and thus would serve to inflate our estimates of crime deterrence. We also test crime reporting directly by estimating effects on deadly shootings, murders, and thefts of expensive goods, such as cars or goods worth over $1,000, since these outcomes are less likely to be affected by changes in reporting. Table 1.5 shows that all outcome coefficients are close to zero and statistically insignificant.

Table 1.5: *Effect of Floyd on Crime Categories Less Affected by Reporting Concerns*

	Pre-period Mean		$\beta_{\tau \geq 1}$	P-value
	Control	Treatment		
	(1)	(2)	(3)	(4)
Killings per 1mi^2/week	0.045	0.043	0.000	0.970
			(0.004)	
All murders per 1mi^2/week	0.058	0.056	-0.006	0.168
			(0.004)	
Grand larceny crimes per 1mi^2/week	6.751	4.839	0.046	0.685
			(0.113)	

Notes: This table reports difference-in-differences estimates on various categories of crimes per square mile using Equation 1.3 with a singular post-*Floyd* indicator. Killings refer to deadly shooting incidents, murders refer to all murder and non-negligent manslaughter crimes. Grand larceny includes all car thefts as well as thefts of property above $1,000 in value. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Robustness

Table 1.6 displays robustness to a variety of alternative specifications, treatment definitions, and sub-samples. Column 1 reports estimates from our preferred specification described in Equation 1.3 using a singular post-treatment indicator variable. Column 2 and 3 replace precinct-by-week-year fixed effects with week-year fixed effects and borough-by-week-year fixed effects, respectively. To confirm that our results are not sensitive to alternative ways of estimating pre-*Floyd* stop residuals, Columns 4 through 6 explore alternative treatment definitions. Column 4 assigns treatment based on precinct-level variation in stop rates during the training period by running Equation 1.2 at the precinct-level.[29] Columns 5 and 6 explore alternative ways to define treatment at the tract-level. Column 5 removes all covariates from Equation 1.2 except for mean shootings, major felonies, non-major felonies, and misdemeanors per 1mi^2/week and week-year fixed effects. Column 6, on the other hand, adds various time-invariant neighborhood characteristics to Equation 1.2.[30] Finally, Columns 7 to 9 restrict the sample to higher-crime neighborhoods where stops are most likely to deter crime. Columns 7, 8, and 9 respectively restrict the sample to neighborhoods that are ranked in the top 75%, 50%, and 25% of neighborhoods in terms of mean felony crimes per 1mi^2/week during the training period.

Across all specifications, reductions in stop rates are substantially larger in treated neighborhoods. The average additional decline in treated relative to control neighborhoods ranges from 46 - 162% of the pre-period control group mean. Panel C displays point estimates of effects on major felonies. Importantly, these coefficients are statistically insignificant and close to zero, spanning +1.0 to -1.3 percent of the treatment group's pre-period mean. In fact, our preferred specification is the largest positive estimate among the group. These findings support the notion that major felonies did not differentially increase in tracts (or

[29] Given that treatment variation then occurs at the precinct-level, we replace precinct-time fixed effects with borough-time fixed effects.

[30] These include indicators for having 0-10% or $\geq$ 10% of tract area covered by public housing, and decile-fixed effects for the following neighborhood characteristics: % Black, % Hispanic, % families living below the poverty line, median household income, and % of adult residents with less than a HS degree.

precincts) that experienced significantly larger declines in stop rates. Across seven of the nine specifications presented in Panel B, felonies and violent misdemeanors are similar across treatment and control groups. Columns 5 and 6 report modest reductions in felonies and violent misdemeanors of about 2% of the pre-period mean, driven by declines in non-major felonies and violent misdemeanors.

Figure A.4 displays a sensitivity analysis per Roth and Rambachan (2021) that tests the robustness of effects on felonies and violent misdemeanors to parallel trend violations. This figure reports 95% confidence intervals allowing for a change in linear slope between any two consecutive periods of up to M. Given the lack of pre-trends, we find the confidence sets for each post-reform coefficient is centered around zero. Setting M at 0.20 crimes per $1mi^2$/week, the adjusted 95% confidence interval on the first-year coefficient rules out increases greater than 1 crime per $1mi^2$/week or 2.3% of the pre-period treatment mean. These confidence intervals gradually grow with each additional year since the reform; year three estimates, for example, can rule out increases greater than 3 crimes per $1mi^2$/week (6.9%). Precision of four and five year estimates requires stronger assumptions on trend violations, closer to the standard assumption of parallel trends.

Table 1.6: *Robustness of Estimated Effects of Floyd Neighborhood Crime*

	Base	Alt. Time FE		Alt. Treat			Alt. Sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Stops per 1mi2/week									
Pre-per. Control Mean	51.471	51.471	51.471	61.336	32.523	59.855	58.695	75.920	107.477
Pre-per. Treat Mean	93.143	93.143	93.143	83.298	112.091	84.758	137.359	188.293	263.931
$\beta_{\tau \geq 1}$	-44.475***	-34.729***	-47.636***	-28.518**	-52.823***	-30.260***	-64.972***	-90.549***	-115.148***
	(7.031)	(8.937)	(8.962)	(12.240)	(5.735)	(6.003)	(7.769)	(8.608)	(11.276)
Panel B: Felonies & violent misdemeanors per 1mi2/week									
Pre-per. Control Mean	27.457	27.457	27.457	25.623	21.866	27.783	31.368	40.588	57.888
Pre-per. Treat Mean	22.938	22.938	22.938	24.772	28.530	22.613	33.018	43.926	61.645
$\beta_{\tau \geq 1}$	-0.097	0.223	0.085	0.141	-0.510**	-0.487*	-0.204	-0.491	-1.003
	(0.229)	(0.268)	(0.264)	(0.518)	(0.232)	(0.276)	(0.317)	(0.494)	(0.910)
Panel C: Major felonies per 1mi2/week									
Pre-per. Control Mean	14.406	14.406	14.406	13.327	12.029	14.173	16.413	21.055	29.589
Pre-per. Treat Mean	11.298	11.298	11.298	12.376	13.674	11.531	16.114	21.082	29.308
$\beta_{\tau \geq 1}$	0.123	0.001	0.090	0.121	-0.181	-0.153	0.094	0.032	-0.097
	(0.144)	(0.169)	(0.178)	(0.374)	(0.131)	(0.145)	(0.194)	(0.300)	(0.557)
Panel D: Non-major felonies & violent misdemeanors per 1mi2/week									
Pre-per. Control Mean	13.052	13.052	13.052	12.296	9.836	13.610	14.955	19.533	28.299
Pre-per. Treat Mean	11.640	11.640	11.640	12.396	14.856	11.082	16.904	22.844	32.337
$\beta_{\tau \geq 1}$	-0.220	0.222	-0.005	0.020	-0.329*	-0.333*	-0.298	-0.523	-0.906
	(0.161)	(0.234)	(0.192)	(0.385)	(0.166)	(0.192)	(0.221)	(0.346)	(0.652)
Observations	1,009,944	1,009,944	1,009,944	1,009,944	1,009,944	1,009,944	757,224	504,972	252,252

Notes: This table shows DD coefficients from estimation of Equation 1.3, replacing event-time-treatment indicators with a post-treatment indicator. Panel A reports coefficients for stops per 1mi^2/week and Panel B explores felony and violent misdemeanor crimes per 1mi^2/week. Panels C and D disaggregate the outcome in Panel B into major felonies per 1mi^2/week and non-major felonies and violent misdemeanors per 1mi^2/week, respectively. Column 1 presents our base specification. Column 2 and 3 replace precinct-time FE with time FE and county-time FE, respectively. Column 4 assigns treatment based on precinct-level variation in stop rates during the training period by running Equation 1.2 at the precinct level. Since treatment varies at the precinct level, we replace precinct-time FE with county-time FE. Columns 5 and 6 explore alternative ways to define treatment at the tract level. Column 5 controls linearly for shootings, major felonies, non-major felonies, and misdemeanors in the training period. Column 6 adds various time-invariant neighborhood characteristics to Equation 1.2. These include indicators for having 0-10% or $\geq$ 10% of tract area covered by public housing or having, and decile fixed effects for the following neighborhood characteristics: % Black, % Hispanic, % families living below the poverty line, median household income, and % of adult residents with less than a HS degree. Columns 7-9 respectively restrict the sample to neighborhoods that are ranked in the top 75%, 50%, or 25% in terms of felony crimes per 1mi^2/week during the training period. All standard errors are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. N = 2,058 Census tracts with 50-50 % split of Treatment and Control neighborhoods.

1.4.3 Police Surges

Do more stops combined with *more patrol officers* deter neighborhood crime? The lack of crime responses to the *Floyd* reform does not provide information on whether street patrols, more generally, are ineffective at deterring crime. Following the reform, patrol officers continued to be assigned to crime hot-spots even though they stopped pedestrians at reduced rates in these neighborhoods. To decouple the crime effect of stop rates from officer presence, we assess the impact of police surges prior to the reform that simultaneously increased officer assignments and stop rates in higher-crime neighborhoods. As discussed earlier, neighborhoods selected as Impact Zones received an influx of officers that increased foot patrols, conducted a high volume of pedestrian stops, and strictly enforced minor infractions. Table 1.7 shows that, prior to assignment, neighborhoods selected as "Impact Zones" have higher rates of crime and stops, and are more disadvantaged along a variety of measurable characteristics (i.e. median household income, federal poverty rate, education).

We begin by estimating crime deterrence effects using a standard difference-in-differences framework:

$$Y_{n,t} = \sum_{\tau \neq 0} \beta_\tau IZ_\tau + \Gamma X'_{nt} + \delta_n + \delta_{p\,\times\,t} + \epsilon_{nt} \tag{1.4}$$

where τ denotes six-month periods relative to neighborhood n's next Impact Zone assignment, and IZ_τ indicates whether neighborhood n has more than 5% of its total area covered by an Impact Zone in the post-assignment window ($\tau \geq 1$). $\delta_{p\,\times\,t}$ are precinct-by-week-year fixed effects and δ_n are neighborhood fixed effects. We restrict our attention to the three years surrounding a given neighborhood's assignment event (i.e. $\tau \in [-5, 6]$). X'_{nt} is a control vector that includes indicators equal to one for periods outside of this three-year window, fixing the reference period as the six months prior to assignment (i.e., $\tau = 0$). Standard errors are clustered at the precinct-level (Bertrand et al., 2004).

Our preferred specification adapts Equation 1.4 by dropping pre-treatment indicators ($\sum_{\tau=-5}^{-1} \beta_\tau IZ_\tau$) and adding a term that allows Impact Zone neighborhoods to exhibit linear differences in pre-trends:

$$Y_{n,t} = \sum_{\tau=1}^{6} \beta_\tau IZ_\tau + \beta_{trend}[IZ(-5 \geq \tau \leq 6) \text{ x } \tau] + \Gamma X'_{nt} + \delta_n + \delta_{p \text{ x } t} + \epsilon_{nt} \tag{1.5}$$

where $[IZ(-5 \geq \tau \leq 6) \text{ x } \tau]$ captures the linear difference in slopes for treated neighborhoods relative to control neighborhoods during the three-year pre-period. We take this approach since Figure 1.5 reveals felony and violent misdemeanor rates are increasing at a faster rate in Impact Zone neighborhoods relative to other control neighborhoods. Our coefficients of interest (β_τ) then represent the mean differential change in Y_{nt} between event-time τ relative to a linear trend observed during the pre-period among treated neighborhoods relative to that same change in Y_{nt} over time among untreated neighborhoods in precinct p. Here, the *identifying assumption* is that differences in neighborhood outcomes between treatment and control neighborhoods would have continued to follow a linear trend absent Impact Zone assignment. Below we relax this assumption by allowing for parallel trend violations a la Roth and Rambachan (2021), finding results to remain qualitatively unchanged.

Figure 1.5 graphs coefficients from Equation 1.4 on various policing and crime measures. Panels A through C illustrate a strong first stage. For $\tau < 0$, there are no visible pre-trend differences in area covered by an Impact Zone, the number of officers conducting stops, or the number of weekly stops per square-mile. Following assignment, the average "treated" neighborhood has 40% of it's area covered by an Impact Zone, and both the number of officers conducting stops and weekly stop rates jump up by about one-third.

Panels D through F respectively display impacts on reported felonies and violent misdemeanors, major felonies, and non-violent misdemeanors and violations. While there are no visible pre-trend differences between study groups for major felonies or non-violent misdemeanors and violations, there is a modest positive pre-trend difference in felonies and violent misdemeanors. This likely stems from the fact that Impact Zones were, in part, selected because of abnormal increases in crime rates.

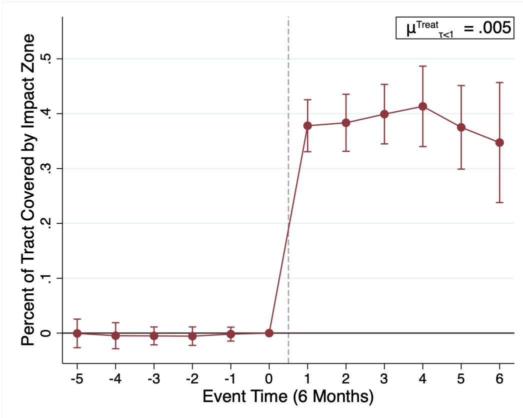
Table 1.7: *Stops, Crimes, and Characteristics of Impact Zone Neighborhoods*

	Not Impact Zone			Impact Zone	
	$>$ 1.0 Mi	1.0 to 0.25 Mi	0.25 to 0.0 Mi	$-5 \leq \tau \leq 0$	$1 \leq \tau \leq 6$
Panel A: Stop, Question, & Frisk					
Stops	34.709	76.267	173.217	155.975	220.969
Frisks and Searches	17.739	43.089	102.157	87.322	126.627
Uses of Force	7.130	17.630	49.300	38.217	51.036
% Stops with Arrest or Weap/Drugs Found	0.105	0.107	0.086	0.087	0.072
% Stops with Arrest	0.073	0.074	0.058	0.057	0.047
% Stops with Weapon Found	0.011	0.013	0.011	0.012	0.012
% Stops with Drugs Found	0.021	0.021	0.017	0.018	0.014
Mean Stops per Officer in 30 Days	1.388	1.470	1.521	1.529	1.571
Number of Officers Making Stops in 30 Days	8.079	13.902	26.190	24.951	33.664
Panel B: Reported Crimes					
Murders	0.019	0.058	0.137	0.139	0.132
Shootings	0.094	0.316	0.656	0.706	0.769
Major Felonies	9.317	17.766	25.357	22.219	22.226
Other Felonies	3.124	6.386	12.372	11.452	11.802
Misdemeanors & Violations	27.286	52.741	94.133	79.996	84.643
Arrest Ratio: Black / White	1.015	2.495	4.703	5.078	5.786
Arrest Ratio: Hispanic / White	0.875	2.124	3.574	4.034	3.602
Major Felony Arrest Ratio: Black / White	0.511	0.808	0.910	1.215	1.365
Major Felony Arrest Ratio: Hispanic / White	0.388	0.617	0.519	0.993	0.780
Panel C: Neighborhood Characteristics					
Census Tract Area in Square Miles	0.169	0.092	0.067	0.072	0.076
Population in 2010 Census	3600.0	3912.0	4384.7	4327.3	4250.5
% Black	0.172	0.298	0.381	0.474	0.517
% Hispanic	0.211	0.337	0.445	0.394	0.351
% of residences in NYCHA Housing	0.015	0.022	0.036	0.043	0.045
Median Household Income	63644.2	53639.3	38119.4	35914.7	35900.5
% Living Below Fed Pov Line	0.123	0.183	0.258	0.274	0.277
% with Bach Degree or Higher	0.380	0.337	0.217	0.197	0.188
% without HS Degree	0.148	0.198	0.269	0.258	0.247
% with SSI	0.060	0.072	0.099	0.107	0.096
% Single-parent Households	0.243	0.346	0.447	0.485	0.488
N (Tract x Week-years)	463,527	146,368	34,929	50,699	36,039

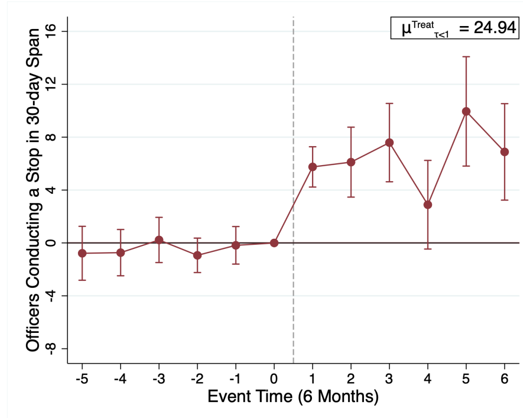
Notes: This table reports means of outcomes and neighborhood characteristics for neighborhoods by proximity to Impact Zones and time relative to Impact Zone assignment. Data are at the tract-week-year level and span Week 2 of 2006 through Week 28 of 2012. Column (1) includes Census tracts more than 1.0 miles away from an Impact zone; Column (2) includes tracts within 1.0 to 0.25 miles of an Impact Zone; Column (3) includes tracts within 0.25 to 0.0 miles of an Impact Zone; Column (4) includes week-year observations of Census tracts during the three years ($-5 \geq \tau \leq 0$) prior to Impact Zone assignment; Column (5) includes weeks in the three years following assignment ($1 \geq \tau \leq 6$). *Panel A* reports SQF outcomes, *Panel B* provides means for various reported crimes and racial arrest ratios, while *Panel C* provides means of various neighborhood characteristics. The final row reports the number of observations at the tract-by-week-year level for each study group.

Figure 1.5: *Standard Difference-in-Differences Impact Zone Estimates*

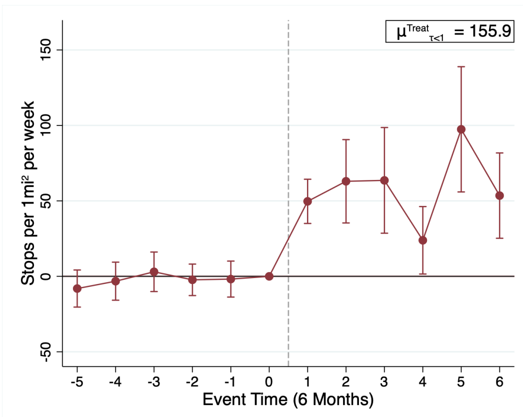
(a) *Fraction of CT Covered by an Impact Zone*



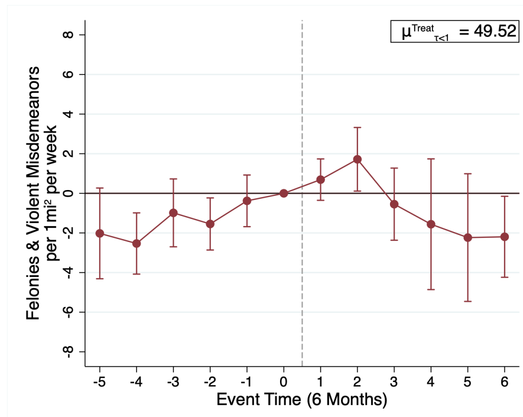
(b) *Officers Conducting a Stop in 30-day Span*



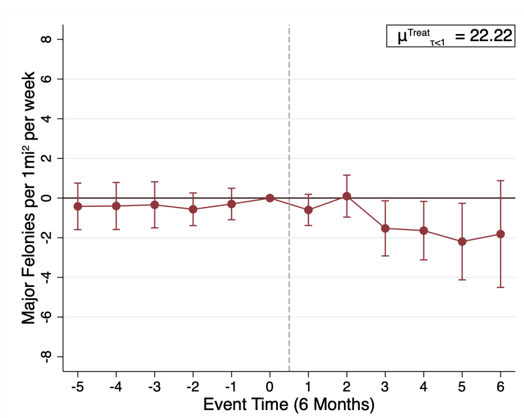
(c) *Stops*



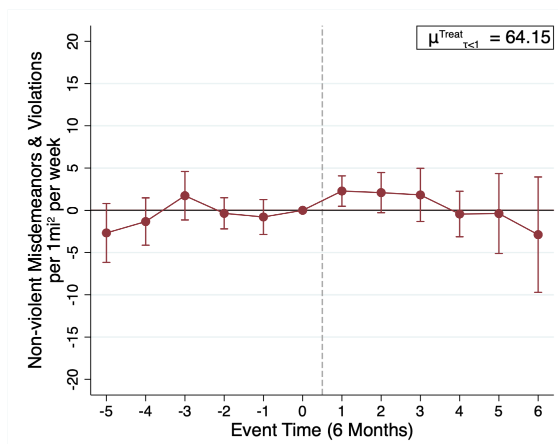
(d) *Felonies & Violent Misdemeanors*



(e) *Major Felonies*



(f) *Non-violent Misd. & Violations*



Notes: This figure graphs coefficients from a Equation 1.4. Maroon dots denote point estimates and whiskers show 95% confidence intervals. The dashed line denotes the timing of Impact Zone assignment. All standard errors are clustered at the precinct level. *Impact Zones* refer to Census tracts with $>$ 5% of their area covered by an Impact Zone. Pre-period outcome means are given for treatment neighborhoods.

Table 1.8 reports coefficients from Equation 1.5 across a variety of policing and crime outcomes. We aggregate treatment coefficients into post-assignment year indicators in order to improve statistical power. Sharp increases in stop rates are primarily driven by increases in the number of patrol officers conducting stops per month, which rose by 25-38% after assignment, compared to only a 4-6% increase in mean number of monthly stops conducted per officer. Frisks and court summonses also increase by more than one-third following assignment, while stop-related arrest rates are unaffected, implying that additional stops did not translate into additional arrests.

One year after assignment, there is no detectable impact on felonies and violent misdemeanors. However, this masks a small negative but statistically insignificant impact on major felonies and a positive and significant impact on non-major felonies and violent misdemeanors ($p < 0.01$). The latter is potentially the result of increased detection and/or stricter enforcement. Treatment effects become negative for both measures in year two, and by year three, major felonies per 1mi^2/week fall by 11% and non-major felonies and violent misdemeanors fall by 8%. Over the entire post-assignment period, reductions in felonies and violent misdemeanors are jointly significant ($p = 0.011$), with the largest reductions observed for major felonies ($p = 0.015$). As we found in the *Floyd* analysis, coefficients on non-violent misdemeanors and violations are positive in the first year post-assignment, potentially due to increased *detection*, but flip signs by year three. Post-assignment coefficients on shootings are positive but are jointly insignificant ($p = 0.236$) and coefficients on killings are close to zero and jointly insignificant.

In contrast to our *Floyd* estimates, we find that simultaneously increasing stop rates and officer presence in higher-crime neighborhoods effectively deters major felony crime. For each additional patrol officer conducting a positive number of stops per month, 21.65 major felonies are deterred each year per 10,000 neighborhood residents. This estimate is strikingly similar to the macro elasticity estimated by Mello (2018), which compares cities just beyond the COPS grant application cutoff to those just below. Mello estimates that each additional

Table 1.8: *Effect of Impact Zone Assignment on Neighborhood Policing and Crime*

	Pre-period Mean	Year 1	Year 2	Year 3	P-value
Policing	(1)	(2)	(3)	(4)	(5)
Fraction of Tract Covered by Impact Zone	0.005	0.379***	0.402***	0.361***	0.000
		(0.024)	(0.030)	(0.044)	
Stops per 1mi^2/week	155.943	55.470***	50.402***	86.492***	0.000
		(9.102)	(14.953)	(19.766)	
Officers with a Stop in 30-day Span	24.947	6.133***	6.099***	9.411***	0.000
		(0.855)	(1.490)	(2.340)	
Mean Stops per Officer	1.529	0.084**	0.061**	0.086**	0.000
		(0.018)	(0.026)	(0.030)	
Frisks per 1mi^2/week	87.305	30.730***	29.456***	31.766***	0.000
		(6.205)	(8.863)	(10.937)	
Uses of Force per 1mi^2/week	38.210	12.700***	6.990	5.020	0.046
		(4.227)	(4.867)	(7.321)	
Stop-related Arrests per 1mi^2/week	7.985	0.696	-0.536	-1.276	0.555
		(0.513)	(0.874)	(0.890)	
Stop-related Court Summonses per 1mi^2/week	11.335	3.731***	5.930***	4.311**	0.001
		(0.898)	(2.053)	(1.659)	
Crime					
Felonies & Violent Misd per 1mi^2/week	49.520	0.550	-2.287**	-4.359***	0.011
		(0.488)	(1.099)	(1.244)	
Major Felonies per 1mi^2/week	22.220	-0.381	-1.764**	-2.422**	0.015
		(0.356)	(0.739)	(1.097)	
Non-major Fel & Violent Misd per 1mi^2/week	27.301	0.931***	-0.523	-1.937**	0.346
		(0.286)	(0.750)	(0.894)	
Non-violent Misd & Violations per 1mi^2/week	64.153	2.333***	0.953	-1.319	0.656
		(0.940)	(1.533)	(2.523)	
Shootings per 1mi^2/week	0.706	0.172**	0.169	0.154	0.236
		(0.085)	(0.137)	(0.219)	
Killings per 1mi^2/week	0.139	-0.022	0.011	-0.024	0.820
		(0.033)	(0.056)	(0.080)	
Observations (tract-weeks)			86,759		
Census Tracts			385		

Notes: This table reports coefficients from Equation 1.5 for various policing and crime measures. We collapse post-assignment six-month treatment indicators into three indicators denoting each year post-assignment. "Pre-period mean" refers to the outcome mean in treated neighborhoods in the three years leading up to assignment. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Treatment is defined as $> 5\%$ of Census tract area covered by an Impact Zone. "P-val" refers to the p-value from an F-test of the null hypothesis that post-treatment indicators are jointly zero.

sworn officer reduces major felony crimes by 19.66 crimes per 10,000 residents.[31]

Robustness

The main concern with controlling for linear pre-trends is regression to the mean. Since treated neighborhoods are, in part, chosen because they display abnormally high felony crime rates in the months leading up to assignment, counterfactual crime rates might have naturally subsided in the post-period regardless of Impact Zone assignment. However, we do not observe large positive pre-trends for major felony crimes, and post-reform coefficients are significantly lower than the lowest pre-trend coefficient (see Figure 1.5). Regression to the mean cannot explain the magnitude of reductions observed for felonies and violent misdemeanors. We show this formally by allowing for parallel trend violations a la Roth and Rambachan (2021).[32] Figure A.5 shows that second- and third-year coefficients are significant at the 5% level if we allow for differential pre-trend slopes of +/- 0.35 crimes per $1mi^2$/week and +/- 0.45 crimes per $1mi^2$/week, respectively. These pre-trend violations are larger than the linear pre-trend difference in slopes observed in the data (+0.26 crimes per $1mi^2$/week).

Section A.1 of the Appendix additionally explores whether crime is displaced rather than reduced, finding no detectable changes for neighborhoods further than 0.25 miles from an Impact Zone. Neighborhoods within this radius report slight increases in non-major felonies and violent misdemeanors. However, when we expand treatment to include neighboring tracts within 0.25 miles of an Impact Zone, we estimate statistically significant reductions in felonies and violent misdemeanors as well as major felonies.

[31] This figure is obtained by adding Mello's IV estimate on violent crime (4.27) and property crime (15.39), since these categories collectively cover the seven offenses in our major felony category.

[32] Formally, Roth and Rambachan (2021) reports confidence sets on DD coefficients when allowing for violations of slope M between any two consecutive periods.

1.5 Effect on High School Dropout Rates

We are interested in estimating impacts of stops on high school dropouts since high school students face critical educational decisions and are one of the most stopped demographic groups overall (see Figure A.2).[33] We proxy for exposure to stops in a student's peer network by exploiting school-level variation in stop exposure. The advantage of this approach is that it is able to detect effects in the presence of within-school peer effects and accounts for the fact that high school students likely spend time in areas outside of their home Census tract.[34]

To this end, we split schools into quartiles based on the number of stops per square mile observed in students' home Census tracts during an earlier training period. More specifically, this process involves the following steps. First, we record the average number of stops per square mile observed for a given Census tract from 2007 to 2009. Then, we link these training-period stop rates to the Census tracts of students attending a given school during the sample period, where we have valid home tracts for over 97% of students. Finally, we calculate average training-period stop rates across all students attending a given school from 2010 to 2017 and split schools into student-weighted quartiles, such that each school quartile contains approximately 25% of the student population.[35]

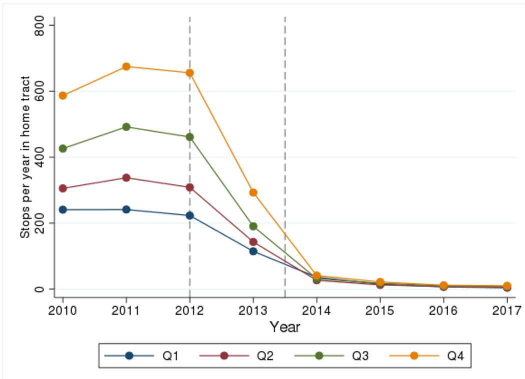
[33]High school students are 21 times more likely to be stopped than middle school students, and Black eighteen year-old residents in particular, face the highest stop rate of any sex-by-age demographic group.

[34]In Section A.2 the Appendix, we report similar but more muted results using student-level variation in neighborhood stop exposure that does not account for peer effects. Unfortunately, we lack data that identifies which individual students are stopped by police.

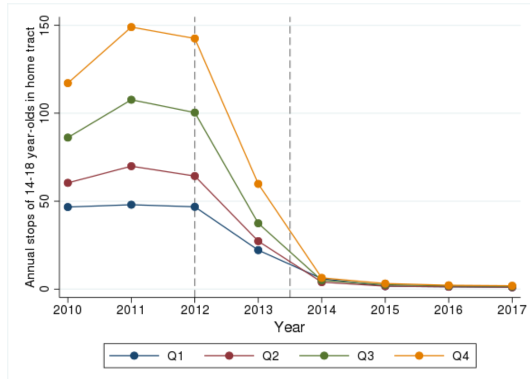
[35]Note that this approach weights training-period stop rates by the endogenous composition of student home addresses observed during the sample period. This could introduce bias if changes in stops differentially affect family location decisions across school quartiles. However, we did not find this to be the case when use tract weights from 2009 and run our main regressions on the sub-sample of schools that are operational from 2009-2017.

Figure 1.6: *Raw Means of Stops, Felonies, and HS Outcomes by School Quartile*

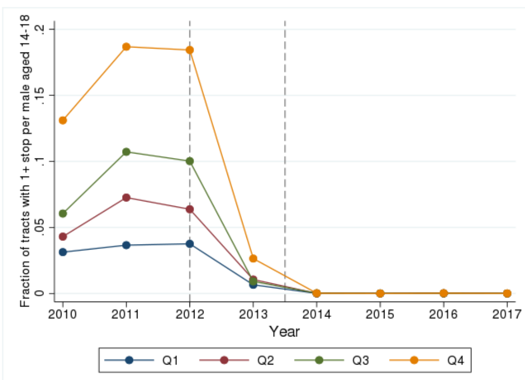
(a) *Stops per Year in Students' Home Census Tract*



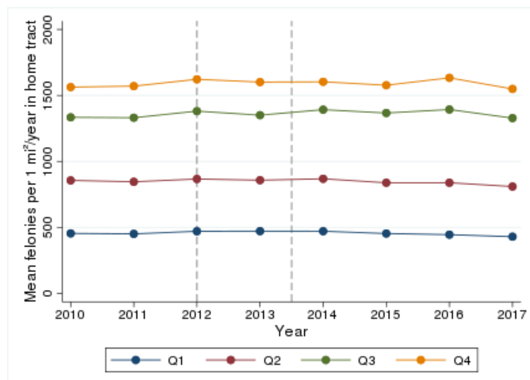
(b) *Stops per year of Pedestrians Aged 14-18 in Students' Home Census Tracts*



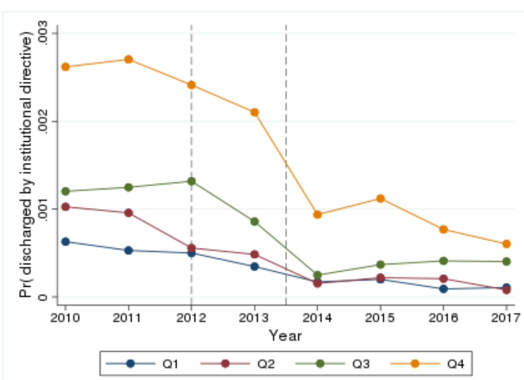
(c) *Fraction of Students' Home Census Tracts with 1+ Stop per Male, Aged 14-18*



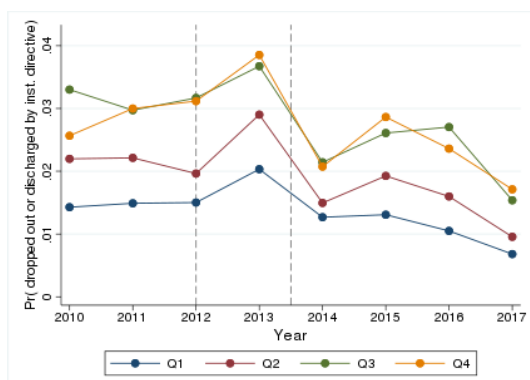
(d) *Felonies per 1 mile^2/year in Students' Home Census Tract*



(e) *Pr(Discharged by Institutional Directive)*



(f) *Pr(Dropped Out or Discharged by Inst. Dir.)*



Notes: Panels A and B display raw means of stops and felonies per 1 mile^2/year in the home Census tract of students attending a given school. We report means by school quartile, where quartiles are generated based on mean stops per 1^2 mile experienced in student home Census tracts during school years 2006-2007 through 2008-2009. Panel C graphs the mean likelihood of enrolled students being discharged from high school by institutional directive during expected grades 9 through 12. Panel D graphs the mean likelihood of enrolled students dropping out or being discharged by institutional directive.

Figure 1.6 depicts observational trends in stops and schooling outcomes by school quartile. Panel A shows that after the *Floyd* decision, the average student attending a fourth (third) quartile school is exposed to 625 (425) fewer stops per year within their home tract, which spans an average area of 0.13 square miles or about four-by-four blocks. In contrast, students attending below-median schools experience a reduction of about 250 stops per year in their home tracts. About 20-25% of these stops would have involved 14 to 18 year-old residents. Panel D confirms that, within each school quartile, felony crimes per square mile remained relatively constant throughout the study period.

Panels E and F display raw outcome means by school quartile. Simple time-series breaks observed in Panel E suggest that the probability a student was discharged by institutional directive declined by 0.15 percentage points per year (or 60% of the pre-period mean) for students attending fourth-quartile schools, and declined by 0.10 percentage points per year (or 75%) for students attending third-quartile schools. Smaller declines are also observable for bottom-half schools, although trends in these schools are considerably noiser. Panel F displays raw changes in dropout rates. While students in third and fourth quartile schools experience sharp declines in 2014, increases in 2013 across all school quartiles complicate a simple time-series analysis. We resolve this issue by implementing a difference-in-differences strategy that compares outcome changes in fourth and third quartile schools to outcome changes in below-median “control" schools. This approach underestimates treatment effects since effects on students in control schools are differenced out.

1.5.1 Empirical Strategy

We estimate the contemporaneous impacts of *Floyd* on dropouts and institutional discharges using a differences-in-differences design. Recall that our outcomes of interest equal one if and only if a student displays a change to their enrollment status during the course of *a given academic year*. This allows us to isolate the precise timing of when school enrollment outcomes were affected by the reform. On our student panel, we run the following estimating equation:

$$Y_{isy} = \sum_{q=3}^{4} \sum_{\tau \neq 0} \beta_{q\tau} I(q)_{\tau} + \delta_s + \delta_{g \text{ x } y} + \delta_{b \text{ x } y} + \Gamma X_i' + \epsilon_{isy} \tag{1.6}$$

Y_{isy} is the outcome indicator for individual i attending school s in year y. A student's expected grade level is denoted by g and is measured from the student's initial enrollment in 9th grade. δ_s are school fixed effects and $\delta_{g \text{ x } y}$ are expected-grade-by-year fixed effects, which control for changes in outcomes as students progress through high school. $\delta_{b \text{ x } y}$ are borough-by-year fixed effects that restricts comparisons to be between students attending schools in the same borough. X_i' is a vector of time-invariant student characteristics that include quintiles of age upon entering 9th grade, Black and Hispanic race indicators, a male indicator, and separate indicators for whether the student speaks Spanish or another non-English language at home. $I(q)_{\tau}$ are relative time to treatment indicators, which equal one for students attending schools in stop quartile q if year y is τ years from the onset of *Floyd*. We set the reference group as schools with below-median exposure to stops during the training period. The coefficients of interest $\{\beta_{q,t}\}$ then represent the average change between year τ and the year prior to *Floyd* among students in quartile q schools relative to the same change over time among students in below-median schools in the same borough b. All standard errors are clustered at the school-cohort level.

To ensure that differences in estimated effects are not driven by variation in pre-treatment trends, we also report results for a linear trend-adjusted version of Equation 1.6:

$$Y_{isy} = \sum_{q=3}^{4} \sum_{\tau=1}^{5} \beta_{q\tau} I(q)_{\tau} + \sum_{q=3}^{4} [I(q) \text{ x } \tau] + \delta_s + \delta_{g \text{ x } y} + \delta_{b \text{ x } y} + \Gamma X_i' + \epsilon_{isy} \tag{1.7}$$

Equation 1.7 is equivalent to Equation 1.6 except that event-time-quartile indicators are dropped for $\tau < 1$ and interactions of school quartiles with linear event-time are added for third and fourth quartile schools (i.e. $\sum_{q=3}^{4} [I(q) \text{ x } \tau]$).

While we primarily focus on reduced-form impacts of the reform, we also estimate effects per changes in neighborhood stop rates. To this end, we instrument for neighborhood stops per year with a treatment indicator set to one in the post decision period $(2014 - 2017)$. The reference period spans 2010 to 2012. Using a two-stage least squares regression, the estimating equation is then:

$$Y_{isy} = \sum_{q=3}^{4} \gamma_s^q Stops_{iy} + \sum_{q=3}^{4} \beta_{q1} I(q)_{\tau=1} + \delta_s + \delta_{g \text{ x } y} + \delta_{b \text{ x } y} + \Gamma X_i' + \epsilon_{isy} \tag{1.8}$$

where $\{\gamma_s^q\}$ estimate the effect of changes to neighborhood stops on outcomes for students attending schools in quartile q. Adding linear pre-trend controls, the equivalent regression becomes:

$$Y_{isy} = \sum_{q=3}^{4} \gamma_s^q Stops_{iy} + \sum_{q=3}^{4} [I(q)_{\tau=1} + I(q) \text{ x } \tau] + \delta_s + \delta_{g \text{ x } y} + \delta_{b \text{ x } y} + \Gamma X_i' + \epsilon_{isy} \tag{1.9}$$

1.5.2 Results

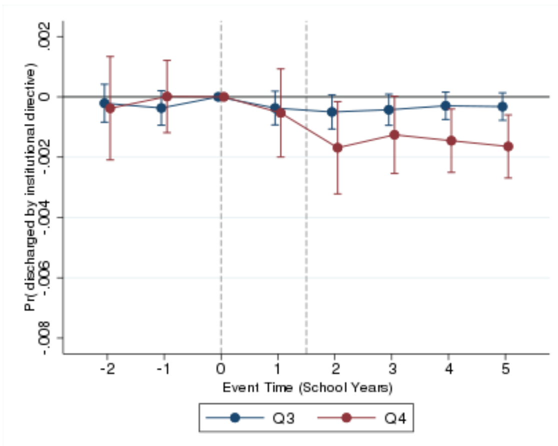
Our main findings are reported in Table 1.9. Following the *Floyd* decision, students in fourth (third) quartile schools experienced 377 (177) fewer stops per year, relative to students in control schools; 82 (36) of these prevented stops would have been of 14 to 18 year-old residents.

Figure 1.7 graphs reduced-form estimates of the impact of stop reductions on educational outcomes using Equations 1.6 and 1.7. The omitted period is the 2012 school year. For our first outcome – the likelihood that a student is discharged by institutional directive in a given school year – there is little evidence of differential group trends prior to *Floyd*, especially for fourth quartile schools. If anything, there is a slight positive pre-trend, leading Equation 1.6 to potentially underestimate treatment effects. Following the *Floyd* decision, we estimate that students in fourth quartile schools were 0.14 percentage points or 54% less likely to leave high school during a given school year by institutional directive ($p = 0.004$). In other words, the reform prevented 84 students per year from being discharged by institutional directive in fourth quartile schools. Adjusting for linear pre-trend differences across study groups increases the estimated reduction to 0.22 percentage points (85%), but this estimate is no longer statistically significant.

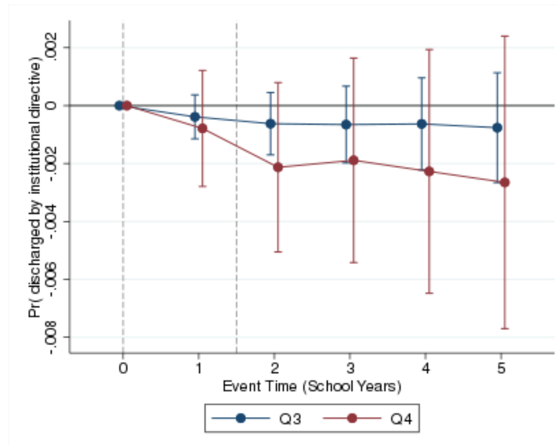
Figure 1.7: *Effect on School Discharges*

DV = Pr(Discharged by Institutional Directive)

(a) *Difference-in-differences*

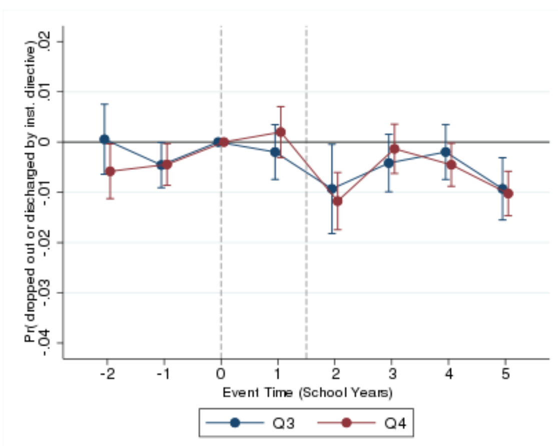


(b) + *Linear Pre-trend Controls*

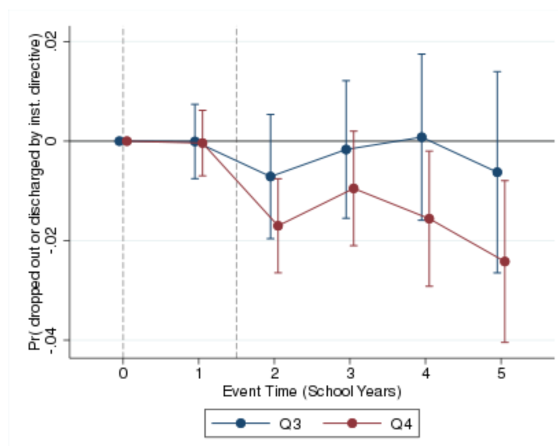


DV = Pr(Dropped Out or Discharged by Inst. Directive)

(c) *Difference-in-differences*



(d) + *Linear Pre-trend Controls*



Notes: This figure graphs coefficients from Equation 1.6 in Panels A and C and coefficients from Equation 1.7 in Panels B and D. Panels A and B display effects on the likelihood of enrolled students being discharged from high school by institutional directive during expected grades 9 through 12. Panels C and D display impacts on the likelihood of enrolled students dropping out or being discharged by institutional directive. The reference group is set to schools with below median exposure to stops during the training period. Blue dots refer to point estimates for schools in the third quartile, while maroon does denote point estimates for schools in the fourth and "most exposed" quartile. Whiskers mark 95% confidence intervals. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling. All standard errors are clustered at the school-cohort level.

Our second outcome examines the holistic impact of *Floyd* on the likelihood that a student drops out or is discharged by institutional directive. Panel C of Figure 1.7 reveals that both third and fourth quartile schools display upward pre-trends relative to control schools. Even so, our difference-in-difference approach still finds statistically significant reductions in drop out rates following the *Floyd* decision. Students from fourth (third) quartile schools are 0.36 (0.49) percentage points or 12% (16%) less likely to drop out or be discharged by institutional directive in a given school year. Panel D of Figure 1.7 shows that once we control for linear pre-trend differences, drop out rates are virtually unaffected during the scale-down year (i.e. 2013), but then fall by about 60% for students in fourth quartile schools following the *Floyd* decision and remain low for the subsequent three school years. As shown in Table 1.9, the likelihood that a student dropped out or was discharged by institutional directive fell by 1.66 percentage points per school year (58%) during the post-ruling period in fourth-quartile schools relative to control schools ($p = 0.008$). Averaging point estimates across specifications, we find that the reform prevented 660 high school students per year from dropping out or being discharged by institutional directive. Point estimates for third quartile schools when controlling for linear pre-trends are negative but considerably smaller and statistically insignificant.

To put these effects in context, Bacher-Hicks and de la Campa (2020) use a precinct commander movers design to show that students attending middle schools in precincts with an additional 350 stops per year exhibit a 0.4 percentage point increase in the likelihood of dropping out from high school. Multiplying our per-year estimate by four years of high school, the equivalent statistic for students in fourth-quartile schools is 4.0 percentage points.[36] It is not surprising that we document substantially larger treatment effects given that our approach differs in a number of key ways. First, we focus on student exposure in high school rather than during middle school. Fourteen to eighteen year-old residents are 21 times more likely to be stopped by police than eleven to thirteen year-old residents. Second, we are able to leverage larger absolute differences in stop exposure. Students

[36]This assumes an effect size of 1.01 percentage points, which is the average across our two specifications.

Table 1.9: *Effect of Floyd on High School Dropout and Institutional Discharge Rates*

	Pre-period	Base		+ Linear Pre-trends	
	Treat Mean	$\beta_{\tau\geq 2}$	P-value	$E[\beta_{\tau\in[2,5]}]$	P-value
	(1)	(2)	(3)	(4)	(5)
DV = Stops per year					
Third Quartile	459.863	-176.720***	0.000	-188.163***	0.000
		(5.191)			
Fourth Quartile	639.028	-376.543***	0.000	-475.060***	0.000
		(5.012)			
DV = Stops of 14-18 year-old residents per year					
Third Quartile	98.094	-35.675***	0.000	-47.812***	0.000
		(1.342)			
Fourth Quartile	136.129	-81.766***	0.000	-123.255***	0.000
		(1.228)			
DV = Pr(Discharged by Institutional Directive)					
Third Quartile	0.0013	-0.0002	0.261	-0.0007	0.368
		(0.0002)			
Fourth Quartile	0.0026	-0.0014***	0.004	-0.0022	0.260
		(0.0005)			
DV = Pr(Dropped Out or Discharged by Inst. Directive)					
Third Quartile	0.0315	-0.0049**	0.029	-0.0036	0.638
		(0.0022)			
Fourth Quartile	0.0289	-0.0036**	0.020	-0.0166***	0.008
		(0.0015)			

Notes: This table reports coefficients from Equations 1.6 and 1.7. Treatment is defined as schools with students from neighborhoods with stop exposure during the training period that ranked between the 50th and 75th percentile – "third quartile" – or 75-100th percentile – fourth quartile. Column 1 reports mean outcomes for each quartile during the 2010 through 2012 school years. Column 2 reports the coefficient on a post-ruling indicator that collapses 2014-2017 treatment indicators in Equation 1.6 into a singular indicator. Standard errors are reported below coefficients in parentheses. Column 4 reports the average across all four post-ruling coefficients from Equation 1.7, and Column 5 reports the p-value from an F-test that the four post-ruling coefficients are jointly zero. Standard errors are clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in fourth quartile schools experienced 377 fewer stops per home tract. Given there are approximately 28 tracts per precinct, our first stage documents changes to stop exposure that is about 30 times larger (in terms of stops per area). Finally, our approach is better able to capture neighborhood stop exposure within a student's peer network. While (Bacher-Hicks and de la Campa, 2020) measured stop exposure by summing stops over school district geographies, we find that our main effects persist even when limiting treatment variation to occur between schools *within the same school district*.[37] This highlights the importance of using detailed information on where students and their peers live to capture neighborhood stop exposure.

How do our treatment effects compare to those observed for students who are exposed to a deadly police shooting? Ang (2021a) finds that high school students living within a 0.5-mile radius of a police shooting were 1.1 percentage points less likely to graduate from high school relative to students living in the same Census block group but further from the shooting.[38] Comparatively, treatment effects observed students attending fourth-quartile schools are almost four times larger. In terms of students affected, they are about 10 times larger.

In Table 1.10, we report coefficients from our IV regressions. In fourth quartile schools, reducing average stop rates across students' home neighborhoods by 100 stops leads to a 0.037 percentage point reduction in the probability a student is institutionally discharged in a given school year ($p = 0.004$). For these same students, the probability of dropping out or being institutionally discharged falls by 0.095 to 0.22 percentage points per year ($p < 0.03$), and is 0.28 percentage points for students in third quartile schools ($p = 0.027$). Given that the net present value of graduating high school is estimated to be around $300,000, these effect sizes are economically large (Vining and Weimer, 2019). That is, our IV estimates suggest that conducting an additional 100 stops per year in a four-by-four block neighborhood

[37] Inclusion of school district fixed effects does little to affect our point estimates. These results are available upon request.

[38] Students who observed a police killing in 9th or 10th grade were the most affected – exhibiting a 1.8 percentage point decline in graduation rates.

Table 1.10: *IV Estimates of Stops on High School Dropout and Institutional Discharge Rates*

	Pre-period	Base		+ Linear Pre-trends	
	Treat Mean	100 Stops	P-value	100 Stops	P-value
	(1)	(2)	(3)	(4)	(5)
DV = Pr (Discharged by Institutional Directive) x 100					
Third Quartile	0.126	0.011	0.260	0.032	0.164
		(0.010)		(0.023)	
Fourth Quartile	0.258	0.037***	0.004	0.042	0.135
		(0.013)		(0.028)	
DV = Pr (Dropped Out or Discharged by Inst. Directive) x 100					
Third Quartile	3.147	0.275**	0.027	0.282	0.405
		(0.124)		(0.339)	
Fourth Quartile	2.893	0.095***	0.019	0.219***	0.026
		(0.041)		(0.098)	

Notes: This table estimates the impact of 100 stops per year in a student's home Census tract on high school dropout and institutional discharge rates. The outcome is in percentage points to ease interpretation. Coefficient are estimated using Equation 1.8 and 1.9. Treatment is defined as schools with students from neighborhoods with stop exposure during the training period that ranked between the 50th and 75th percentile – "third quartile" – or 75-100th percentile – fourth quartile. Column 1 reports mean outcomes for each quartile during the 2010 through 2012 school years. Column 2 instruments for post-ruling changes in stops with a singular indicator that equals one for treatment groups in school years 2014-2017. Column 4 instruments for post-ruling changes in stops with year indicators for years 2014-2017, controlling for pre-trend differences in trends. Standard errors are clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

carries a social cost of about $23 million due to spillovers on high school graduation rates alone.[39] These findings highlight just how harmful frequent stop encounters can be to the life trajectory of high school students.

Robustness

It is possible that the timing of *Floyd* coincided with other other educational reforms that affect our outcomes of interest. Most prominent of which were changes to school discipline policy. During the 2013 school year, the Department of Education removed low-level suspensions for disorderly behavior, such as profane language or persistent non-compliance. Instead of being suspended for 1 to 5 days, less disruptive punishments were adopted, such as removing students from a single class. Craig and Martin (2019) study this reform by

[39]This estimate assumes 18% of students who dropout or are discharged by institutional directive will go on to graduate in six years, and assumes a conservative effect size of 0.1575 percentage points.

comparing middle schools that prior to the reform displayed high versus low suspension rates. They find the reform increased math (reading) scores of middle school students by 0.05 (0.03) standard deviations. To test whether our findings can be explained by this suspension reform, we split schools into high (above-median) and low (below-median) suspension high schools using mean disorderly suspension rates observed from 2008 and 2009. We adapt Equations 1.6 and 1.7 by fulling interacting an "above-median" stop-exposure indicator with an above-median suspension indicator. This approach compares outcome changes in high-stop high-suspension schools, high-stop low-suspension schools, and low-stop high-suspension schools to low-stop low-suspension schools. Figure A.6 displays difference-in-differences coefficients and Table A.4 reports post-ruling coefficients.

Figure A.6 shows that reductions in suspension rates sharply change in 2013, while changes in dropouts and institutional discharges bottom out in 2014. In terms of levels, Panel B shows that high-stop high-suspension schools display the largest declines in institutional discharges, suggesting that reductions in disorderly suspensions may have complemented reduced stop rates. However, high-stop low-suspension schools also display substantial and statistically significant declines, especially when viewed as a percent of pre-reform mean; the post-ruling difference-in-differences coefficient implies a 32% decline ($p = 0.007$). What's more, drop out rates in high-stop low-suspension high schools closely track those of high-stop high-suspension schools, while low-stop high-suspension schools do not. Therefore, we argue that the simultaneous change to disorderly suspension policy may have complemented the reform but is unlikely to be the driving force behind our main findings.

A second set of disciplinary reforms was ushered in by the newly-elected Mayor Bill de Blasio and his School Chancellor Carmen Fariña. While Chancellor Fariña openly promoted less punitive discipline procedures after taking office in January 2014, it was not until February 2015 that she officially revised the school discipline code.[40] This was a full year

[40]These reforms were moderate, however, changing only the most contentious aspects of the school discipline code. Most notably, under the new policy principals needed approval from the Education Department's central office before suspending a student and the NYPD were forced to track and report any instance in which a student is handcuffed (Harris, 2015).

after dropout and institutional discharge rates had bottomed out, suggesting that these reforms do not explain our findings. More broadly, there does not appear to be sharp changes to other important school inputs. For example, Appendix Figure A.11 shows that summonses and arrests made by school safety agents steadily declined from 2011 to 2016.

We also probe robustness by exploring parallel trend violations per Roth and Rambachan (2021). Figures A.7 and A.8 show that our findings are robust to fairly large violations so long as we assume pre-trend violations introduce positive rather than negative bias. Finally, Section A.2 of the Appendix reveals that our main findings persist if we instead exploit tract-level variation in stop exposure rather than school-level variation. Students that experience larger reductions to stop rates in their home neighborhoods due to the reform also display significantly larger reductions in dropout rates, with effect sizes that are similar in magnitude to our main findings.

Heterogeneity

How do effects differ by student race and sex? It is first important to note that 92% of students attending schools in the fourth quartile are Black or Hispanic, meaning that the *Floyd* decision disproportionately benefited students of color. We formally explore whether demographic sub-populations differentially responded to the reform by running Equations 1.6 and 1.7 separately on each race, sex, and race-by-sex subgroup. To improve precision, we pool white, Asian, and other-race students into one race category.

Table 1.11: *Effect of Floyd on Institutional Discharge Rates by Student Race and Sex*

Y = Pr(Discharged by Institutional Directive)	Pre-period Treat Mean	Base $\beta_{\tau\geq 2}$	Base P-value	+ Linear Pre-trends $E[\beta_{\tau\in[2,5]}]$	+ Linear Pre-trends P-value
	(1)	(2)	(3)	(4)	(5)
Student Race					
Black	0.0036	-0.0019**	0.0155	-0.0034	0.2869
		(0.0008)			
Hispanic	0.0018	-0.0011*	0.0619	-0.0019	0.4177
		(0.0006)			
White/Asian/Other	0.0020	-0.0009	0.1235	-0.0036	0.1729
		(0.0006)			
Student Sex					
Male	0.0042	-0.0021**	0.0101	-0.0035	0.3354
		(0.0008)			
Female	0.0010	-0.0007***	0.0020	-0.0011	0.3549
		(0.0002)			
Student Race x Sex					
Black Male	0.0059	-0.0027*	0.0537	-0.0054	0.4294
		(0.0014)			
Black Female	0.0013	-0.0011**	0.0020	-0.0019	0.2979
		(0.0004)			
Hispanic Male	0.0028	-0.0018*	0.0819	-0.0040	0.2773
		(0.0010)			
Hispanic Female	0.0008	-0.0005	0.1133	0.0000	0.9989
		(0.0003)			
White/Asian/Other Male	0.0029	-0.0009	0.2670	-0.0038	0.3836
		(0.0008)			
White/Asian/Other Female	0.0009	-0.0007	0.2345	-0.0027	0.2065
		(0.0006)			

Notes: This table reports post-ruling coefficients for fourth quartile schools from Equations 1.6 and 1.7 ran separately restricting the data to each sub-group. The outcome of interest is an indicator for whether the student switched from being enrolled at the beginning of the school year to being discharged by institutional directive. Column 1 reports mean outcomes for each sub-group group during the 2010 through 2012 school years. Column 2 reports race-specific post-ruling impacts (i.e. years 2014 through 2017) using Equation 1.6. Standard errors are reported below coefficients in parentheses. Column 4 reports the average across all four post-ruling coefficients from Equation 1.7, and Column 5 reports the p-value from an F-test that the four post-ruling coefficients are jointly zero. Standard errors are clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.12: *Effect of Floyd on Dropout Rates by Student Race and Sex*

Y = Pr(Dropped Out or Discharged by Inst. Dir.)	Pre-period Treat Mean	Base		+ Linear Pre-trends	
		$\beta_{\tau\geq 2}$	P-value	$E[\beta_{\tau\in[2,5]}]$	P-value
	(1)	(2)	(3)	(4)	(5)
Student Race					
Black	0.0290	-0.0095***	0.0007	-0.0157	0.1338
		(0.0028)			
Hispanic	0.0297	-0.0021	0.3664	-0.0143	0.1002
		(0.0023)			
White/Asian/Other	0.0237	-0.0021	0.2919	-0.0194**	0.0397
		(0.0020)			
Student Sex					
Male	0.0326	-0.0035*	0.0672	-0.0196**	0.0143
		(0.0019)			
Female	0.0253	-0.0036**	0.0133	-0.0139**	0.0222
		(0.0015)			
Student Race x Sex					
Black Male	0.0335	-0.0114***	0.0016	-0.0161	0.2608
		(0.0036)			
Black Female	0.0247	-0.0081***	0.0021	-0.0149	0.1289
		(0.0026)			
Hispanic Male	0.0329	-0.0022	0.4513	-0.0221*	0.0518
		(0.0029)			
Hispanic Female	0.0265	-0.0012	0.5972	-0.0075	0.4182
		(0.0023)			
White/Asian/Other Male	0.0265	-0.0021	0.4291	-0.0205	0.1111
		(0.0027)			
White/Asian/Other Female	0.0203	-0.0017	0.4998	-0.0189	0.1084
		(0.0025)			

Notes: This table reports post-ruling coefficients for fourth quartile schools from Equations 1.6 and 1.7 ran separately restricting the data to each sub-group. The outcome of interest is an indicator for whether the student switched from being enrolled at the beginning of the school year to being listed as dropped out or discharged by institutional directive. Column 1 reports mean outcomes for each sub-group group during the 2010 through 2012 school years. Column 2 reports race-specific post-ruling impacts (i.e. years 2014 through 2017) using Equation 1.6. Standard errors are reported below coefficients in parentheses. Column 4 reports the average across all four post-ruling coefficients from Equation 1.7, and Column 5 reports the p-value from an F-test that the four post-ruling coefficients are jointly zero. Standard errors are clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Results on the likelihood a student is discharged by institutional directive are reported in Table 1.11 and results for the likelihood a student drops out or is discharged by institutional directive are reported in Table 1.12. We only report coefficients on fourth-quartile school indicators, so as to reduce the number of hypotheses tested. Point estimates across both specifications suggest that impacts on discharges by institutional directive are slightly larger for black students, although differences between estimated coefficients are not statistically significant. If we average post-ruling coefficients across the two specifications, we find that the yearly likelihood of dropping out or being discharged by institutional directive falls by 1.1 percentage points (38%) per year for Black students, 0.7 percentage points (25%) per year for Hispanic students, and 0.9 percentage points (38%) per year for white/Asian/other-race students. As one might expect, these tables also reveal that male students benefit more from the *Floyd* decision than female students, especially with regard to institutional discharge rates. Black male students in particular display the largest absolute reductions. Scaling estimates by the number of students, we estimate that the reform prevented about 50 Black male students per year from becoming involved with the carceral state.

1.6 Conclusion

Through the lens of the *Floyd* reform, this paper examines the effectiveness of using police stops as a crime deterrence tool. We find that sizeable reductions in pedestrian stops have no impact on serious crimes, and *decrease* the detection of minor offenses. These results contrast the effectiveness of police surges that simultaneously increase officer presence and stop rates in crime hot-spots. Together, our results suggest that *patrol officer presence but not the concentration of pedestrian stops* matter for deterring serious neighborhood crime.

The concentration of stops in higher-crime neighborhoods leads to large racial disparities in stop rates, where young men of color are stopped at alarming rates. These excessive stop rates spillover into the classrooms of neighborhood high school students. We estimate that the reform prevented about 660 students per year from dropping out or being discharged by institutional directive, which has a net present value over $200 million per year.[41] Over ninety percent of these teenagers were Black or Hispanic. Given that our approach differences out effects for students attending less-exposed schools and that we do not account for educational effects on younger children or labor market impacts on young adults, the net social benefit of the reform is likely much larger.

This chapter highlights the need for careful analysis when examining the social impact of police practices. Much of the social value of policing is not directly observed. Instead, accurate estimation of social value requires causal estimation of first-order behavioral responses, such as crime and educational responses. The first two chapters of this dissertation additionally illustrate the outsize role the police play in urban communities of color, and especially, for young men of color. Future research should examine the downstream consequences of stops and other lower-level police encounters that have the potential to generate long-run social costs by disrupting the life trajectory of young people. By estimating these social costs, future research could inform better police practices that reduce existing racial and place-based opportunity gaps.

[41] We arrive at this figure by multiplying the total effect per year by the average number of students enrolled in fourth-quartile schools during the post-ruling period, the fraction of dropouts who do not go on to graduate within six years (82%), and the net present value of graduating from high school (Vining and Weimer, 2019).

Chapter 2

Measuring Racial Bias in Police Stopping Decisions[1]

2.1 Introduction

Over 3.5 million pedestrians are stopped by police in the United States every year (Harrell and Davis, 2020). While the primary goal of pedestrian stops is to protect civilians from imminent harm, police departments often concentrate stops in higher-crime neighborhoods with the aim of deterring serious crime. Chapter 1 revealed that excessively high stop rates are ineffective at deterring serious crime and impose substantial costs on Black and Hispanic high school students who live in higher-crime neighborhoods. Pedestrian stops may also impose additional costs on local communities, as they can be traumatizing (Geller *et al.*, 2014; Boyd, 2018), breed institutional distrust (Kirk and Papachristos, 2011a), disrupt educational investments (Legewie and Fagan, 2019a; Bacher-Hicks and de la Campa, 2020), and harm

[1]Co-authored with Jeffrey Fagan. Dr. Fagan was instrumental in securing and understanding NYPD policing data. We also thank Edward Glaeser, Lawrence Katz, Joscha Legewie, and Andrei Shleifer for unrelenting advice and encouragement on this project. We are grateful for helpful comments from Desmond Ang, Noam Angrist, Caroline Chin, Natalia Emanuel, Paul Fontanier, Benny Goldman, Dev Patel, Robert Fluegge, and seminar participants at Harvard Public/Labor lunches and seminars. The NYU Research Alliance of NYC Schools provided the education data for this project. Tebes also benefited from generous financial support from the NSF Graduate Research Fellowship, the Stone Scholar Fellowship for Social Policy and Inequality Research, and the Horowitz Foundation. All errors are our own.

labor market opportunities. Stops also mechanically increase the likelihood of more serious downstream police actions, such as uses of force or arrests for minor offenses (Knox and Mummolo, 2020).

While Black and Hispanic pedestrians comprise 83% of stops in New York City, only 51% of city residents are Black or Hispanic. The concentration of stops in higher-crime neighborhoods contributes to these racial disparities – residents of higher-crime neighborhoods are stopped at five times the rate of residents in lower-crime neighborhoods.[2] In fact, assigning the residential racial composition of a Census tract to stops conducted in that tract accounts for over half of the observed racial difference in stop rates. Young men of color, in particular, are disproportionately subjected to stops. We estimate that, on average, a Black male teenager living in a higher-crime neighborhood would be stopped 30 times over the course of his high school tenure. These stops rarely lead to the detection of criminal behavior – only 6% result in an arrest and 1 in 770 stops discover an illegal firearm. Nonetheless, such racial disparities could arise as the byproduct of officers stopping pedestrians who are more likely to be committing crime.

This chapter explores whether these racial disparities reflect true racial differences in criminal behavior or are the result of unfair targeting by patrol officers. To answer this question, we compare the *accuracy* of stops that were deterred by the 2012 reform across racial groups as a test for racial bias. Our main stop outcome is whether or not the stop resulted in an arrest or the discovery of a weapon or drugs, although we find similar results using other stop outcomes. Our analysis is motivated by a simple model of officer stopping decisions where statistically-rational officers face diminishing marginal returns to stops.[3] Officers equate marginal returns across racial groups so as to maximize overall returns to stops. Under standard assumptions, racial bias is then identified by racial differences in outcomes of marginal stops (Becker, 1957).

[2] See Figure 1.1

[3] Specifically, adapt Anwar and Fang's model of statistical discrimination (2006). See also Becker (1957, 1993); Knowles and Todd (2001); Feigenberg and Miller (2021); Fryer Jr (2019a); Abrams *et al.* (2021).

When taking this model to the data, directly regressing stop outcomes on race indicators and a vector of controls generates two sources of bias: omitted variables bias and *infra-marginality bias*. The latter occurs when racial differences in outcomes of always-stopped pedestrians cause average racial differences to differ from marginal differences (Ayres, 2002). Using the reform as an instrument for neighborhood stop rates in the classic framework of Angrist *et al.* (1996), one can consistently estimate the likelihood that marginally-deterred stops (i.e. stops that would have occurred prior to the reform but do not occur post-reform) would have resulted in the detection of criminal behavior.[4]

For each race, we trace out the marginal return curve during the post-reform period by estimating the mean hit rate of stops that were additionally deterred in each year of the post-reform period. Curves are upward-sloping and convex when graphed against reductions in stop rates, validating the notion that officers face diminishing marginal returns when conducting stops. Racial differences in returns to marginally-deterred stops in the first year of the reform provide an estimate of racial bias in pre-reform stop rates. These estimates are substantially larger than OLS estimates – hit rates of marginally-deterred Black and Hispanic stops are 24% and 34% smaller than white estimates, while those same differences using OLS are 7% and 6%, respectively. Black-white and Hispanic-white pairwise differences are statistically significant at the 1% level using both conventional and permuted p-values. Across the entire post-reform period, we find that deterred Black stops yield a hit 8.2% of the time, compared to 10.0% of the time for deterred white stops and 7.9% of the time for deterred Hispanic stops.[5]

Our marginal return curves also allow us to estimate the point at which marginal Black and Hispanic returns equal the initial return estimated for marginally-deterred white

[4]This approach critically assumes that pedestrians do not respond to the reform by increasing their criminal behavior. Although such responses are conceivable, we provide three pieces of evidence to the contrary: (1) there is a lack of aggregate changes in crime levels surrounding the reform; (2) the reform has no impact on serious crimes and only mechanically affects the detection of minor crimes; and, (3) Black-white and Hispanic-white arrest ratios are relatively flat throughout the study period, suggesting racial groups do not differentially respond to the reform.

[5]These Black-white and Hispanic-white pairwise differences are statistically significant at the 0.1% level using both conventional and permuted p-values.

stops. Doing so reveals that Black and Hispanic stops had to be reduced by an additional 55% to reach initial white estimates. Put differently, in the first year of the reform, after aggregate stop rates had already fallen by about 35%, the police still over-stopped Black and Hispanic pedestrians by a combined *298,000 stops per year*. Since our approach allows officers to statistically discriminate or “rationally" racial profile, we interpret these results as a *conservative estimate* of racial discrimination (Arnold and Hull, 2020).

This chapter contributes to research adapting the Becker outcome test to investigate racial bias.[6] Papers examining racial bias in policing have primarily focused on vehicular stops (Knowles and Todd, 2001; Anwar and Fang, 2006; Feigenberg and Miller, 2021) and, more recently, uses of force (Fryer Jr, 2019a; Hoekstra and Sloan, 2021). Little attention has been paid to racial bias in pedestrian stopping decisions, even though bias at this decision point influences bias estimates of many “downstream" policing decisions.[7] Our approach also adds to an emerging literature that uses the Angrist *et al.* (1996) framework to estimate race-specific marginal hit rates as a test for racial bias (Arnold and Hull, 2020; Marx, 2018; Rose, 2021). By tracing out race-specific marginal return curves, we are able to estimate the extent to which the police over-stopped minority pedestrians by finding the point at which marginal returns are equated across racial groups.

More broadly, the first two chapters of this dissertation shed light on the mechanisms underlying the impact of neighborhoods on social mobility (Ludwig *et al.*, 2013; Chetty and Katz, 2016; Chetty *et al.*, 2020). Chetty *et al.* (2020) show that, conditional on parental income, Black boys have lower incomes in adulthood in 99% of Census tracts. Not only do we find that Black male teenagers were substantially over-stopped by police, but also that these interactions disproportionately translated into high school dropouts and interactions with the carceral state. Combined with previous studies showing the indirect harm of police killings (Ang, 2021a), ticketing (Mello, 2019; Goncalves and Mello, 2021), and prosecution of

[6]See, for example: Becker (1957, 1993); Arnold and Hull (2020); Arnold *et al.* (2018); Knowles and Todd (2001); Feigenberg and Miller (2021); Fryer Jr (2019a); Rose (2021).

[7]For example, estimates of racial bias in police use of force typically condition on a stop occurring and thus will tend to understate racial bias if there exists bias in stopping decisions (Knox and Mummolo, 2020).

low-level arrests (Agan *et al.*, 2021), it becomes clear that the police play a pivotal role in the life trajectory of young men of color. Together, Chapters 1 and 2 provide actionable policy advice about one reform that could help close place-based racial opportunity gaps.

The remainder of this chapter proceeds as follows. Section 2.2 describes the model of officer stopping decisions that motivates our test for racial bias. Section 2.3 explains common statistical biases that arise from a direct regression approach and the instrumental variables framework that we employ to circumvent these issues. Section 2.4 describes our main results and potential mechanisms, while Section 2.5 concludes.

2.2 Theoretical Framework

Officers observe pedestrian race and a noisy signal of criminal activity, θ, such that the probability of detecting criminal activity is monotonically increasing in θ.[8] This motivates a simple decision rule for officers, where they stop pedestrians of a given race R if and only if $\theta \geq \theta_R^*$. At the margin, officers are indifferent between stopping an additional pedestrian of either race:

$$p_W(\theta_W^*) - C_W = p_M(\theta_M^*) - C_M \tag{2.1}$$

where, for simplicity, we consider white (W) and minority (M) pedestrians.[9] $p_R(\theta)$ denotes the probability of detecting criminal behavior for a pedestrian of race R with signal θ and C_R is the marginal cost of conducting a stop of pedestrians of race R. Following Becker (1957, 1993), we define racial bias against minority pedestrians as the marginal cost of stopping a minority pedestrian is lower than the marginal cost of stopping a white pedestrian, $C_M < C_W$.

Equation 2.1 implies the standard outcomes test result: racial bias can be detected by

[8]Formally, this requires the standard monotone likelihood property assumption that implies a higher θ informs the officer that the pedestrian is more likely to be committing a crime. See Section B.1 of the Appendix for a more detailed discussion of our motivating model.

[9]This framework can be easily extended to include multiple racial groups. For ease of exposition, we will only consider two racial groups, where one should think of minority pedestrians as referring to Black and Hispanic pedestrians.

estimating *racial differences in the outcomes of marginal stops*. If an officer is racially prejudiced against minorities (i.e. $C_M < C_W$), then he should be more likely to detect criminal behavior during a stop of a marginal white pedestrian than during a stop of a marginal minority pedestrian:

$$p_W(\theta_W^*) > p_M(\theta_M^*) \tag{2.2}$$

Absent racial bias, marginal detection rates or "hit" rates should be equal.[10]

This approach allows for statistical discrimination – stop rates of minority pedestrians may be higher than white pedestrians due to true racial differences in criminal behavior (e.g., see Arrow (1973), Phelps (1972), and Coate and Loury (1993)). However, the legality of statistical discrimination remains controversial.[11] An alternative definition, for example, could define discrimination as the average disparity in stop outcomes between white and minority individuals with identical criminal propensities (Arnold and Hull, 2020). Instead, our definition compares officers' observed actions to a counterfactual officer who, given the same information set, acts as a statistically rational decision-maker when deciding whether or not to conduct a potential stop. To this end, our definition provides a *conservative benchmark* and acknowledges the fact that, in practice, race is easily observed by officers and may affect their decision-making if it serves as an informative signal of pedestrian criminality. As discussed in Feigenberg and Miller (2021), this approach is equivalent to asking whether closing racial disparities in stop rates poses an equity-efficiency trade-off for patrol officers.

---

[10]Arnold *et al.* (2018) demonstrate that this model is isomorphic to one where officers do not hold a "taste" for discrimination but rather are racially biased in their perceptions of risk, i.e. $p_M(\theta) = E[Y|\theta, R = M] + \tau_M(\theta)$, where $\tau_M(\theta) > 0$ is racially-biased prediction error.

[11]For example, constitutional challenges to racial profiling have historically been unsuccessful, since plaintiffs had to prove "discriminatory purpose". (Feigenberg and Miller, 2021) Despite this, the Department of Justice has filed a number of cases against police departments for racially-targeted stops of pedestrians and motorists under this basis, often pointing to significant city-wide racial disparities in stop rates and particularly alarming cases of misused force or illegal searches (Anderson, 2020). More recently, Tiwara (2019) has argued that disparate impact liability under the Omnibus Crime Control and Safe Streets Act of 1968 challenges the legality of racial profiling. Under this criteria, a police practice that has a disparate impact on ethnic and racial minorities is permissible if and only if "the police can demonstrate that it has a legitimate law-enforcement-related necessity" (Tiwara, 2019; Feigenberg and Miller, 2021). Our analysis of crime effects is an example of one such analysis.

2.3 Empirical Strategy

2.3.1 Identification Issues with OLS Approach

We are interested in comparing the *accuracy* of marginal stops across racial groups. A common approach in the literature is to directly regress stop outcomes on race indicators, controlling for as many observable characteristics as possible.[12] Consider the regression

$$Y_i = \beta_B Black_i + \beta_H Hispanic_i + \beta_A Asian_i + \beta_O Other + \beta_X X_i' + \epsilon_i \tag{2.3}$$

where Y_i is the outcome of a stop of pedestrian i and X_i' is a vector of situational characteristics recorded in police records and available to researchers.

Using data on stops conducted prior to *Floyd*, we report coefficients from Equation 2.3 in Table 2.1. Our primary outcome is an indicator for whether a stop led to an arrest or the discovery of drugs or a weapon (hereafter "hit").[13] We also report impacts on each stop outcome individually, as well as hit rates of pedestrian frisks and searches.[14] Overall, hits are rather rare – only 7.3% of White stops result in a hit. Column 1 shows that Black and Hispanic stops are, respectively, 13% and 10% less likely to yield a hit. Controlling for pedestrian age quintiles, pedestrian sex, time-of-day, and day-of-week shrinks racial differences by 30-40%. Adding precinct fixed effects shrinks differences by an additional 10-20%, while adding tract and officer fixed effects have relatively no impact. With the full set of controls, Column 5 reports Black and Hispanic stops are, respectively, 7.0% and 5.5% less likely to result in a hit ($p < 0.01$).

This direct regression approach suffers from two potential biases. First, absent control variables, race coefficients capture racial differences in stop outcomes between the average pedestrian of each group. Racial bias, however, is defined at the margin and average

[12] See, for example, Fryer Jr (2019a); Knowles and Todd (2001); Anwar and Fang (2006).

[13] While our main analysis will focus on the 93% of stops involving white, Black or Hispanic pedestrians, it is important to note that hit rates are also lower for other racial minority groups. For example, the largest disparity occurs for the 3.5% of stops where race is listed as "other" or is missing.

[14] See Table 2.3 for IV estimates and Table B.1 for OLS results.

Table 2.1: *Racial Disparities in Pre-Floyd Stop Outcomes*

Y = Arrest or Weapon/Drugs Found	(1)	(2)	(3)	(4)	(5)
Black	-0.010***	-0.007***	-0.005***	-0.005***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Hispanic	-0.007***	-0.004***	-0.005***	-0.005***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Asian, Pacific Islander,	-0.008***	-0.007***	-0.011***	-0.010***	-0.007***
or Native American	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Other or Missing Race	-0.022***	-0.016***	-0.017***	-0.017***	-0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
White Mean	0.0733	0.0733	0.0733	0.0733	0.0733
Observations	3,502,679	3,423,793	3,423,793	3,423,793	3,421,516
R-squared	0.000	0.003	0.011	0.017	0.098
Specification	None	+X′	+Pct FE	+Tract FE	+Off FE

Notes: This table reports race coefficients from Equation 2.3. The reference group are white non-Hispanic pedestrians who are stopped by police. "Other or missing race" cover 3.52% of all stops and include stops where police reported race as "other" or did not provide racial information. Column (1) reports coefficients without any situational controls. Column (2) adds vector X' which includes a male indicator, age quintiles, day-of-week indicators, and eight three-hour time-of-day indicators. Columns (3), (4), and (5) sequentially add precinct, tract, and officer fixed effects. Standard errors are clustered at the date-time-location level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

racial differences may not reflect marginal differences. Termed "inframarginality" bias, this bias will arise whenever always-stopped pedestrians differ by race in their likelihood of criminal behavior. In New York City, Black and Hispanic civilians are disproportionately represented in the tail of the criminal risk distribution, potentially biasing OLS estimates against detecting discrimination against minority pedestrians. Second, we do not have an instrument for race nor does the the control vector – X_i' – fully capture situational factors that enter the officer's objective function. This generates omitted variable bias, where the direction bias is unclear. We address these concerns by implementing an instrumental variables strategy that compares mean outcomes of stops deterred by the *Floyd* reform across racial groups as a test for racial bias.

2.3.2 Instrumental Variables Approach

Consider our main dataset that contains stop data collapsed to the tract-week level. Let Y_n represent the number of successful stops per week in neighborhood n (i.e. those that lead to the detection of criminal activity). Consider the binary instrument, Z_n, which is equal to 1 in the post-*Floyd* period and 0 otherwise. Finally, let Z_n index the number of potential stops, $S_n(0), S_n(1)$, and the number of potential successful stops, $Y_n(0), Y_n(1)$, per week in neighborhood n.

We then assume the following assumptions are satisfied a la Angrist *et al.* (1996):

1. **Independence and exclusion:** $(Y_n(0), Y_n(1), S_n(0), S_n(1)) \perp\!\!\!\perp Z_n$
2. **First stage:** $E(S_n|Z_n = 0) > E(S_n|Z_n = 1)$
3. **Monotonicity:** $S_n(0) \geq S_n(1)$

In words, these assumptions imply that Z_n is independent of potential stop outcomes and stop rates, only affects stop outcomes by decreasing the number of stops conducted, and weakly lowers the number of stops conducted in all neighborhoods.

To satisfy the exclusion restriction, it must be that the entire change in stop outcomes is attributable to changes in stop rates. Imposing this restriction rules out that pedestrians alter their behavior in response to changes in stop rates. These responses are definitely possible. For example, pedestrians may increase their tendency to carry drugs or weapons when officers reduce their stop rates. Although untestable, we provide numerous pieces of evidence that support this assumption. First, Panel A of Figure B.6 shows that the level of reported crimes across all neighborhoods remained relatively flat over the study period we examine. Second, Section 1.4 revealed that *Floyd* did not deter serious crime – shootings, murders, and major felonies were unaffected, while minor misdemeanors and violations fell mechanically due to decreased stop-related detection. Third, we use arrest data that contain complete race information to compare how Black-white and Hispanic-white arrest ratios vary over the study period for a variety of offense categories. Panels B and C of Figure B.6 demonstrates that Black-white and Hispanic-white arrest ratios are relatively flat throughout the study period. If anything, these ratios are slightly downward sloping, which would lead our approach to understate racial discrimination.[15] In order to compare mean outcomes of deterred stops as a test for racial bias, behavioral responses among infra-marginal pedestrians must be the same across racial groups. Note that the exclusion restriction is sufficient for this to be true (since zero behavioral responses implies responses are the same).

The plausibility of our monotonicity assumption rests on the notion that post-*Floyd*, officers in each neighborhood n were (weakly) less likely to conduct a given stop relative to prior to *Floyd*. Panel A of Figure B.5 reveals that stop rates fell significantly within every neighborhood and Panel B shows that stop rates also fell for over 99% of all officers.[16] Finally, Figure B.4 shows that stop rates declined by 70-80% across all sex-by-race-by-age sub-groups over the three years post-*Floyd*.

Under these assumptions, it is possible to causally estimate the *accuracy* of deterred stops for racial group r:

$$E[Y_n(0)|S_n(0) > S_n(1), R_n = r] = \frac{E[Y_n|Z_n = 0, R_n = r] - E[Y_n|Z_n = 1, R_n = r]}{E[S_n|Z_n = 0, R_n = r] - E[S_n|Z_n = 1, R_n = r]} \tag{2.4}$$

This quantity measures the share of pedestrian stops of racial group r that were conducted when $Z_n = 0$ that would not have been conducted when $Z_n = 1$ that led to the detection of criminal activity. In other words, this approach estimates the *local average treatment effect* (LATE) of the reform on outcomes of stops deterred by the reform and provides no information on infra-marginal stops.

Empirically, we restrict the sample to the two years before and three years after the reform, since reported crime rates are constant during this period.[17] For each racial group r,

[15] In essence, downward-sloping arrest ratios suggest that we are under-estimating the hit rate of Black/Hispanic inframarginal stops in the pre-period by using post period hit rates relative to White inframarginal stops.

[16] By year three, over 99% of officers who conducted more than 10 stops in the year prior to *Floyd* conducted fewer stops, compared to just 64% of officers during our pre-*Floyd* placebo period.

[17] Specifically, this period spans the 5 x 52 weeks between April 15th, 2010 to April 15th, 2015.

we estimate the following two-stage least-squares (2SLS) regression

$$Y_{nt} = \beta_S S_{nt} + \delta_n + \delta_{p \times w} + \epsilon_{nt} \tag{2.5}$$

where we instrument for weekly stops ($S_{n,t}$) with an indicator denoting the post-*Floyd* period (Z_t). δ_n and $\delta_{p \times w}$ are neighborhood and precinct by week-of-year fixed effects. Then, β_S provides an estimate of the mean hit rate of stops that would have occurred before *Floyd* but did not occur after *Floyd*. We are also interested in tracing out the marginal return curve to examine how returns to stops evolve for each race as the police scaled back stops in response to the reform. To do so, we run three separate 2SLS regressions that estimate the mean hit rate of deterred stops in each post-*Floyd* year relative to the year before.[18] That is, the estimating equation is:

$$Y_{nt} = \pi_S^{\tau} S(Z = I(\tau)) + \sum_{t \neq \tau, \tau-1} \delta_{\tau} I(t) + \delta_n + \delta_{p \times w} + \epsilon_{nt} \tag{2.6}$$

where $\sum_{t \neq \tau, \tau-1} \delta_{\tau} I(t = \tau)$ are indicators containing all years other than the reference year and the post-*Floyd* treatment year of interest.

2.4 Racial Disparities in Outcomes of Deterred Stops

2.4.1 Main Results

Figure 2.1 displays results from Equation 2.6 and traces out the marginal return curve for each racial group as the police scaled back stops in response to *Floyd*. For each race, we display three IV estimates that sequentially estimate the mean hit rate of stops that were *additionally deterred* in each year of the post-period.[19] The x-axis reports the average percent reduction in stops conducted for a given post-period year relative to the two years prior to the reform. Racial differences in IV estimates from the first year of the reform provide an

[18] We break the post-period into yearly increments so that we can include week-of-year fixed effects that control for seasonal variation in the productivity of stops.

[19] Note that in order to properly control for seasonal variation in stop returns, we cannot partition the post-period into smaller increments than years. Doing so would require us to drop week-of-year fixed effects.

estimate of initial disparities in marginal returns. We find that deterred Black stops yield an arrest or the discovery of a weapon or drugs 6.7% of the time, compared to 8.8% of the time for deterred white stops, and 5.8% of the time for deterred Hispanic stops. White-Black and white-Hispanic pairwise comparisons are statistically significant at about the 1% level using conventional p-values and at the 0.1% level using permuted p-values (see Table 2.2).[20] These disparities are substantially larger than OLS estimates. Respectively, Black and Hispanic estimates are 24% and 34% smaller than white estimates, while those same differences using OLS with the full set of controls are 7% and 6%. This likely reflects the fact that the hit rate of always-stopped Black and Hispanic pedestrians is higher than their white counterparts.

For each racial group, the marginal return curve is upward-sloping and convex, validating the notion that officers face diminishing marginal returns when conducting stops (Abrams *et al.*, 2021). However, the flatness of these curves in the first two years after the reform is somewhat surprising, given that stop rates fell by 85% over this period. The curve is particularly flat for Black pedestrians, suggesting that, for the bottom 85% of Black stops, pedestrians display a sufficiently low risk signal ($\theta_B \leq \bar{\theta}_B$) such that officers are unable to predict stop outcomes. This result accords with the modified statistical discrimination model put forward by Feigenberg and Miller (2021), who find that highway patrol officers face constant returns to searching vehicles of sufficiently high-risk motorists. While constant returns imply there is no equity-efficiency trade-off to closing racial disparities in their setting, our analysis identifies racial disparities in marginal returns that suggest reducing racial disparities would generate efficiency gains for officers.

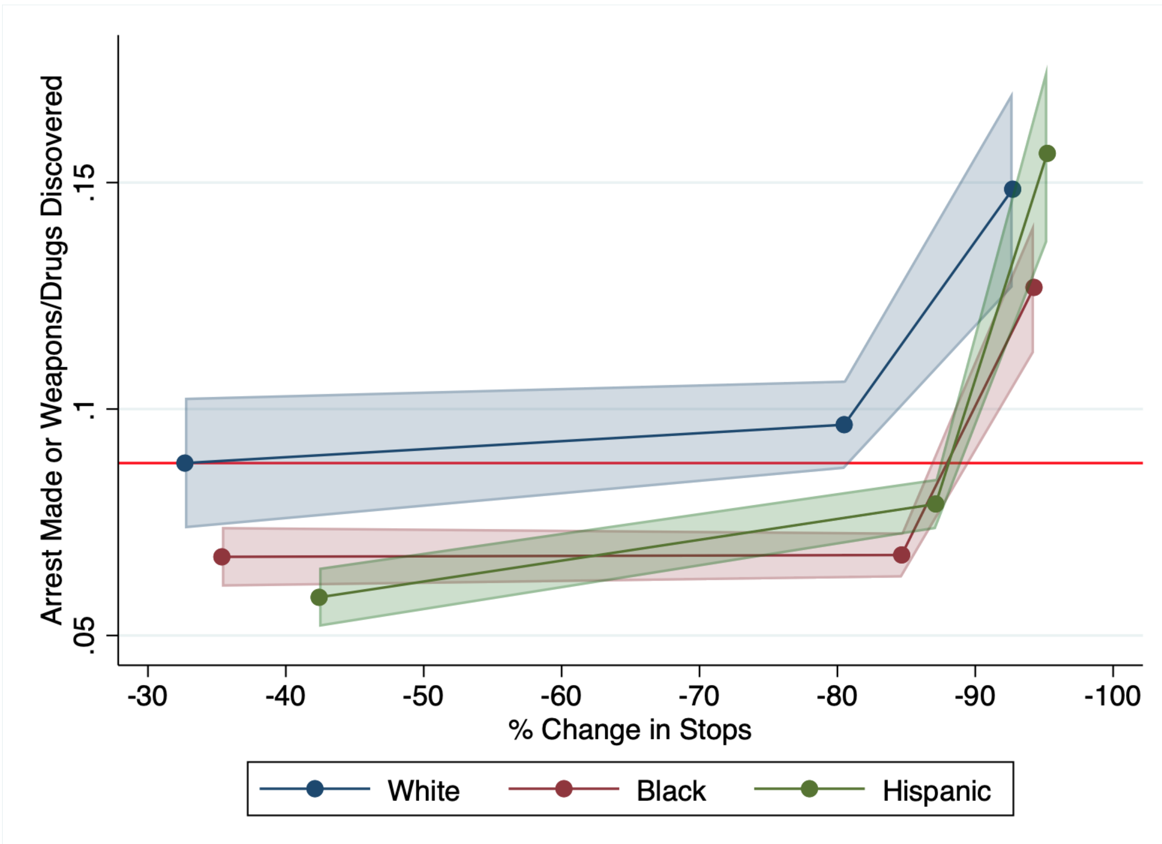
We can estimate how much the police would have to reduce stop rates of Black and Hispanic pedestrians in order to equate marginal returns across races. The red line in Figure 2.1 denotes the mean outcome of white stops that were deterred by the reform in the first

[20]Drawing on the fact that regression coefficients are normally distributed in large samples, we first compute a conventional two-sided Z-test. The white-Black Z-score, for example, is computed by $Z_{WB} = \frac{\beta_W - \beta_B}{\sqrt{(SE^2_{\beta_W} + SE^2_{\beta_B})}}$. The p-value tests the null that $Z_{WB} = 0$ using a two-sided Z-test. Our second and preferred approach permutes p-values by randomly re-assigning race and re-running our main specification 1,000 times. Permuted p-values are given by the fraction of iterations where the absolute value of simulated Z-scores is weakly greater than the absolute value of our observed Z-score.

year after the reform. Where the red line intersects the marginal return curve marks the point at which estimated marginal returns are equated across racial groups. Noting that the reference estimate is computed from a 33% reduction in white stops, we find that for marginal returns to be equated across races, stops of both Black and Hispanic pedestrians would have to be reduced by an additional 55%. Put differently, in the first year of the reform, after aggregate stops had already fallen by 35%, the police still over-stopped Black and Hispanic pedestrians by a combined 298,000 stops per year.

Using Equation 2.5, we estimate the mean return to all stops deterred by *Floyd* for each racial group. We find that deterred Black stops, on average, yield a hit 8.2% of the time, compared to 10.0% of the time for deterred white stops, and 7.9% of the time for deterred Hispanic stops. Conceptually, one should think of these estimates as a weighted average of marginal racial differences across the interval of potential stops affected by the reform. Table 2.2 shows that white-Black and white-Hispanic differences are statistically significant at the 0.1% level using both conventional and permuted p-values. Black-Hispanic differences are statistically significant at the 1% level using permuted p-values but are not distinguishable at the 10% level using the standard approach.

Figure 2.1: *Race-specific Marginal Return Curves*



Notes: This figure plots race-specific estimates of hit rates for stops that were deterred each year post-*Floyd* relative to the year prior. Coefficients are computed using a two-stage-least squares regression as specified in Equation 2.6 and the shaded region displays the 95% confidence interval. The x-axis reflects the average change in stops relative to the two years prior to the reform. The horizontal red line is set to the mean outcome of White stops that were deterred in the first year of *Floyd*.

Table 2.2: *Racial Disparities in Mean Outcomes of Deterred Stops by Post-Floyd Year*

	White	Black	Hispanic	Pairwise Comparisons					
				White - Black		White - Hisp		Black - Hisp	
				Standard	Permuted	Standard	Permuted	Standard	Permuted
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Year 1	0.088***	0.067***	0.058***	0.011	0.001	0.000	0.001	0.059	0.006
	(0.007)	(0.003)	(0.003)						
Year 2	0.097***	0.068***	0.079***	0.000	0.000	0.002	0.000	0.003	0.000
	(0.005)	(0.003)	(0.003)						
Year 3	0.149***	0.127***	0.156***	0.101	0.029	0.594	0.512	0.017	0.000
	(0.011)	(0.007)	(0.010)						
Entire Post Period	0.100***	0.082***	0.079***	0.000	0.000	0.000	0.000	0.193	0.007
	(0.003)	(0.002)	(0.002)						

Notes: This table reports mean outcomes of stops that were additional deterred in each year of the post-period as well as over the entire three-year post-period, as specified in Equations 2.5 and 2.6. The stop outcome of interest is whether or not a stop led to an arrest or the discovery of weapons or drugs. Columns (1) to (3) report race-specific IV coefficients. Columns (4) to (9) report pairwise comparisons of race-specific IV estimates. We report standard/conventional p-values that test whether we can reject the null that the normalized difference in coefficients is different from zero using a two-sided Z-test. We also report permuted p-values from randomly re-assigning race, running our main specification, and computing a two-sided Z-test 1,000 times. All standard errors are clustered at the tract level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.3: *Racial Disparities in Mean Outcomes of Deterred Stops and Frisks*

	White	Black	Hispanic	Pairwise Comparisons		
				White - Black	White - Hisp.	Black - Hisp.
Stop Outcomes						
Arrest Made or Weapon/Drugs Found	0.100***	0.082***	0.079***	0.000	0.000	0.193
	(0.003)	(0.002)	(0.002)			
Arrest Made	0.062***	0.056***	0.054***	0.013	0.000	0.164
	(0.002)	(0.001)	(0.001)			
Weapon Found	0.015***	0.009***	0.010***	0.001	0.010	0.009
	(0.002)	(0.000)	(0.000)			
Drugs Found	0.022***	0.017***	0.015***	0.000	0.000	0.004
	(0.001)	(0.001)	(0.001)			
Court Summons Issued	0.075***	0.069***	0.072***	0.018	0.245	0.161
	(0.002)	(0.002)	(0.002)			
Frisk Outcomes						
Arrest Made or Weapon/Drugs Found	0.075***	0.041***	0.039***	0.000	0.000	0.320
	(0.004)	(0.001)	(0.001)			
Arrest Made	0.032***	0.015***	0.017***	0.000	0.000	0.034
	(0.004)	(0.001)	(0.001)			
Weapon Found	0.046***	0.028***	0.024***	0.000	0.000	0.002
	(0.002)	(0.001)	(0.001)			

Notes: This table reports mean outcomes of stops or frisks that were deterred by *Floyd*, as specified by the two-stage least squares regression in Equation 2.5. We report effects for our main stop outcome – whether or not a stop led to an arrest or the discovery of weapons or drugs – as well as for each potential stop outcome, separately. For deterred frisks of pedestrians, we report whether the encounter led to the discovery of drugs, weapons, or either of these outcomes. Columns (4) to (6) report pairwise comparisons of race-specific IV estimates. We report conventional p-values that test whether we can reject the null that the normalized difference in coefficients is different from zero using a two-sided Z-test. All standard errors are clustered at Census tract level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

2.4.2 Mechanisms

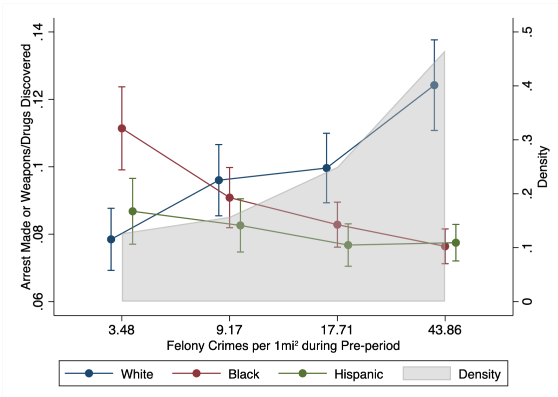
It is unclear what mechanisms lead officers to over-stop Black and Hispanic pedestrians. Do managerial pressures to maintain high stop rates in higher-crime neighborhoods drive racial bias? Or does racial bias stem from officer prediction errors, or even racial animus? To disentangle mechanisms, we examine how racial bias estimates differ by neighborhood characteristics, crime rates, and policing practices prior to the reform. Specifically, we partition the data into neighborhood quartiles and estimate Equation 2.5 on each subsample. Results are displayed in Figure 2.2.

Panel A explores effects by felony crimes per square mile. We find that racial bias estimates increase with felony crime rates, potentially reflecting the fact that the NYPD encouraged officers to maintain high stop rates in these neighborhoods. Marginal white stops are particularly productive in the highest crime neighborhoods, potentially due to behavioral biases that lead officers to underestimate returns to stopping whites.[21] As the felony crime rates fall, Black-white racial bias estimates subside, and even reverse in the bottom quartile. Panel B reports effects by the average number of officers conducting at least one stop per month. We observe similar but less pronounced trends as in Panel A. Panel C reports racial bias estimates by neighborhood racial composition, finding that racial differences emerge for the second quartile and remain rather constant thereafter.

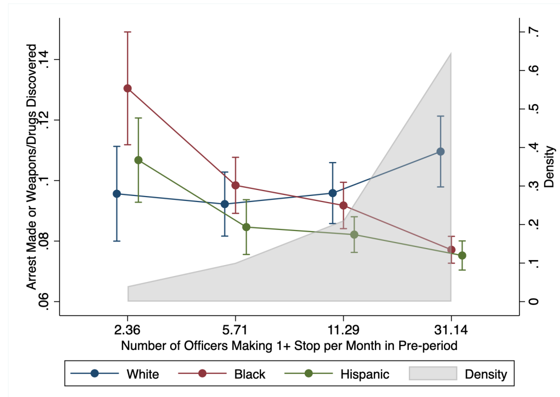
[21]It is not the case that there are no white pedestrians in these neighborhoods. The top quartile has a residential population that is over 25% non-Black and non-Hispanic.

Figure 2.2: *Race-specific Marginal Returns by Pre-Reform Crime Rates and Policing Practices*

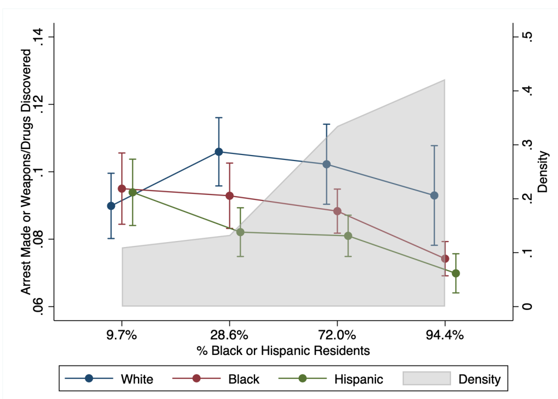
(a) *Felony Crime Rates*



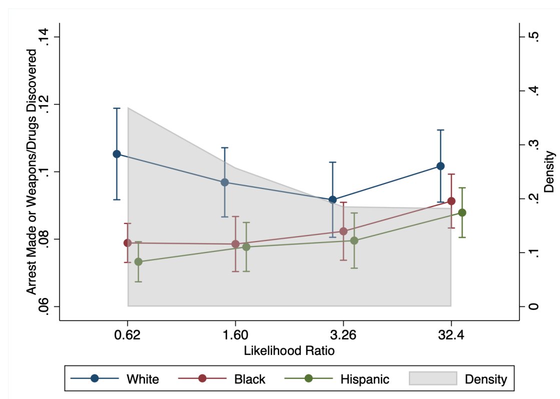
(b) *Number of Officers Conducting Stops*



(c) *Neighborhood Racial Composition*



(d) *Likelihood Ratio*



Notes: This figure plots coefficients from Equation 2.5 run separately for quartile sub-samples of neighborhoods based on pre-reform crime rates and policing practices. The pre-reform period spans Jan 1 2006 to Apr 15 2012. Panel A splits data by felony crimes per square mile. Panel B splits data by the average number of officers conducting 1 or more stop per month. We drop 64 of 2158 neighborhoods in Panels C and D since they report less than 500 residents and thus have noisy measures of racial composition. Panel C splits the sample by the fraction of residents who are listed as Black or Hispanic in the 2010 Census. Panel D splits data by the likelihood ratio which approximates how representative being Black or Hispanic is of arrested felons. Formally, we follow Bordalo *et al.* (2016) in defining this ratio as

$$\frac{\left[\frac{\text{\# Bl and Hisp Felony Arrests per year}}{\text{\# Bl and Hisp Residents}}\right]}{\left[\frac{\text{\# Wh Felony Arrests per year}}{\text{\# Wh Residents}}\right]}$$

Panel D reports estimates by how representative race is of felony offending in a given neighborhood (Bordalo *et al.*, 2016). If minority residents are over-represented in the tail of the risk distribution, then officers may overestimate the likelihood that a given minority pedestrian is committing a crime due to a "representativeness" heuristic where the offending type comes to mind when observing a pedestrian is a racial minority. This model of stereotypes has been found to explain racial bias in judges' bail decisions (Arnold *et al.*, 2018). In our data, we estimate the likelihood ratio for each neighborhood as the relative likelihood a Black or Hispanic pedestrian is arrested for a felony crime relative to that for a white pedestrian.[22] Panel D shows that racial bias estimates are actually largest when the racial minority is least representative of being arrested for a felony. This result stems from the fact that neighborhood crime rates negatively correlate with our likelihood ratio: as the residential population of a neighborhood becomes more Black and Hispanic, the less predictive race becomes of criminal behavior. It is possible that officer beliefs are influenced by racial disparities outside of a particular patrol neighborhood. For example, Black and Hispanic residents are nine times more likely than whites to be arrested for a felony crime across all neighborhoods. In turn, stereotypical beliefs or other behavioral biases may lead officers to overestimate the likelihood that minority residents are committing crimes in higher-crime, predominantly minority neighborhoods – the exact neighborhoods where being Black or Hispanic is relatively *less predictive* of criminal behavior.

Putting these results together, we find that racial bias is most strongly correlated with felony crime rates and that in these higher-crime neighborhoods, officers significantly under-stopped white pedestrians. If racial animus was the sole contributor to racial bias, we would expect constant differences in marginal returns that do not vary with neighborhood crime rates. While racial animus may have played a role, these results suggest that racial bias may

[22]That is, we estimate:

$$R(felon, BH, -BH) = \frac{\left[\frac{\text{\# Black and Hispanic Felony Arrests per year}}{\text{\# Black and Hispanic Residents}}\right]}{\left[\frac{\text{\# White Felony Arrests per year}}{\text{\# White Residents}}\right]}$$

For a given neighborhood, this ratio provides an approximation of how much more likely a minority resident is of being arrested for a felony than a white resident.

have, in part, resulted from orders to concentrate stops in higher-crime neighborhoods and patrol officers' tendencies to make racially-biased prediction errors.

2.5 Conclusion

Through the lens of the *Floyd* reform, we show that racial disparities in stop rates do not reflect true racial differences in criminal behavior. Black-white and Hispanic-white racial differences in IV estimates from the first year of the reform provide evidence that the police racially discriminated prior to the reform. Even after the first year of the reform, we estimate that patrol officers would have to reduce stops of Black and Hispanic pedestrians by a combined 298,000 stops per year in order to equalize marginal returns across racial groups. Since our approach allows officers to "rationally" racial profile, these findings understate the total level of racial discrimination. As marginal white stops are more productive than marginal Black or Hispanic stops, our findings also suggest that closing racial disparities in stop rates would generate efficiency gains for the police.

Chapters 1 and 2 illustrate the outsize role the police play in urban communities of color, and especially, for young men of color. Future research should examine the downstream consequences of pedestrian stops and other lower-level police encounters that have the potential to generate long-run social costs by disrupting the life trajectory of young people. By estimating these social costs, this research could inform better police practices that reduce existing racial and place-based opportunity gaps.

Chapter 3

Civic Responses to Police Violence[1]

3.1 Introduction

In recent years, acts of police violence have garnered significant public attention. The high-profile killings of George Floyd and Breonna Taylor compelled an estimated 15 to 26 million Americans to protest against police brutality and systemic racism (Buchanan *et al.*, 2020b). This movement was part of a larger national reckoning, which saw violent counter-protests in Portland, player walk-outs across professional sports leagues, and widespread calls for police reform. These events are not without historical precedent. Dating back to the 1965 Watts riots, the four largest episodes of urban unrest in America were all triggered by police use of force (DiPasquale and Glaeser, 1998).

While recent events have raised questions about the role of state-linked violence in democratic societies, researchers know little about the effects of police violence on local political participation. A large literature has shown that interactions with the criminal justice can have drastic demobilizing consequences (Weaver and Lerman, 2010; Lerman and Weaver,

[1]Co-authored with Desmond Ang. We thank the editor and three anonymous referees of the American Political Science Review for their constructive feedback. This article has benefited from helpful comments and suggestions from Amy Finkelstein, Edward Glaeser, Lawrence Katz, Amy Lerman, Andrei Shleifer, Ariel White, seminar participants at University of California, Berkeley, Massachusetts Institute of Technology, and the 2020 APSA conference. Tebes benefited from financial support from a James M. and Cathleen D. Stone PhD Scholar fellowship from the Multidisciplinary Program in Inequality & Social Policy at Harvard University. All errors are our own.

2014a; White, 2019b,a). These studies tend to focus on individuals or families with direct contact with law enforcement or carceral systems. But as recent history demonstrates, police killings are often public and visible events that may incite concerns about institutional trust, racial discrimination and procedural justice even among individuals with little or no direct relation to the deceased. Understanding the political ramifications of state-linked violence is thus central to questions of democratic governance and may bear important implications for our understanding of recent and future elections.

This paper provides the first causal evidence of the impact of police violence on voter participation. To do so, we combine highly detailed voter registration data from Los Angeles County with novel incident-level data on the timing, location and context of nearly three hundred police killings spanning nearly a decade. As the occurrence of police killings is not random, a simple comparison of civic engagement in neighborhoods with high and low rates of police violence is likely to be confounded by a number of correlated factors. We instead employ a dynamic difference-in-differences (i.e., event study) design to leverage hyper-local variation in exposure to police violence. This approach compares changes in participation before and after a police killing in the exact Census block where the killing occurred to pre-post changes in adjacent blocks in the same neighborhood. Causal identification then hinges on the assumption that post-killing voting patterns in neighboring blocks serve as a valid counterfactual for those in incident blocks. We provide empirical support for this assumption by showing that registration and voting trends closely mirror each other across treatment and control areas in the elections leading up to a police killing.

We find that police killings may mobilize local residents to engage with the electoral process. Registrations and votes in incident blocks increase by roughly 5% in the elections following a police killing. While gains in voter turnout are relatively short-lived, registration effects continue more than a decade later. However, these effects are highly-localized, pointing to the potential role of information in mediating responses. For the 80% of sample killings that went unmentioned in local newspapers, we find no evidence of spillovers extending beyond the block of the incident. In contrast, for killings that did receive media coverage, we observe registration increases up to half a mile away.

To interrogate the role of family and household members, we first compare across police killings that occurred near and far from the home of the deceased. We find similar increases in registration and voting regardless of incident location. Searching individual-level voter registration files, we also find little evidence of increased participation among individuals sharing the same last name as the deceased. Together, these findings suggest that changes in registration and voting are unlikely to be driven by family or household members of the deceased, but rather other local residents who may have seen or heard about the killing.

However, the aggregate effects do mask significant heterogeneity across demographics. Increased civic engagement is driven entirely by Black and Hispanic citizens, who are 8% and 5% more likely to register as a result of exposure to local police killings, respectively. We find no statistical or practical impact on the political behavior of nearby whites and Asians. We also document striking differences along other dimensions. The largest effects come from younger voters, new registrants and Democrats. We find no significant impact among Republicans or individuals over age 35. These findings accord with a host of survey evidence documenting deep racial and partisan divisions in views of law enforcement, with minorities and liberals far more concerned about police use of force than whites and conservatives.

To unpack mechanisms, we first explore differences in civic responses based on whether the person killed by police possessed a weapon. If changes in turnout reflect concerns about crime and support for more intensive policing, we would expect larger effects for police killings of armed individuals. In fact, our data support the opposite narrative. Point estimates of civic spillovers are roughly three times larger following killings of unarmed individuals – those events in which police actions may have seemed the least justifiable. We corroborate these findings with data on local referenda voting and find that police killings significantly increase support for propositions designed to reduce criminal penalties for non-serious offenses. Together, our findings suggest that acts of police violence may drive some residents to the polls in an attempt to reform the criminal justice system.

3.2 Motivation

The core question of this paper is how police killings affect local registration and voting. A large literature documents the potential demobilizing effects of the criminal justice system. Examining Add Health and Fragile Families survey data, Lerman and Weaver (2010; 2014a) find that previously arrested, convicted or incarcerated individuals are significantly less likely to report having registered or voted in recent elections. These patterns are causally validated by White (2019b), who exploits random courtroom assignment and find that short jail spells reduce future turnout for first-time misdemeanor defendants. Related work documents demobilization spillovers among other individuals living in the same household (Lee *et al.*, 2014; White, 2019a). These effects are likely driven in part by the downstream economic costs of a criminal record (Western and Pettit, 2010; Agan and Starr, 2018) as well as the stigma, emotional trauma, and institutional distrust that may accompany it. Consistent with the latter, recent work has shown that police killings may cause nearby adolescents to disengage from formal institutions by dropping out of school (Ang, 2021b).

At the same time, research has demonstrated that some groups may be mobilized by perceived injustice and political threats. While much of the empirical work centers on immigration policy (Pantoja *et al.*, 2001; White, 2016; Zepeda-Millán, 2017; Gutierrez *et al.*, 2019), there is significant reason to believe that police violence may fuel related concerns about fairness and justice (Weitzer and Tuch, 2006). The recent surge in Black Lives Matter demonstrations following the murder of George Floyd as well as the historic protests that erupted after the police beatings of Rodney King and Marquette Frye provide anecdotal support for such a pathway. Similarly, Walker (2014, 2020) finds a positive correlation between "proximal contact" with the criminal justice system – knowing a close friend or family member who has been arrested, charged, or questioned by police – and self-reported measures of informal political participation – such as signing a petition, attending a community meeting, or attending a demonstration or protest.[2] Consistent with the

[2] Walker (2020) also finds a positive, but insignificant, relationship between proximal contact and self-reported turnout in National Crime and Politics Survey data but a near-zero relationship in American National

potential role of perceptions of injustice, she provides evidence that part of this relationship is mediated by the impact of proximal contact on an individual's belief about unfair targeting by law enforcement.

These narratives point to opposite conclusions about the effects of police killings on local electoral participation. Nonetheless, there exists little causal evidence on this topic. One of the central challenges is that criminal justice contact – whether direct or indirect – may be correlated with a number of observed and unobserved factors that may affect an individual's likelihood of voting. This is compounded by the scarcity of reliable data capturing direct or proximal contact much less linking those measures to voting behavior. While researchers have attempted to address these concerns by employing multivariate analysis of self-reported, cross-sectional survey data, such methods may still suffer from measurement error and do not address the fundamental concern regarding selection bias and the endogeneity of police encounters.

Empirical research leveraging quasi-random variation and time-series or panel analysis has sought to answer related questions about how officer use of force affects citizen cooperation with law enforcement (Lerman and Weaver, 2014b; Desmond *et al.*, 2016; Cohen *et al.*, 2019; Zoorob, 2020; Ang *et al.*, 2021). However, this work finds mixed effects on 311 and 911 calls for service and leaves unanswered the central question of this paper: whether citizens strategically respond to police violence by engaging with electoral systems. A separate body of work interrogates the effects of political protests, including the Rodney King riots and other events that arose in the wake of high-profile use of force incidents (Enos *et al.*, 2019; Wasow, 2020). However, as protests are themselves a consequence of police violence, such case studies may be unable to disentangle the impact of one from the other. By focusing on high-profile events, these studies may also provide limited insight into the political ramifications of the vast majority of use of force incidents that receive little or no media attention.

Beyond the core question of how acts of police violence affect electoral participation

Election Studies data.

is who is affected by these events and why. While researchers disagree about the role of racial bias in officer use of force (Fryer Jr, 2019b; Knox *et al.*, 2020; Hoekstra and Sloan, 2020; Durlauf and Heckman, 2020), relative to their population shares, Black and Hispanic individuals are significantly more likely to die at the hands of police than their white and Asian counterparts (Edwards *et al.*, 2019). Perhaps not surprisingly, Black and Hispanic individuals are far more likely to believe that use of force is exercised in a racially-biased manner and that police violence is a pressing social concern (Leiber *et al.*, 1998; Weitzer and Tuch, 2002; AP-NORC, 2015).[3] In this light, one might expect that Black and Hispanic political participation would be particularly sensitive to police violence and that these responses may be driven by concerns about police accountability and racial discrimination.

3.3 Data

Police Killings

Incident-level data on police killings come from the Los Angeles Times Homicide Database, which tracks all known officer-involved killings in Los Angeles county and includes 294 incidents between the 2002 and 2010 general elections.[4] For each killing, the data includes the name, age and race of the deceased as well as the exact address and date of the event. We supplement this with information on whether the incident was reported by local newspapers.[5] For roughly 85% of killings, we were also able to determine whether a weapon was recovered from the deceased. This information was hand-coded from Los Angeles County District Attorney reports as well as police reports and other sources.

[3] For example, a 2015 survey found that 75% of Black respondents and over 50% of Hispanic respondents believe that police violence is a very or extremely serious issue, relative to 20% of whites. Similar disparities arose when asked whether police "deal more roughly with members of minority groups" (AP-NORC, 2015).

[4] In a handful of cases, multiple individuals were killed in a single incident. The total number of distinct incidents is 286.

[5] We searched for each incident by the name of the deceased in the print versions of six local newspapers: the Los Angeles Times, the Los Angeles Daily News, Pasadena Star News, San Gabriel Valley Tribune, Torrance Daily Breeze and Whittier Daily News. The combined daily circulation of the papers is roughly 1 million copies.

Note that these contextual measures may provide an incomplete picture of the surrounding events. Often officers acted under faulty information. For example, in one incident, police killed a man who was reported to have a gun but who was actually holding a water hose nozzle. In other cases, killings were precipitated by seemingly innocuous encounters that quickly escalated – such as, when a man lunged for an officer's gun after he was stopped for riding a bicycle on the sidewalk. Nonetheless, weapon information has the benefit of being objectively verifiable and can be found in all available incident reports.

Panel A of Table 3.1 provides a summary of the police killings data. 53% of deceased individuals were Hispanic, 29% were Black, 15% were white and 3% were Asian.[6] Relative to their population shares, Black (Hispanic) individuals are roughly six (two) times more likely to be killed by police than whites. The vast majority of individuals (96%) were male and the average age was 30 years old.

Consistent with national statistics, 47% of those killed were armed with a firearm (including BB guns and replicas), while 23% possessed some other type of weapon. This includes items like knives and pipes as well as individuals who attempted to hit someone with a vehicle. 14% of individuals were completely unarmed. We were unable to find contextual information for the remaining 16% of incidents.

Notably, the vast majority of killings received little or no media coverage. Only 18% of sample killings were ever mentioned in any of six local newspapers. Conditional on coverage, the median number of articles is two. The most mentions of any incident was 28, far below the level of media attention garnered by recent high-profile police killings.

Examining contextual factors separately by race, Black and Hispanic individuals killed by police were younger on average than white and Asian individuals (29 vs. 36 years old, respectively) and more likely to possess a firearm (53% vs. 24%). However, rates of media coverage are similar between groups (19% vs. 16%).

Regardless of circumstance, involved officers were never prosecuted. The District Attorney did not pursue criminal charges against police following any of the 294 sample

[6]Race categories are mutually exclusive.

Table 3.1: *Summary Statistics*

Panel A: Police Killings

	All	Deceased Race: Black/ Hispanic	Deceased Race: White/ Asian
Deceased Demographics			
Age	30.34	28.97	36.30
Black	0.28	0.35	0.00
Hispanic	0.53	0.65	0.00
White	0.16	0.00	0.85
Asian	0.03	0.00	0.15
Male	0.96	0.97	0.93
Newspaper Mentions			
Any	0.18	0.19	0.16
Total	0.70	0.77	0.38
Median (if any)	2.00	2.00	1.00
Deceased Weapon			
Unarmed	0.14	0.13	0.20
Knife	0.23	0.20	0.31
Gun	0.47	0.53	0.24
Unknown	0.16	0.14	0.25
Killings	294	238	55

Panel B: Block Demographics and Registration

	All	w/ Killing	w/o Killing: Treat Blk Grp	w/o Killing: Ctrl Blk Grp
Demographics (2000 Census)				
Pop. 18+	99.28	190.55	100.26	98.84
Black %	0.09	0.16	0.16	0.08
Hispanic %	0.38	0.57	0.52	0.37
White %	0.39	0.16	0.20	0.40
Asian %	0.12	0.10	0.10	0.12
Voter Registration (2002)				
Registrations	56.30	79.93	48.20	56.53
Black %	0.13	0.27	0.26	0.12
Hispanic %	0.28	0.39	0.37	0.28
White/Asian %	0.56	0.31	0.35	0.57
Age 18-34 %	0.26	0.32	0.30	0.26
Age 35-54 %	0.40	0.39	0.39	0.40
Democrat %	0.53	0.62	0.61	0.52
Republican %	0.27	0.17	0.19	0.28
New Reg %	0.36	0.42	0.39	0.35
Votes	24.64	29.46	18.51	24.87
Blocks	68,326	285	2,675	65,366

Panel A provides summary statistics for the police killings data, separately for killings of Blacks and Hispanics and killings of whites and Asians (we are unable to obtain race for one individual). Newspaper mentions come from a search of each incident by deceased name in six local newspapers, including the Los Angeles Times. Any indicates mention in any article. Total is the number of articles mentioning the incident. Median is the median number of articles, conditional on being mentioned. Unarmed refers to suspects that did not have a weapon, gun refers to suspects with firearms (including BB guns and replicas), knife refers to suspects with any other type of weapon, unknown refers to incidents where weapon type was not obtainable from District Attorney reports and other sources.

Panel B provides summary statistics for the Census blocks included in the main analysis. The remaining columns are mutually-exclusive sub-samples. “w/ Killing” are blocks where a killing occurred during the sample period. “w/o Killing” refers to blocks without police killings, separated by blocks in Census block groups that experienced a killing (“Treat Blk Grp”) and blocks in Census block groups that did not experience a killing (“Ctrl Blk Grp”). Registration data include total registration and vote counts as well as the share of registrations by race, age, party affiliation, and new registrants (i.e., those who registered within the prior 4 years) during the 2002 general election. The reason there are 285 killing blocks versus 294 killings is because some incidents involved multiple deaths. In total, there were 286 distinct incidents, two of which occurred in the same block on separate dates.

killings. This is consistent with national statistics, which find that criminal charges are filed against police in fewer than half a percent of all officer-involved shootings.

Voter Registration, Turnout and Preferences

Police killings are geocoded to Census blocks and merged to voting information from the California Statewide Database. The database contains information on the number of individuals registered to vote and the number of ballots cast at the 2010 Census block-level for each general election from 2002 to 2010.[7] The advantage of these data relative to standard voter registration files is that they capture registration and voting *at the date of each election*, which allows us to more precisely measure effects.[8] This is particularly important when considering geographies as small as Census blocks, which average less than 0.10 square miles and fewer than 100 adults. In addition to total registration and vote counts, disaggregated counts by ethnicity (i.e., Hispanic, Asian), party affiliation (i.e., Democrat, Republican, Independent/other), age and duration of registration are also available. We combine these data with block-level demographic information on the voting age population from the 2000 and 2010 Censuses.

In order to examine the impact of police killings on voter preferences, we leverage voting data from two referenda that proposed changes to the severity of criminal sentencing laws. The first – Proposition 66 in 2004 – would have limited California's "three strikes" law to apply only to violent and serious felonies. The second – Proposition 5 in 2008 – would have enacted numerous measures to reduce the criminal penalties for drug offenses, including the reduction of marijuana misdemeanors to infractions, and the expansion of drug treatment and rehabilitation programs. Block-level estimates of the share of ballots cast for and against

---

[7] Data at the 2010 Census block-level are not available for elections prior to 2002. Total vote counts and demographic-specific registration counts are only available through 2010, while total registration counts are available through 2016. In the Online Appendix, we extend the registration sample beyond 2010 and find highly persistent effects lasting more than a decade.

[8] Due to irregular registration purges and resident mobility, voter files obtained months after an election can have registration counts that differ substantially from known election-day aggregates.

each proposition comes from the California Statewide Database.[9]

Analysis Sample

Since vote counts are only available until 2010, our main analysis focuses on the 2002-2010 general elections. To improve precision, we restrict the sample to blocks with five or more residents of voting age in the 2000 and 2010 Censuses. In robustness analysis, we show similar results under alternative sample restrictions and when extending the registration analysis to include more recent elections.[10]

Panel B of Table 3.1 provides a summary of the voter registration data. Blocks that experienced a police killing had, on average, 190 adult residents in 2000, compared to roughly 99 residents in other blocks. Notably, treated blocks are quite similar to untreated blocks in the same Census block group (i.e. the effective control group) in terms of the racial, political and age characteristics of residents and registered voters. These areas also experienced similar rates of turnout (37 vs. 38%) and registration (42% vs. 48%) in 2002.

3.4 Empirical Strategy

Exposure to Police Killings

A primary concern for identification is that police shootings are likely not random. Thus, a cross-sectional comparison of turnout rates in areas with high and low prevalence of police violence could be confounded by the fact that police killings are most likely to occur in neighborhoods with high shares of minority or impoverished households, both of which are associated with low turnout. Furthermore, if local trends in crime or law enforcement

---

[9]These estimates are generated using ecological inference, which combines precinct-level election results with individual-level address and turnout data from the county registrar. The resulting individual-level voting propensities are then aggregated to the Census block-level. Additional information is available at https://statewidedatabase.org/.

[10]Research documentation and data that support the findings of this study will be openly available in the APSR Dataverse at https://doi.org/10.7910/DVN/EOIIEV.

predict the timing and location of police killings, biases could remain even if one were to include area fixed effects in panel analysis.

To address these concerns, we adopt a difference-in-differences approach similar to Ang (2021b) and exploit *within-neighborhood* variation in exposure to police killings. Causal identification comes from comparing changes in voting before and after a police killing in the blocks nearest where the incident occurred to changes in voting among neighboring blocks in the same neighborhood. Pairing this approach with finely disaggregated data allows us to leverage the exact time and location of each police killing and to construct the counterfactual from demographically- and geographically-similar areas experiencing the same local conditions and shocks, except for the killing itself.

The validity of this strategy is aided by two factors. First, police killings are quite rare and difficult to predict. Over 300,000 arrests and nearly 60,000 violent crimes occur in Los Angeles each year, compared to fewer than 50 officer-involved killings. Furthermore, many police killings are entirely unaccompanied by violent crime, as only a quarter of events involved armed suspects who fired at others. Thus, while crime rates and policing intensity may differ across neighborhoods, the exact timing and location of officer-involved shootings within those neighborhoods is plausibly exogenous.

Second, in contrast to the handful of high-profile events in recent years, the vast majority of police killings receive little or no media coverage. Consistent with Ang (2021b), who finds that educational spillovers of police violence are limited to less than 0.50 miles, living in one block versus the other is likely highly correlated with even knowing about the existence of a police killing. This provides meaningful treatment variation within neighborhoods.

Graphical Evidence

As evidence, we first examine how effects differ by geographical proximity to a killing. To do so, we construct the following distance metric to capture a given Census block's proximity to a police killing: for each police killing, we estimate the minimum radius needed for a circle centered on the police killing to cover at least 75% of a block's area.[11] Using this metric, we run the following generalized difference-in-differences regression to explore how impacts on registration count dissipate with distance from a killing:

$$y_{b,t} = \delta_b + \delta_{n,t} + \sum_d \alpha_d Distance_d + \delta_{POP_b xELEC_t} + \epsilon_{b,t} \tag{3.1}$$

$y_{b,t}$ is the number of registered voters in block b at election t. δ_b and $\delta_{n,t}$ are Census block fixed effects and neighborhood-by-election fixed effects. Because block-level population counts are only available from the decennial Census, we include interactions between election fixed effects and deciles of estimated voting age population in 2002 ($\delta_{POP_b xELEC_t}$) to account for the possibility of differential population growth between blocks. However, as we will demonstrate, results are robust to excluding these population controls. $\sum_d Distance_d$ are a set of mutually-exclusive treatment indicators that track a block's distance to the nearest police killing that occurred prior to election t.[12] We partition shootings first by whether or not the block was directly exposed to a killing (i.e. occurred in the block), and then into 0.1-mile bins up to 2 miles from the shooting.

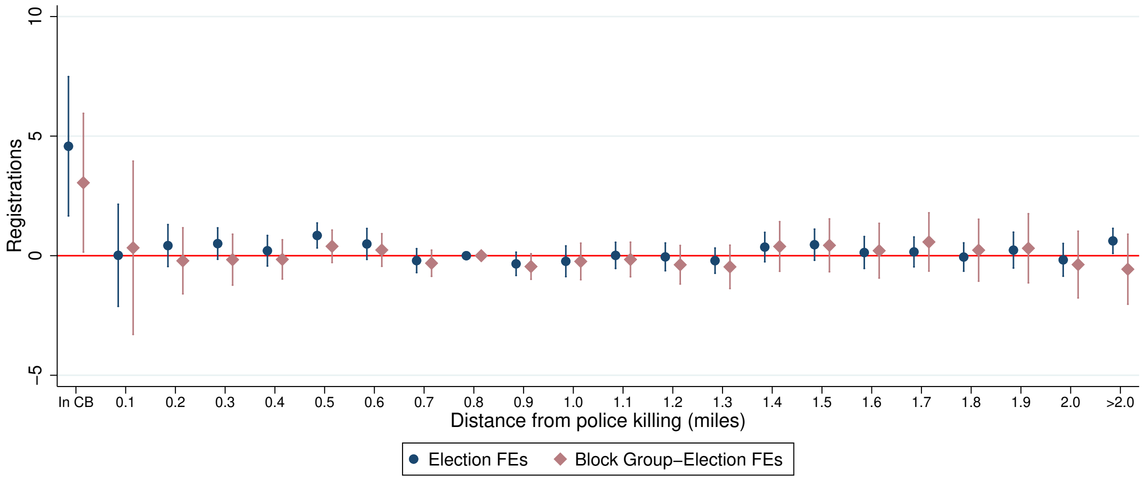
Figure 3.1 plots α_d coefficients under two different specifications: controlling for election fixed effects across the entire sample and at the Census block group level. These coefficients represent the average difference between the pre-post change in registrations experienced by blocks with distance d and blocks in the omitted group (i.e., blocks in the "zero-impact" region between 0.7 and 0.8 miles from a killing).[13] The graphical evidence suggests that impacts are indeed hyper-local. Across specifications, treated Census blocks experience a significant increase in registrations of between 3.5 and 5 counts (4-6% of the pre-killing mean). Consistent with the under-publicized nature of police killings, effects fall off dramatically

[11] We estimate minimum radius in 0.05-mile increments up to 2 miles. Reassuringly, over 95% of blocks containing a police killing are within 0.35 miles of the killing based on this measure. We found other metrics, such as the distance to a block's centroid, to be substantially noisier.

[12] For example, if a block's nearest killing before 2007 was 1.5 miles away then experienced another killing 0.5 miles away in 2007, distance is 1.5 miles for elections before 2008 and 0.5 miles for 2008 forward.

[13] In the block group specification, distance indicators are set to 0 within treated block groups so that neighboring blocks are not mechanically biased downwards through the neighborhood-election fixed effects.

Figure 3.1: *Effects by Distance From Police Killing*



Notes: Figure reports coefficients from estimation of Equation 3.1 on registration counts. Sample spans 2002 to 2010 general elections. Indicators are mutually-exclusive and track a Census block's minimum distance to police killings that occurred prior to a given election."In CB" refers to blocks where killings occurred. Other blocks are partitioned by 0.1-mile bins. Blue dots include election fixed effects. Red diamonds include Census block group by election fixed effects. Standard errors are clustered at the Census block group-level. Full regression results are included in Table A.I of the Online Appendix, Columns 1 and 2.

with spatial distance, with near-zero estimates for neighboring blocks only slightly further away.

Estimating Equation

To estimate effects on civic engagement, we next employ a dynamic difference-in-differences (i.e. event study) model. Drawing on the distance analysis, treatment is defined as Census blocks that experienced a police killing and neighborhood is defined at the Census block group-level. We estimate the following base equation on the block-level panel data:

$$y_{b,t} = \delta_b + \delta_{n,t} + \sum_{\tau \neq -1} \beta_\tau Shoot_{t,\tau} + \delta_{POP_b xELEC_t} + \epsilon_{b,t}. \tag{3.2}$$

This is essentially analogous to Equation 3.1 except we replace the set of treatment distance indicators ($\sum \alpha_d Distance_d$) with a set of time to treatment indicators ($\sum \beta_\tau Shoot_{t,\tau}$), fixing treatment to the first killing that occurred in a Census block between the 2002 and

2010 general elections.[14] Neighborhood-time fixed effects ($\delta_{n,t}$) are defined by Census block groups, which include about 11 blocks and measure less than one square mile in area, on average. The coefficients of interest (β_τ) represent the differential change between relative time τ and the last period prior to the police killing in the incident block relative to that same change over time among other blocks in the same neighborhood. Drawing on Bertrand *et al.* (2004), standard errors are clustered by Census block groups, allowing for correlation of errors within each of the sample's 6,400 Census block groups.

Crime and Migration

One potential threat to identification is that trends in violent crime may influence both the level of civic engagement (Bateson, 2012; Sønderskov *et al.*, 2020) and the presence of police killings in an area. However, given that we control for block group-election fixed effects, any biases would have to be hyper-local, affecting individuals on one street but not the next within the same neighborhood. To test this, Panel A of Figure A.I of the Online Appendix estimates Equation 3.2 on criminal homicides in a block-year. Since geocoded data for *all* crimes and arrests is only available after 2010, we replicate the exercise using information on the timing and location of police killings from 2010 to 2016 in Panels B and C. In all cases, we find little support for differential trends in local crime or policing activity before or after police killings, reinforcing the plausible exogeneity of these events.

Another potential threat is selective migration in response to police violence. As block-level population counts are only measured every decade, we are unable to directly test for differential migration using our main event study model. However, several pieces of corroborating evidence suggest that it is unlikely to be a serious concern. First, given the positive effects on registration and vote counts, the main threat would be if police killings *increased* population in a neighborhood. However, examining administrative schooling data from Los Angeles, Ang (2021b) finds little effect of exposure to police killings on school

[14] Only one block experienced multiple separate incidents (i.e. killings that occurred on different days) over the sample period. Results are robust to excluding that block.

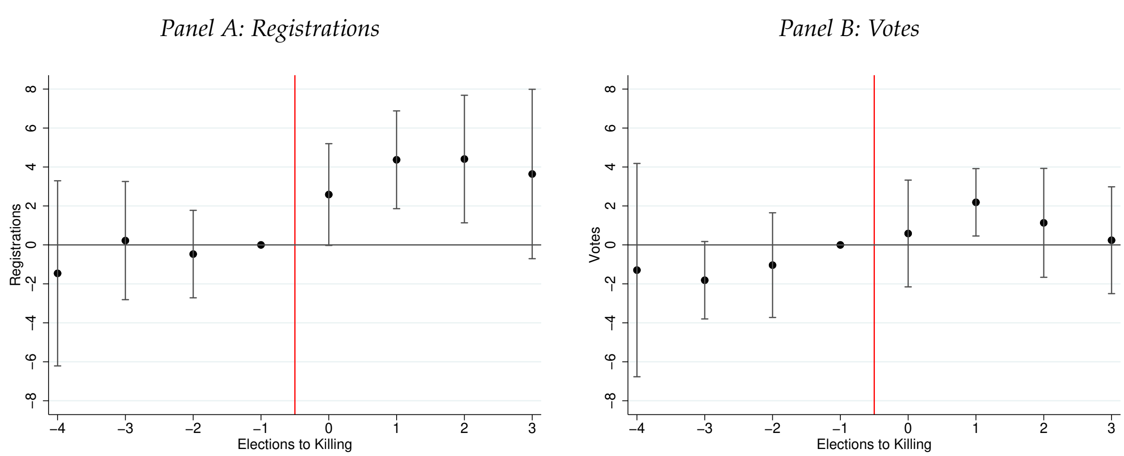
transfers among nearby students. In fact, point estimates are near zero and suggest, if anything, small *reductions* in local population. Second, 2006-2010 ACS data indicates that the share of individuals who reported residing at the same house one year prior is virtually identical between Census blocks that did and did not experience a police killing (86.6% and 86.8%, respectively). Third, a simple difference-in-differences regression comparing changes in log population from 2000 to 2010 across blocks that did and did not experience a killing returns a precise zero estimate (β = 0.006, p-value=0.78).

3.5 Effects on Registration and Turnout

3.5.1 Main Results

Turning to our primary results, Panel A of Figure 3.2 examines the impact of police killings on local registration counts. The omitted period is the last election prior to a killing and the sample spans the 2002 to 2010 general elections.

Figure 3.2: *Effects on Civic Engagement*



Notes: Figure shows treatment estimates and 95 percent confidence intervals from estimation of Equation 3.2 on registrations (pre-treatment mean = 81.6) and votes (pre-treatment mean = 42.9). Unit of observation is registrations/votes in a Census block-election. Standard errors are clustered by Census block group. The sample spans the 2002 to 2010 general elections and treatment is defined by blocks where police killings occurred during the sample period. Red vertical line represents time of treatment. Full regression results are included in Column 1 of Table A.III in the Online Appendix.

In the elections prior to a killing, we find strong evidence of parallel trends between treatment and control areas. Treatment coefficients for $\tau < 0$ are near zero and statistically insignificant, both individually and jointly ($F = 0.33, p = 0.806$). These findings reinforce the plausible exogeneity of police killings and provide support for parallel trends in the counterfactual.

Following police killings, registration increases significantly among nearby citizens. Treated blocks gain, on average, about 2.5 additional registrants in the election immediately following the killing and about 4.5 registrants within four years. Given that treated blocks contain an average of 80 registrants prior to treatment, these effects represent a meaningful increase of 3 to 5%. The stability of point estimates four to eight years after exposure suggests that effects on registration are persistent over time. As corroboration, Figure A.II of the Online Appendix expands the sample to include police killings and elections through 2016 and finds significant effects on registration more than a decade after a killing.[15]

Panel B presents analogous results for vote counts. We again find little evidence of differential trends in ballots cast in the lead-up to police killings. The pre-treatment coefficients are individually and jointly insignificant ($F = 1.09, p = 0.350$). After killings, we find a significant, if short-lived, increase of approximately 2 votes (5% of the pre-killing mean).

Table 3.2 presents a series of robustness checks under alternative specifications. For concision, Column 1 estimates our base model with a single post-treatment indicator, representing the average treatment effect across all elections. We find that, on average, police killings lead to 3.6 more registrations and 1.7 more votes per election in treated blocks. To account for potential confounds due to local crime, Column 2 controls for the number of homicides in a block in the two years preceding each election. Column 3 includes quintiles of minority population share by election fixed effects to allow for differential voting patterns among minority neighborhoods, which may be more likely to experience police killings. Given that younger individuals may have aged into voting eligibility during the sample

[15] As 2010 block-level total vote counts and age/ethnicity/party-specific registration counts are not available for elections after 2010, we are unable to extend the voting and heterogeneity analysis to more recent periods.

period, Column 4 includes interactions between election fixed effects and the population of 10 to 17 year-olds in a block in 2000. To account for differential population growth, Column 5 includes interactions between percent population change from 2000 to 2010 and election fixed effects. In contrast, Column 6 excludes all population controls from the regression.[16] To demonstrate robustness to sample selection, Column 7 drops the single Census block that experienced multiple distinct police killing incidents over the sample period, while Column 8 expands the sample to include all Census blocks, even those with less than five adults in 2000 or 2010. Alternatively, Column 9 restricts the sample to Census blocks with at least 10 registered voters in 2002.

Table 3.2: *Effects on Civic Engagement: Alternative Specifications*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: DV=Registrations								
Treat x Post	3.636***	3.638***	3.632***	3.614***	3.483***	5.357***	3.495***	3.431***	3.293**
	(1.308)	(1.308)	(1.308)	(1.309)	(1.295)	(1.364)	(1.306)	(1.285)	(1.332)
Mean	81.59	81.59	81.59	81.59	81.59	81.59	80.72	81.59	85.13
	Panel B: DV=Votes								
Treat x Post	1.743**	1.745**	1.722**	1.742**	1.666*	3.511***	1.748**	1.234	1.631*
	(0.867)	(0.867)	(0.866)	(0.869)	(0.861)	(0.909)	(0.871)	(0.876)	(0.877)
Mean	42.87	42.87	42.87	42.87	42.87	42.87	42.23	42.87	44.75
Model	Main	Homicide Ctrls	Minority% x Elect	Pop 10-17 x Elect	Pop Δ x Elect	w/o Pop Ctrls	w/o Multi-Treaters	Full Sample	2002 Reg ≥ 10
Obs.	341,420	341,420	341,420	341,420	341,420	341,420	341,415	547,815	306,340

Table shows results from estimation of Equation 3.2 on registrations and votes, replacing time to treatment indicators with a single post-treatment dummy. Column 1 examines our preferred specification. Column 2 controls for the number of homicides in a block in the two years preceding each election, yielding $\beta_{homicide}$ of 0.320 (p-value=0.079) for registrations and 0.215 (p-value=0.079) for votes. Column 3 includes minority share quintile by election fixed effects. Column 4 adds quintiles of population aged 10-17 in 2000 by election fixed effects. Column 5 includes interactions between percent population change (from 2000 to 2010) and election fixed effects. Column 6 excludes population decile by election fixed effects. Column 7 drops the single Census block that experienced more than one police killing during the sample period. Column 8 includes all Census blocks, even those with less than 5 people over the age of 18 in 2000 or 2010. Column 9 restricts the sample to Census blocks with 10 or more registered voters in 2002. Unit of observation is the block-election. The sample period spans the 2002 to 2010 general elections. Standard errors clustered by Census block group. Mean registrations and votes for treatment blocks during the election prior to treatment are listed. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

[16] Full event study results excluding the population decile-by-election fixed effects are discussed in Mechanisms and included in Figure A.III of the Online Appendix.

We find similar results across all specifications, with significant and positive effects on voter registration and turnout. These findings provide evidence of the robust causal relationship between police killings and local political participation. In particular, we find that individuals are mobilized to register and vote by extreme acts of police violence.

3.5.2 Heterogeneity

Voter and Deceased Race

Given large demographic and partisan differences in views of law enforcement, we examine heterogeneous responses to police violence. To explore how effects differ across voter race, we make use of vote and registration counts by ethnicity provided by the California Statewide Database, which predicts Hispanic ethnicity from voter surname using the Census Bureau's Passel-Word list and Asian ethnicity using the surname dictionary of Lauderdale and Kestenbaum (2000).[17] From these measures, we generate estimates of Black (white) vote and registration counts using the following formula:

$$\begin{aligned} VoteBlk_{b,t} = {} & (VoteTot_{b,t} - VoteHisp_{b,t} - VoteAsn_{b,t})\times \\ & \left(\frac{\%Blk_{b,2010} \times \%VoteBlk_t}{(\%Blk_{b,2010} \times \%VoteBlk_t) + (\%Wht_{b,2010} \times \%VoteWht_t) + (\%Oth_{b,2010} \times \%VoteOth_t)}\right) \end{aligned} \tag{3.3}$$

where $VoteTot_{b,t}$, $VoteHisp_{b,t}$ and $VoteAsn_{b,t}$ are the number of total votes, Hispanic votes and Asian votes in block b at election t and $\%Blk_{b,2010}$, $\%Wht_{b,2010}$, and $\%Oth_{b,2010}$ are the share of residents over age 18 who are Black, white, and other race from the 2010 Census. To account for racial differences in voter turnout rates, we weight by each racial group's turnout rate in California during election t as estimated by the CPS Voting and Registration Supplement ($\%VoteBlk_t$, $\%VoteWht_t$, and $\%VoteOth_t$).[18] Essentially, we weight

[17] The Passel-Word list has been shown to be more predictive of Hispanic ethnicity than directly-collected Medicare measures (Morgan *et al.*, 2004; Wei *et al.*, 2006).

[18] Due to the small sample size of the "other" racial group, we collapse CPS turnout rates for "other" into presidential and mid-term election averages.

non-Hispanic, non-Asian votes and registrations in a given block-election by each racial group's predicted vote and registration share relative to the other remaining racial groups.[19]

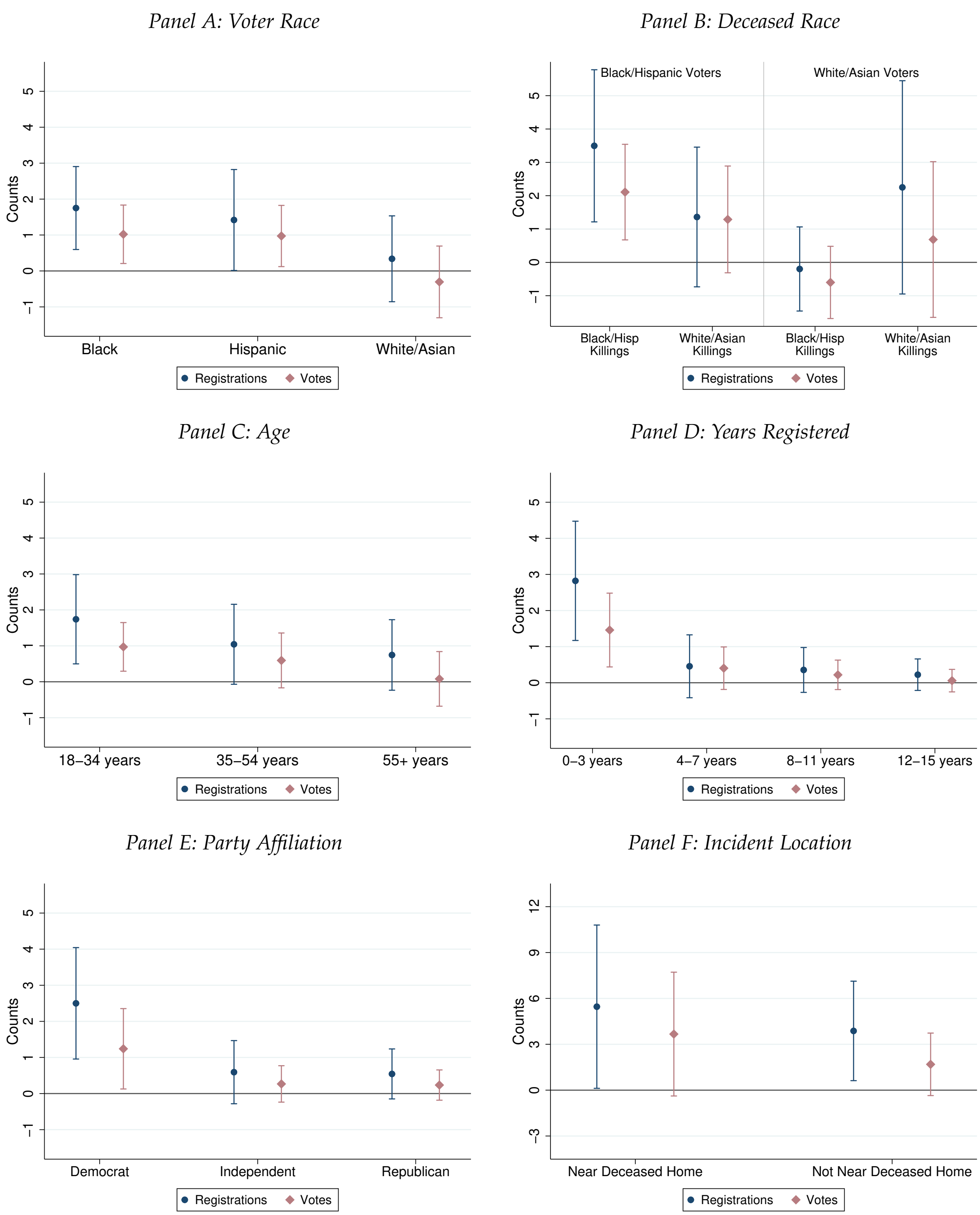
We then estimate our simplified version of Equation 3.2 on predicted counts by race. As shown in Panel A of Figure 3.3, a striking pattern emerges. Police killings lead to large increases in Black and Hispanic participation. On average, each police killing increases Black registrations by 1.7 and Black votes by 1.0. These estimates are highly significant and represent an 8 to 11% increase over the pre-killing mean (20.2 registrations and 9.2 votes). We find similar, if proportionally smaller, responses among Hispanics with increases of 1.5 registrations (5% of mean) and 1.0 votes (6% of mean). In contrast, we find no significant impact on White and Asian participation, with point estimates representing less than 2% of the mean.

Further disaggregating by race of the deceased in Panel B, we find suggestive evidence that racial concordance between the voter and the deceased may lead to larger effects. Among Black and Hispanic voters, point estimates for registrations (votes) are roughly 150% (60%) larger for killings of Black and Hispanic individuals than for killings of white and Asian individuals. Similarly, for white and Asian voters, we find positive, if noisy, estimates for killings of white and Asian individuals, but near zero estimates for killings of different-raced individuals.

That effects are primarily concentrated among Black and Hispanic citizens is consistent with a host of evidence documenting large racial differences in perceptions of law enforcement. Researchers have found that race is the single strongest predictor of trust in police (Taylor *et al.*, 2001) and Black and Hispanic individuals are far more likely than others to believe that use of force is excessive, unjustified or a serious social concern (Weitzer and Tuch, 2002; AP-NORC, 2015; Davis *et al.*, 2018).

[19] To address concerns with ecological inference (King, 2013), Figure A.III of the Online Appendix compares our race estimates for 2010 against estimates obtained by predicting individual race from surname and address in the full voter registration file and aggregating to the block-level (Consumer Financial Protection Bureau, 2014). Estimates are highly similar with mean differences near zero in both treatment and control areas.

Figure 3.3: *Heterogeneous Effects*



Notes: Panels report results from estimating Equation 3.2 separately for each voter group (i.e., by race, age, registration length and party affiliation), replacing the time to treatment indicators with a single post-treatment indicator. Panels B and F include distinct post-treatment indicators corresponding to each incident type (i.e., by deceased race in Panel B and by proximity to deceased home in Panel F). Standard errors are clustered by Census block group. Unit of observation is the Census block-election. The sample spans the 2002 to 2010 general elections and treatment is defined by blocks where police killings occurred during the sample period. Full regression results are included in Table A.V of the Online Appendix.

Voter Age, Years Registered, and Political Affiliation

In Panel C, we find that gains in voter participation are driven by younger individuals. Following police killings, registrations among nearby 18- to 34-year-olds increase by about 7% (pre-killing mean = 26.8), while votes increase by roughly 10% (pre-killing mean = 10.4). In contrast, treatment estimates for individuals over 35 years old are statistically insignificant and small in magnitude (less than 4% of the pre-killing mean). Consistent with this, Panel D demonstrates that increases in turnout come entirely from individuals who registered within 3 years of a given election. That point estimates for registration are near zero for longer-registered voters also provides evidence that the registration effects are not driven by differential migration (i.e. previously registered voters moving into treatment areas).

In Panel E, we show that effects are also concentrated among registered Democrats. We find no significant impact on registration or vote counts among Republicans or independents. These results are reflective of longstanding partisan gaps in views of law enforcement. Survey evidence from 1970 found that Democrats were more likely to oppose police use of force than Republicans (Gamson and McEvoy, 1970) and Democrats today remain much more skeptical of police accountability and discretion (Morin *et al.*, 2017).[20]

Family and Household Responses

To interrogate the possibility that effects are driven by family and household members of the deceased, we obtained full-count voter registration files from 2004 to 2010 and searched for new registrants who lived in the Census block of a killing and bore the same last name as the deceased. Across 204 killings that occurred in this window, we identified 25 total matches or roughly 0.12 new surname-match registrants per killing. While this is obviously an imperfect measure of kinship networks, it suggests that increased participation among family members of the deceased cannot account for the significant effects on voter participation, which average 3.5 new registrants in the surrounding block.

[20] For example, 27% of Democrats versus 74% of Republicans believe police do a good job "using the right amount of force."

As further corroboration, we examine whether effects differ depending on whether a police killing occurred near the home of the deceased. While we do not observe the exact home address of all individuals killed by police, we are able to infer whether a killing occurred near the deceased's residence based on DA incident report descriptions.[21] If increased participation was driven primarily by the deceased's household, then we would expect police killings that occurred far from the deceased's home to have little impact on registration rates in the block of the incident. However, as shown in Panel F of Figure 3.3, we find significant effects regardless of incident location. This suggests that less proximal individuals, such as witnesses, neighbors and other local community members – as opposed to family or household members – likely account for much of the mobilization effects.

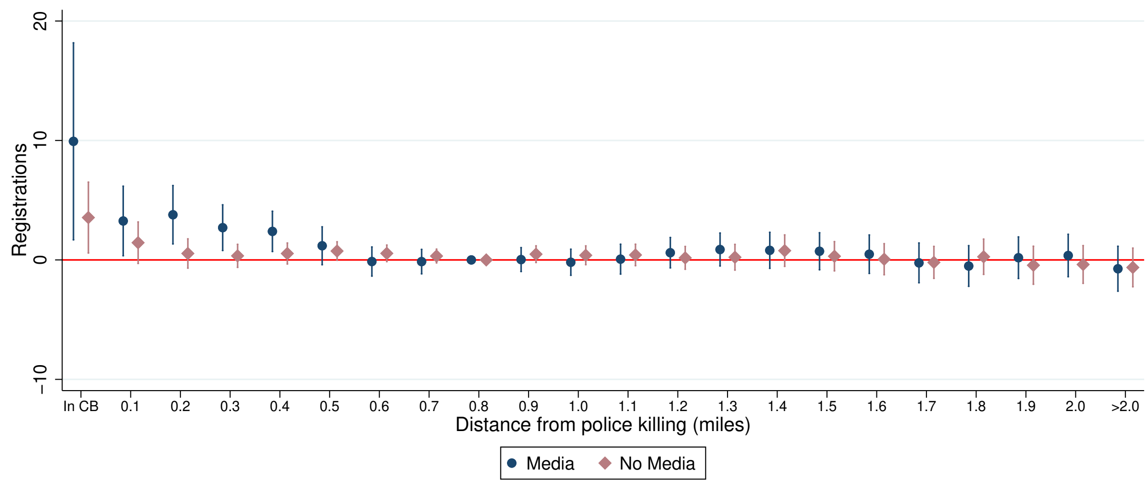
Role of Media

Given the recent proliferation of viral footage capturing acts of police violence, we interrogate the role of media in community responses to these events. Specifically, we examine how spatial spillovers vary by media coverage. To do so, we estimate separate distance gradients for killings that were and were not mentioned in local newspapers using Equation 3.1.

As shown in Figure 3.4, we find suggestive evidence of wider spread effects following police killings that were covered in the media than those that were not. While effects for unmentioned killings are contained to the immediate block, we find statistically significant effects extending to blocks nearly 0.50 miles away for killings mentioned in the media. While media coverage may simply reflect (rather than influence) community awareness or perceptions of an incident, these patterns may nonetheless help to explain the discrepancy between the national responses to recent high-profile police killings and the highly-localized effects we observe here.

[21] Of the 294 sample killings, 48 were identified as occurring in or outside the deceased's home.

Figure 3.4: *Effects by Media Coverage*



Notes: Figure reports coefficients from estimation of Equation 3.1 on registration counts, separately for killings that were and were not reported in local newspapers. Sample spans 2002 to 2010 general elections. Indicators are mutually-exclusive and track a Census block's minimum distance to police killings that occurred prior to a given election."In CB" refers to blocks where killings occurred. Other blocks are partitioned by 0.1-mile bins. Block group by election fixed effects and block group-clustered standard errors are included. Full regression results are included in Columns 3 and 4 of Table A.I of the Online Appendix.

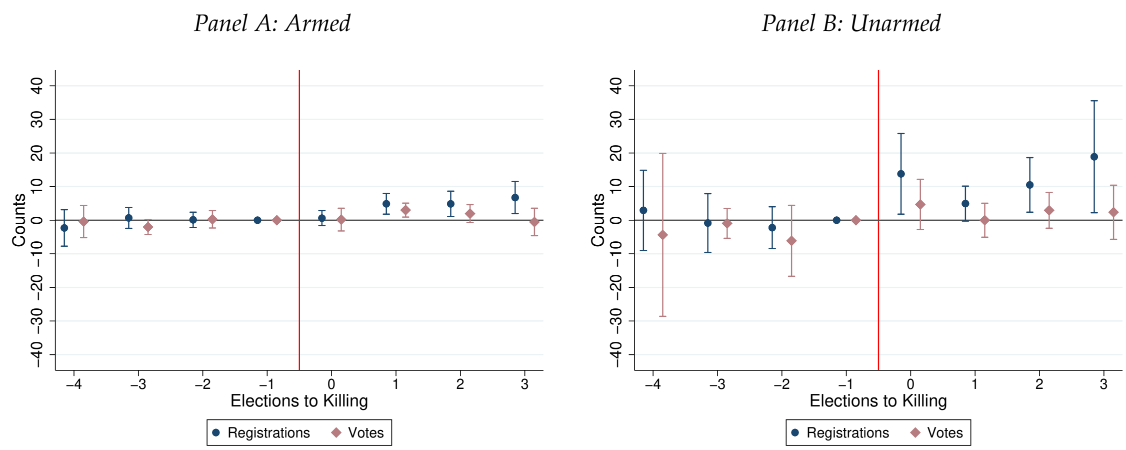
3.6 Mechanisms

While our results indicate that police violence may increase local civic participation, the motivation behind these responses is theoretically ambiguous. For example, if officer-involved killings cause citizens to perceive higher rates of local crime, changes in turnout could reflect support for more intensive policing (Cummins, 2009). On the other hand, these events may raise concerns about institutional discrimination or police accountability such that citizens are spurred to reform the criminal justice system.

To disentangle mechanisms, we examine differential effects based on whether the person killed by police possessed a weapon. If voters are motivated by heightened concerns about crime, we would expect larger effects following police killings of armed suspects, which likely involved more gunfire or individuals who posed greater danger to the community. If instead voters are mobilized by perceptions of government misconduct, we would expect the largest effects to stem from killings of unarmed individuals.

Figure 3.5 presents our event study results separately for police killings of armed and

Figure 3.5: *Effects by Deceased Weapon*



Notes: Figure shows treatment estimates and 95 percent confidence intervals from estimation of Equation 3.2 on registrations (pre-treatment mean = 81.6) and votes (pre-treatment mean = 42.9). Unit of observation is registrations/votes in a Census block-election. Panel A restricts treatment group to killings of individuals armed with a knife or gun. Panel B restricts treatment group to killings of individuals who were unarmed. Estimates are similar when including incidents with unknown weapon type. Standard errors are clustered by Census block group. The sample spans the 2002 to 2010 general elections and treatment is defined by blocks where police killings occurred during the sample period. Red vertical line represents time of treatment. Full regression results are included in Columns 2 and 3 of Table A.III of the Online Appendix.

unarmed individuals.[22] Notably, estimates are small or insignificant for police killings of individuals armed with a knife or a gun (Panel A), with average treatment effects of 2.7 registrations ($p = 0.042$) and 1.5 votes ($p = 0.133$).[23] However, police killings of unarmed individuals lead to large increases in participation (Panel B). The average treatment effects of 11.5 registrations ($p = 0.025$) and 5.0 votes ($p = 0.025$) correspond to nearly 15% of the pre-killing means and are three to four times as large as effects for armed killings.[24] That police killings of unarmed individuals generate such large relative spillovers suggests that voters are responding to the perceived "reasonableness" of officer actions as much as to the violence itself.

As corroboration, we test whether police killings affect support for criminal justice

[22]Figure A.IV of the Online Appendix shows similar results when excluding population controls.

[23]Average treatment effects are derived from regressions with a single post-treatment indicator in place of the full set of leads and lags.

[24]Similar relative effects exist when including the 14% of sample killings with unknown weapon status.

reforms using data on referenda voting. Specifically, we examine block-level vote shares for California Proposition 66 in 2004 and California Proposition 5 in 2008, both of which sought to reduce criminal penalties for lower-level offenses. While both propositions were narrowly defeated, they provide a local measure of policy preferences and potential insight into beliefs about law enforcement.[25]

We estimate the following difference-in-differences model:

$$y_{b,t} = \delta_b + \delta_t + \beta Treat_b \times Post_t + \epsilon_{b,t}, \tag{3.4}$$

where $y_{b,t}$ is the share of Yes ballots cast for Proposition 66 in 2004 and the share of Yes ballots cast for Proposition 5 in 2008. $Treat_b$ is an indicator for Census blocks that experienced a police killing between the 2004 and 2008 elections, while $Post_t$ is a 2008 indicator. To improve internal validity, the sample is restricted to blocks in treated block groups. Standard errors are clustered by block group.

As shown in Column 1 of Table 3.3, we find that support for criminal justice reform increased significantly in blocks that experienced a police killing relative to other blocks in the same neighborhoods. These effects represent a meaningful change in policy preferences – the 5.3 percentage point increase in pro-reform ballot share is equivalent to nearly 15% of the 2004 treatment mean.

Column 2 disaggregates this effect across armed and unarmed killings and finds significant gains in support for criminal justice reform following both types of events. However, point estimates for unarmed killings are roughly three times as large as those for armed killings (0.115 versus 0.041). This is consistent with the differential effects on registration and turnout, and suggests that acts of police violence that appear less justified may provoke more skepticism of the criminal justice system.

Given that this analysis relies on only two elections, we are unable to examine pre-trends. To address lingering validity concerns, we instead conduct a placebo test examining how support for criminal justice reform changed in *future* treatment areas. That is, we

[25] Proposition 66 failed by a 52.7% to 47.3% margin and Proposition 5 by a 59.5% to 40.5% margin.

Table 3.3: *Effects on Support for Criminal Justice Reform*

	DV = Support for Criminal Justice Reform (%)			
	Actual Treatment		**Placebo Treatment**	
	(1)	(2)	(3)	(4)
Treat x Post	0.053***		-0.004	
	(0.019)		(0.021)	
x Armed		0.041*		0.020
		(0.021)		(0.033)
x Unarmed		0.115**		-0.019
		(0.053)		(0.072)
Mean	0.397		0.422	
Killings	2004-2008		2008-2010	
Obs.	2,614		1,302	

Table reports results from estimation of Equation 3.4 on the fraction of ballots cast in favor of referenda that would have weakened penalties for lower-level offenses (i.e. California Proposition 66 in 2004 and California Proposition 5 in 2008). Data are restricted to Census block groups that experienced a killing between 2004-2008 in the "Actual Treatment" sample and that experienced a killing between 2008-2010 in the "Placebo Treatment" sample. Treatment is defined by blocks where a police killing occurred. Columns 1 and 3 report changes in pro-reform support between 2004-2008 for blocks with actual (placebo) police killings, while Columns 2 and 4 interact the treatment (placebo) indicator with indicators for whether the deceased was armed or unarmed. Standard errors are clustered at the block group level. Mean support for treatment and placebo blocks in 2004 are listed. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.10$.

estimate Equation 3.4 for referenda voting in 2004 and 2008 on the sample of neighborhoods that experienced a police killing between 2008 and 2010. The placebo treatment group is comprised of blocks treated after 2008, while the control group is limited to untreated blocks from those same block groups.

These results are shown in Columns 3 and 4. Notably, placebo treatment estimates are insignificant and very near zero in all cases. This is analogous to support for parallel pre-trends and suggests that our actual treatment estimates reflect the impact of police killings as opposed to other confounds correlated with the location of those events.

Taken together, these results indicate that civic responses to police violence are driven by individuals opposed to law enforcement actions. Our findings suggest that these individuals may be mobilized by killings that appear the least justified and may seek to reform the criminal justice system, at least partly, through the electoral process.

3.7 Conclusion

This paper documents the causal impact of police killings on local political participation. We find that acts of extreme police violence significantly increase voter registration and turnout among nearby residents. These effects are driven by new registrations among historically under-enfranchised groups – young Black and Hispanic individuals. Strikingly, gains in civic engagement are largest following police killings of unarmed individuals and are accompanied by increased support for criminal justice reforms. Together, our results add to growing evidence of the social consequences of police use of force (Bor *et al.*, 2018; Legewie and Fagan, 2019b; Ang, 2021b; González and Prem, 2020).

The direction of our findings differs from research documenting the demobilizing impact of direct and familial criminal justice contact. One explanation for this divergence is that our results do not appear to be driven by family or household members of the deceased – those individuals who may bear the greatest personal and economic loss from a police killing. Given the public nature of many police shootings, it seems instead that concerned residents, witnesses, and other, less proximal individuals are responsible for the gains in turnout. This is consistent both with work by Walker (2020) suggesting that vicarious police contact can mobilize groups that feel targeted by law enforcement and with research by Lerman and Weaver (2014b) showing that the effects of policing on local citizen engagement can vary – even directionally – according to the intensity and nature of police-civilian interactions.

With the proliferation of social media and the advent of the Black Lives Matter movement in 2013, an important question is whether our results generalize to more recent years. While this can only be answered with further research, there are several reasons to expect qualitatively similar effects today. First, in Figure A.V of the Online Appendix, we examine whether registration effects differ between more and less recent police killings. Notably, local registration effects are positive and remarkably stable in magnitude over the extended sample period (i.e., 2002 to 2016). Second, we presented evidence that police killings covered in local newspapers produced wider-spread – but directionally-similar – effects relative to unpublicized incidents. This suggests that growing public attention to police violence may lead to the propagation of similar mobilization effects at the state- or national-level, as opposed diverging local and regional responses. Third, data indicate that Joe Biden's 2020 presidential victory was fueled partly by record turnout among minority youth, the same demographic that is most likely to support the Black Lives Matter movement and for which we observe the largest responses to police violence (Tufts College, 2020; Pew Research Center, 2020).

Considering whether this paper's results may extend to other cities is a similarly important and speculative exercise. Given Los Angeles' particular history with police brutality as well as evidence of the mediating role of perceived injustice, officer-involved killings in areas with less fraught police-community relations may not incite the same electoral responses. On the other hand, the racial and demographic pattern of effects we observe are mirrored in national surveys examining group perceptions of law enforcement. Thus, while specific electoral contexts may differ across cities, significant concerns about police accountability and use of force are shared by minority communities throughout the country.

Together, these findings highlight the pivotal role that law enforcement and social justice concerns may play in shaping civic engagement and provide empirical complement to longstanding concerns about race and policing (Weitzer and Tuch, 2006; Kirk and Papachristos, 2011b; Tyler *et al.*, 2014). In 1968, the Kerner Commission reported on the deep-rooted belief in a "double-standard of justice and protection" underlying widespread civil unrest. Our findings suggest that such beliefs continue to permeate communities of color today and are exacerbated by acts of police violence. They tell a nuanced story about Black and Hispanic citizens responding to these concerns by strategically engaging with formal electoral systems in an effort to hold institutions accountable.

At the same time, we caution that our estimates represent average effects and may fail to capture the wide range of responses that police violence may engender. Our findings do not rule out the possibility that some individuals are demobilized by police killings or that already disenfranchised groups are further disempowered. Rather, they tell us only that any demobilizing effects are outweighed by increased participation among other citizens. Indeed, that our results are driven by new registrations among younger voters suggests that individuals may grow numb to state violence over time. As fatal shootings comprise less than one-tenth of one per cent of all use of force encounters (Davis *et al.*, 2018), it is possible that the long-run consequences of repeated exposure to traumatic police encounters may be far different from the marginal effects we document here. Further interrogation of this and other questions is critical to understanding the role of law enforcement in democratic engagement and representation.[26]

[26] Research documentation and data that support the findings of this study will be openly available in the APSR Dataverse at https://doi.org/10.7910/DVN/EOIIEV.

References

Abrams, D. S., Fang, H. and Goonetilleke, P. (2021). Do cops know who to stop? assessing optimizing models of police behavior with a natural experiment. *Working paper.*

ACLU of Illinois (2015). *Investigative stop and protective pat down settlement agreement*. Chicago Police Department, ACLU of Illinois, City of Chicago.

ACLU of Pennsylvania (2011). *Bailey, et al. v. The City of Philadelphia, et al.* U.S. District Court for the Eastern District of Pennsylvania.

ACLU of Pennsylvania (2021). Federal judge orders new approach to eliminate racial bias in stop and frisk policing in philadelphia.

Agan, A. and Starr, S. (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *Quarterly Journal of Economics*, **133** (1), 191–235.

Agan, A. Y., Doleac, J. L. and Harvey, A. (2021). Misdemeanor prosecution. *No. w28600. National Bureau of Economic Research.*

Anderson, A. J. (2020). Racial profiling: Constitutional and statutory considerations for congress.

Ang, D. (2021a). The effects of police violence on inner-city students. *Quarterly Journal of Economics*, **136** (1), 115–168.

— (2021b). The effects of police violence on inner-city students. *Quarterly Journal of Economics*, **136** (1), 115–168.

—, Bencsik, P., Bruhn, J. and Derenoncourt, E. (2021). Police violence reduces civilian cooperation and engagement with law enforcement.

Angrist, J. D., Imbens, G. W. and Rubin, D. B. (1996). Identification of causal effects using instrumental variable. *Journal of the American Statistical Association*, **91** (434), 444–455.

Anwar, S. and Fang, H. (2006). An alternative test of racial prejudice in motor vehicle searches: Theory and evidence. *American Economic Review*, **96** (1), 127–151.

AP-NORC (2015). *Law Enforcement and Violence: The Divide Between Black and White Americans*. Tech. rep.

Arnold, D., Dobbie, W. and Yang, C. S. (2018). Racial bias in bail decisions. *The Quarterly Journal of Economics*, **133** (4), 1885–1932.

Arnold, W. S. D., David and Hull, P. (2020). Measuring racial discrimination in bail decisions. *No. w26999. National Bureau of Economic Research*.

Arrow, K. J. (1973). The theory of discrimination. *In Discrimination in Labor Markets, Ashenfelter, O. and A. Rees*, pp. 3–33.

Ayres, I. (2002). Outcome tests of racial disparities in police practices. *Justice Research and Policy*, **4** (1-2), 131–142.

Bacher-Hicks, A. and de la Campa, E. (2020). Social costs of proactive policing: The impact of nyc's stop and frisk program on educational attainment. *Working paper.*

Bateson, R. (2012). Crime victimization and political participation. *American Political Science Review*, pp. 570–587.

Becker, G. S. (1957). The economics of discrimination. *Chicago: University of Chicago Press.*

— (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, **76** (2), 169–217.

— (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy*, **101** (3), 385–409.

Bertrand, M., Duflo, E. and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, **119** (1), 249–275.

Bor, J., Venkataramani, A. S., Williams, D. R. and Tsai, A. C. (2018). Police killings and their spillover effects on the mental health of Black Americans: A population-based, quasi-experimental study. *The Lancet*, **392** (10144), 302–310.

Bordalo, P., Coffman, K., Gennaioli, N. and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, **131** (4), 1753–1794.

Boyd, R. W. (2018). Police violence and the built harm of structural racism. *Lancet*, **392**, 258–259.

Braga, A. A. and Bond., B. J. (2008). Policing crime and disorder hot spots: A randomized controlled trial. *Criminology*, **46** (3), 577–607.

Bronstein, N. (2014). Police management and quotas: Governance in the compstat era. *Columbia Journal of Law and Social Problems*, **48**, 543.

Buchanan, L., Bui, Q. and Patel, J. (2020a). Black lives matter may be the largest movement in U.S. history. *The New York Times*.

—, — and — (2020b). Black Lives Matter May Be the Largest Movement in U.S. History. *The New York Times*.

Caetano, G. S. and Maheshri, V. (2014). Identifying dynamic spillovers in criminal behavior. *Available at SSRN 2460952*.

Chalfin, A. and McCrary, J. (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature*, **55** (1), 5–48.

— and — (2018). Are US cities underpoliced? Theory and evidence. *Review of Economics and Statistics*, **100** (1), 167–186.

Chetty, H. N., Raj and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, **106**, 855–902.

Chetty, R., Hendren, N., Jones, M. R. and Porter, S. R. (2020). Race and economic opportunity in the united states: An intergenerational perspective. *Quarterly Journal of Economics*, **135**, 711–783.

Coate, S. and Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review*, pp. 1220–1240.

Cohen, E., Gunderson, A., Jackson, K., Zachary, P., Clark, T. S., Glynn, A. N. and Owens, M. L. (2019). Do officer-involved shootings reduce citizen contact with government? *Journal of Politics*, **81** (3), 1111–1123.

Consumer Financial Protection Bureau (2014). Using publicly available information to proxy for unidentified race and ethnicity: A methodology and assessment. *Washington, DC: CFPB, Summer*.

Craig, A. C. and Martin, D. C. (2019). Discipline reform, school culture, and student achievement. *Unpublished working paper*.

Cummins, J. (2009). Issue voting and crime in gubernatorial elections. *Social Science Quarterly*, **90** (3), 632–651.

Davis, E., Whyde, A. and Langton, L. (2018). Contacts between police and the public, 2015. *Bureau of Justice Statistics. US Department of Justice*.

Dept. of Justice (2016). *The United States of America v. The City of Ferguson*. U.S. District Court for the Eastern District of Missouri.

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), 857–876.

Di Tella, R. and Schargrodsky, E. (2004). "do police reduce crime? Estimates using the allocation of police forces after a terrorist attack.". *American Economic Review*, **94** (1), 115–133.

DiPasquale, D. and Glaeser, E. L. (1998). The Los Angeles riot and the economics of urban unrest. *Journal of Urban Economics*, **43** (1), 52–78.

Durlauf, S. N. and Heckman, J. J. (2020). An empirical analysis of racial differences in police use of force: A comment. *Journal of Political Economy*, **128** (10), 3998–4002.

Edwards, F., Lee, H. and Esposito, M. (2019). Risk of being killed by police use of force in the United States by age, race–ethnicity, and sex. *Proceedings of the National Academy of Sciences*, **116** (34), 16793–16798.

Enos, R. D., Kaufman, A. R. and Sands, M. L. (2019). Can violent protest change local policy support? Evidence from the aftermath of the 1992 Los Angeles riot. *American Political Science Review*, **113** (4), 1012–1028.

Eterno, J. A. and Silverman, E. B. (2019). *The crime numbers game: Management by manipulation*. CRC Press.

Evans, W. N. and Owens, E. G. (2007). COPS and Crime. *Journal of Public Economics*, **91** (1-2), 181–201.

Fan, J., Gijbels, I., Hu, T.-C. and Huang, L.-S. (1996). A study of variable bandwidth selection for local polynomial regression. *Statistica Sinica*, pp. 113–127.

Feigenberg, B. and Miller, C. (2021). Would eliminating racial disparities in motor vehicle searches have efficiency costs? *The Quarterly Journal of Economics*.

Fryer Jr, R. G. (2019a). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, **127** (3), 1210–1261.

— (2019b). An empirical analysis of racial differences in police use of force. *Journal of Political Economy*, **127** (3), 1210–1261.

Gamson, W. A. and McEvoy, J. (1970). Police violence and its public support. *The Annals of the American Academy of Political and Social Science*, **391** (1), 97–110.

Geller, A., Fagan, J. F., Tyler, T. and Link, B. (2014). Aggressive policing and the mental health of young urban men. *American Journal of Public Health*, **104** (12), 2321–2327.

Golden, M. and Almo, C. (2004). Reducing gun violence: An overview of New York City's strategies. *Vera Institute of Justice*.

Goncalves, F. and Mello, S. (2021). A few bad apples? Racial bias in policing. *American Economic Review*, **111** (5), 1406–41.

González, F. and Prem, M. (2020). Police Repression and Protest Behavior: Evidence from Student Protests in Chile. *Available at SSRN 3705486*.

Gutierrez, A., Ocampo, A. X., Barreto, M. A. and Segura, G. (2019). Somos Mas:Hhow racial threat and anger mobilized Latino voters in the Trump era. *Political Research Quarterly*, **72** (4), 960–975.

Harrell, E. and Davis, E. (2020). *Contacts Between Police and the Public, 2018 – Statistical Tables*. Bureau of Justice Statistics, U.S. Department of Justice.

Harris, E. A. (2015). Suspension rules altered in New York City's revision of school discipline code. *The New York Times*.

Hoekstra, M. and Sloan, C. (2020). Does race matter for police use of force? Evidence from 911 calls. *National Bureau of Economic Research*, **No. w26774**.

— and — (2021). Does race matter for police use of force? evidence from 911 calls. *American Economic Review*, p. [Forthcoming].

King, G. (2013). *A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data*. Princeton University Press.

Kirk, D. S. and Papachristos, A. V. (2011a). Cultural mechanisms and the persistence of neighborhood violence. *American journal of sociology*, **116** (4), 1190–1233.

— and — (2011b). Cultural mechanisms and the persistence of neighborhood violence. *American Journal of Sociology*, **116** (4), 1190–1233.

Knowles, N. P., John and Todd, P. (2001). Racial bias in motor vehicle searches: Theory and evidence. *Journal of Political Economy*, **109** (1), 203–229.

Knox, D., Lowe, W. and Mummolo, J. (2020). Administrative records mask racially biased policing. *American Political Science Review*, **114** (3), 619–637.

Knox, W. L., Dean and Mummolo, J. (2020). Administrative records mask racially biased policing. *American Political Science Review*, **114** (3), 619–637.

Lauderdale, D. S. and Kestenbaum, B. (2000). Asian American ethnic identification by surname. *Population Research and Policy Review*, **19** (3), 283–300.

Lee, H., Porter, L. C. and Comfort, M. (2014). Consequences of family member incarceration: Impacts on civic participation and perceptions of the legitimacy and fairness of government. *The ANNALS of the American Academy of Political and Social Science*, **651** (1), 44–73.

Legewie, J. and Fagan, J. (2019a). Aggressive policing and the educational performance of minority youth. *Sociological Review*, **84** (1), 220–247.

— and — (2019b). Aggressive policing and the educational performance of minority youth. *American Sociological Review*, **84** (1), 220–247.

Leiber, M. J., Nalla, M. K. and Farnworth, M. (1998). Explaining juveniles' attitudes toward the police. *Justice Quarterly*, **15** (1), 151–174.

Lerman, A. E. and Weaver, V. (2014a). *Arresting citizenship: The democratic consequences of American crime control*. University of Chicago Press.

— and — (2014b). Staying out of sight? Concentrated policing and local political action. *The ANNALS of the American Academy of Political and Social Science*, **651** (1), 202–219.

Levitt, S. D. (2004). Understanding why crime fell in the 1990s: Four factors that explain the decline and six that do not. *Journal of Economic Perspectives*, **18** (1), 163–190.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. and Sanbonmatsu, L. (2013). Long-term neighborhood effects on low-income families: Evidence from moving to opportunity. *American Economic Review*, **103** (3), 226–31.

MacDonald, J., Fagan, J. and Geller, A. (2016). The effects of local police surges on crime and arrests in new york city. *PLoS one*, **11** (6), e0157223.

Marx, P. (2018). An absolute test of racial prejudice. *The Journal of Law, Economics, and Organization*.

Mello, S. (2019). More COPS, less crime. *Journal of Public Economics*, **172**, 174–200.

— (2021). Fines and financial wellbeing. *Working paper.*

Morgan, R. O., Wei, I. I. and Virnig, B. A. (2004). Improving Identification of Hispanic Males in Medicare: Use of surname matching. *Medical Care*, pp. 810–816.

Morin, R., Parker, K., Stepler, R. and Mercer, A. (2017). Behind the badge. *Pew Research Center*, **11**.

National Academies of Sciences Engineering and Medicine (2018). *Proactive Policing: Effects on Crime and Communities*. The National Academies Press, Washington, DC.

Pantoja, A. D., Ramirez, R. and Segura, G. M. (2001). Citizens by choice, voters by necessity: Patterns in political mobilization by naturalized Latinos. *Political Research Quarterly*, **54** (4), 729–750.

Pew Research Center (2020). American trends panel, Wave 74.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, **62** (1), 659–661.

Rose, E. (2021). Who gets a second chance? Effectiveness and equity in supervision of criminal offenders. *The Quarterly Journal of Economics*, **136** (2), 1199–1253.

Roth, J. and Rambachan, A. (2021). An honest approach to parallel trends. *Working paper*.

Sønderskov, K. M., Dinesen, P. T., Finkel, S. E. and Hansen, K. M. (2020). Crime Victimization Increases Turnout: Evidence from Individual-Level Administrative Panel Data. *British Journal of Political Science*.

Taylor, T. J., Turner, K. B., Esbensen, F.-A. and Winfree, L. T. (2001). Coppin' an attitude: Attitudinal differences among juveniles toward police. *Journal of Criminal Justice*, **29** (4), 295–305.

Tiwara, A. (2019). Disparate-impact liability for policing. *The Yale Law Journal*, **129** (1).

Tufts College (2020). Election Week 2020: Young People Increase Turnout, Lead Biden to Victory.

Tyler, T. R., Fagan, J. and Geller, A. (2014). Street stops and police legitimacy: Teachable moments in young urban men's legal socialization. *Journal of empirical legal studies*, **11** (4), 751–785.

Vining, A. R. and Weimer, D. L. (2019). The value of high school graduation in the united states: Per-person shadow price estimates for use in cost–benefit analysis. *Administrative Sciences*, **9** (4), 81.

Walker, H. L. (2014). Extending the effects of the carceral state: Proximal contact, political participation, and race. *Political Research Quarterly*, **67** (4), 809–822.

— (2020). Targeted: The Mobilizing Effect of Perceptions of Unfair Policing Practices. *Journal of Politics*, **82** (1), 119–134.

Wasow, O. (2020). Agenda Seeding: How 1960s Black protests moved elites, public opinion and voting. *American Political Science Review*, pp. 1–22.

Weaver, V. and Lerman, A. E. (2010). Political consequences of the carceral state. *American Political Science Review*, pp. 817–833.

Wei, I. I., Virnig, B. A., John, D. A. and Morgan, R. O. (2006). Using a Spanish surname match to improve identification of Hispanic women in Medicare administrative data. *Health Services Research*, **41** (4p1), 1469–1481.

Weisburd, D., Wyckoff, L. A., Ready, J., Eck, J. E., Hinkle, J. C. and Gajewski, F. (2006). Does crime just move around the corner? a controlled study of spatial displacement and diffusion of crime control benefits. *Criminology*, **44** (3), 549–92.

Weitzer, R. and Tuch, S. A. (2002). Perceptions of racial profiling: Race, class, and personal experience. *Criminology*, **40** (2), 435–456.

— and — (2006). *Race and policing in America: Conflict and reform*. Cambridge University Press.

Western, B. and Pettit, B. (2010). Incarceration & Social Inequality. *Daedalus*, **139** (3), 8–19.

White, A. (2016). When threat mobilizes: Immigration enforcement and Latino voter turnout. *Political Behavior*, **38** (2), 355–382.

— (2019a). Family matters? Voting behavior in households with criminal justice contact. *American Political Science Review*, **113** (2), 607–613.

— (2019b). Misdemeanor disenfranchisement? The demobilizing effects of brief jail spells on potential voters. *American Political Science Review*, **113** (2), 311–324.

Zepeda-Millán, C. (2017). *Latino mass mobilization: Immigration, racialization, and activism*. Cambridge University Press.

Zimroth, P. (2016). Submission of second report of the independent monitor. *Southern District of New York*.

Zoorob, M. (2020). Do police brutality stories reduce 911 calls? Reassessing an important criminological finding. *American Sociological Review*, **85** (1), 176–183.

Appendix A

Appendix to Chapter 1

A.1 Robustness of Crime Responses to Police Surges

We probe the robustness of estimated crime effects of police surges in a number of ways. First, it is possible that Operation Impact displaces crime rather than deters it. To get a sense of geographic spillovers to nearby Census tracts, we explore the impact of becoming a "neighbor" of an Impact Zone on neighborhood crime rates in Appendix Table A.2. Columns 1 to 5 define treatment as neighborhoods that have $\leq$5% of their area covered by an Impact Zone and are within 0.25-miles of an Impact Zone. Event time is then defined in six-month periods relative to the time when neighborhood n switches from being outside a 0.25-mile radius of an Impact Zone to being within 0.25-mile radius. We include indicators that control for whether a neighbor is part of an Impact Zone during other assignment periods and whether non-neighbors are part of an Impact Zone. Coefficients on measures of neighborhood policing confirm that policing intensity is unaffected by nearby Impact Zone assignment. We find no evidence of spillover effects for major felonies, shootings, killings or non-violent misdemeanors and violations. However, there are small increases in non-major felonies and violent misdemeanors that are jointly significant at the 10% level. This may be the result of crime spillovers or be due to increased detection or crime reporting. In Columns 6 to 10, we repeat the same analysis except that treatment is defined as Census tracts that are within a 0.25-1.0 mile radius of an active Impact Zone (and were previously more than 1.0 mile away). Point estimates are close to zero and statistically insignificant, suggesting that crime does not spillover to neighborhoods situated more than 0.25 miles from an Impact Zone.

Given some evidence of spillovers to neighborhoods within a 0.25-mile radius of Impact Zones, Appendix Table A.3 reports results when redefining treatment as an indicator for whether a neighborhood is within a 0.25-mile radius of an active Impact Zone (including those that are partially covered by an Impact Zone). Compared to our main specification, point estimates are similar in direction but are smaller in magnitude. Post-assignment coefficients on major felonies are still around -10% of the pre-period mean and are jointly significant ($p = 0.017$). The same is true for non-major felonies and violent misdemeanors, which increase by about 4% of the pre-period mean ($p = 0.030$). Non-violent misdemeanors and violations increase initially by 3% only to taper off in years two and three. All other crime coefficients are small and jointly insignificant during the post-period.

Second, we redefine treatment as the first time a Census tract becomes part of Impact Zone (i.e. has $> 5\%$ of its area covered) during our study period. While this reduces the number of assignment events, it removes selection into treatment during the post-period, since event-time is no longer dependent on police management continuing to designate the tract as an Impact Zone after the initial six-month assignment period. Results are reported in Appendix Table A.1. Across most crime outcomes, point estimates are similar to our main results in terms of direction and magnitude, although standard errors are considerably larger. We interpret the similarity of these results as validation of our primary specification.

Finally, Appendix Figure A.5 tests robustness of effects on felonies and violent misdemeanors to parallel trends violations a la Roth and Rambachan (2021). Consistent with our trend-adjusted specification, we find that standard difference-in-difference estimates (given in blue) underestimate crime deterrence effects of Impact Zone assignment due to a positive treatment-control differences in pre-trend slopes. Once we allow for parallel trend violations, confidence sets around first-year estimates are centered around zero. Confidence intervals of second and third-year coefficients fall well below zero. Panels B and C reveal that the "break-down" slope for year two and year three estimates, respectively, is 0.35 and 0.45 crimes per 1mi^2/week. During the pre-period, Impact Zones display a differential slope of +0.26 crimes per 1mi^2/week, substantiating the claim that these police surges significantly reduced serious crime rates.

A.2 Estimating Effects on High School Dropout Rates using Student-level Variation

As an alternative approach for estimating the impact of the reform on high school dropout rates, we compare students living in neighborhoods that prior to the reform had substantially different stop rates. To this end, we assign each student-year observation in our main sample to a stop-exposure quartile based on the mean stop rate observed for the home Census tract of a given student during the training period (i.e., from 2007 to 2009). We then estimate the following difference-in-differences (DD) regression:

$$Y_{iny} = \sum_{q=3}^{4} \sum_{\tau \neq 0} \beta_{q\tau} I(q)_{\tau} + \delta_n + \delta_{g \times y} + \delta_{b \times y} + \Gamma X_i' + \epsilon_{iny} \tag{A.1}$$

Y_{iny} is the outcome indicator for individual i living in neighborhood n in year y. Note that this specification is the same as Equation 1.6 except that (*i*) quartile q is now defined at the tract-level (and not the school-level) based on where students live in a given school year, and (*ii*) school fixed effects are replaced with tract fixed effects. We also cluster standard errors at the tract-by-cohort level instead of at the school-by-cohort level. If students frequently encounter police outside of their home neighborhood and/or the school concentration of stop exposure is a particularly important mediator of observed effects (e.g., via peer effects), then this approach will underestimate reductions in dropout rates.

Results are reported in Table A.5. Relative to students living in low-exposure neighborhoods, students living in fourth (third) quartile neighborhoods experience 763 (258) fewer stops per year following the *Floyd* decision. These effect sizes are almost twice as large as our main “first-stage" estimates. However, it is important to note that tract-level differences likely overstate true differences in exposure, since students do not spend all of their time in their home Census tract. Using Equation A.1, we find evidence of substantial declines in both outcomes for students living in both third- and fourth-quartile neighborhoods, all of which are significant at the 1% level. Effects sizes are similar in magnitude to our main estimates, but are about half as large when controlling for linear pre-trend differences. DD coefficients are displayed in Figure A.10 and closely resemble DD coefficients from our main analysis. While we prefer the school-level analysis since it better captures stop exposure in a student's peer network, these findings confirm that our main findings persist when restricting the analysis to comparisons between students that experienced differently-sized shocks to stop rates in their home neighborhoods.

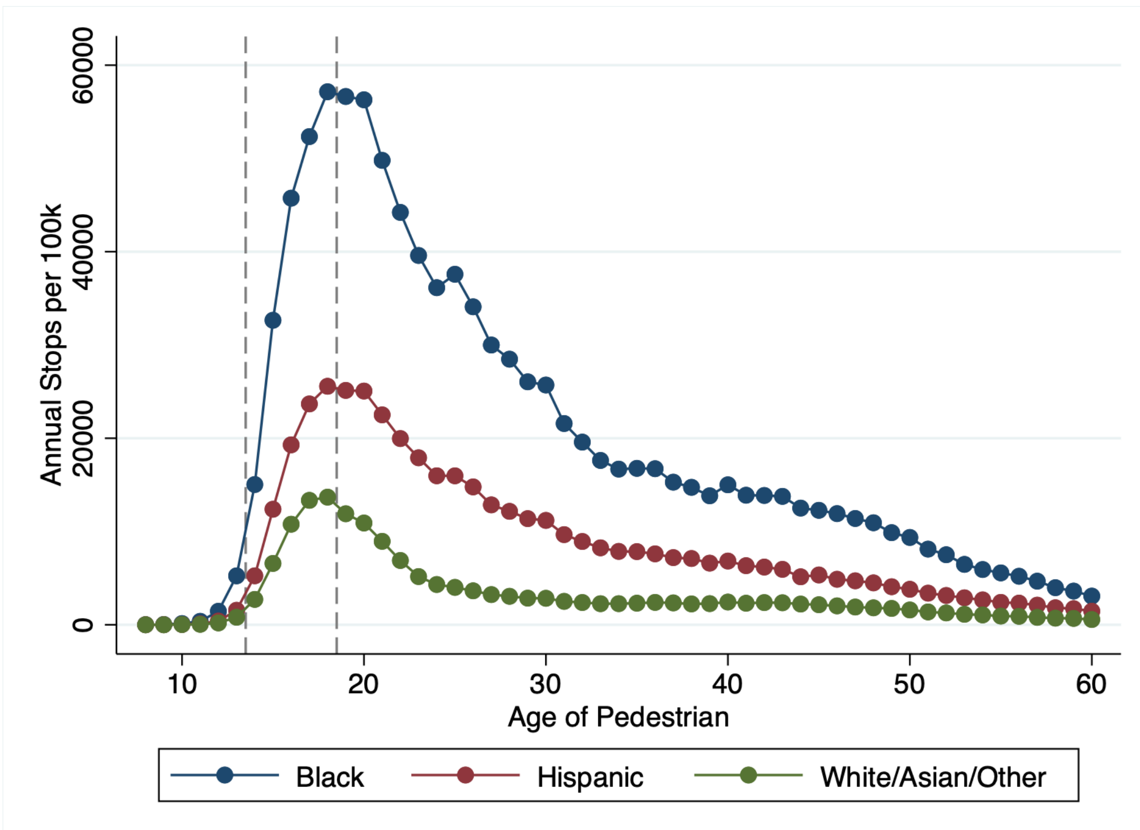
A.3 Supplemental Figures

Figure A.1: *Timeline of NYPD Stop and Frisk Policies*

Jan 1, 2003	The NYPD launches Operation Impact, deploying a majority of its Academy's graduating officers to twenty Impact Zones throughout the city.
Jan 31, 2008	*Floyd*: The Center for Constitutional Rights (CCR) files initial complaint.
Sept 10, 2008	*Floyd*: Court orders the City and NYPD to turn over to CCR all stop-and-frisk data for past decade.
Oct 17, 2011	Commissioner Kelly issues department-wide order that formalizes use of stop quotas in officer performance reviews.[1]
Nov 23, 2011	Judge Scheindlin grants motion to reconsider *Floyd* case in federal district court.
Apr 16, 2012	*Floyd*: Court mostly denies City's motion to exclude Dr. Fagan's expert testimony.
May 16, 2012	*Floyd*: Court allows "Class Action" status. Commissioner Kelly issues memo scaling back stops.
Jan 8, 2012	*Ligon*: Court grants plaintiffs' motion for a preliminary injunction ordering the NYPD to immediately cease its practice of unlawful trespass stops outside Clean Halls buildings in the Bronx.
Aug 12, 2013	*Floyd*: Judge Scheindlin finds the NYPD liable for a pattern and practice of racial profiling and unconstitutional stops, and orders broad reforms to be supervised by a court-appointed monitor.
Jan 30, 2014	*Floyd*: Newly-elected Mayor de Blasio accepts Court's ruling and withdraws appeals.

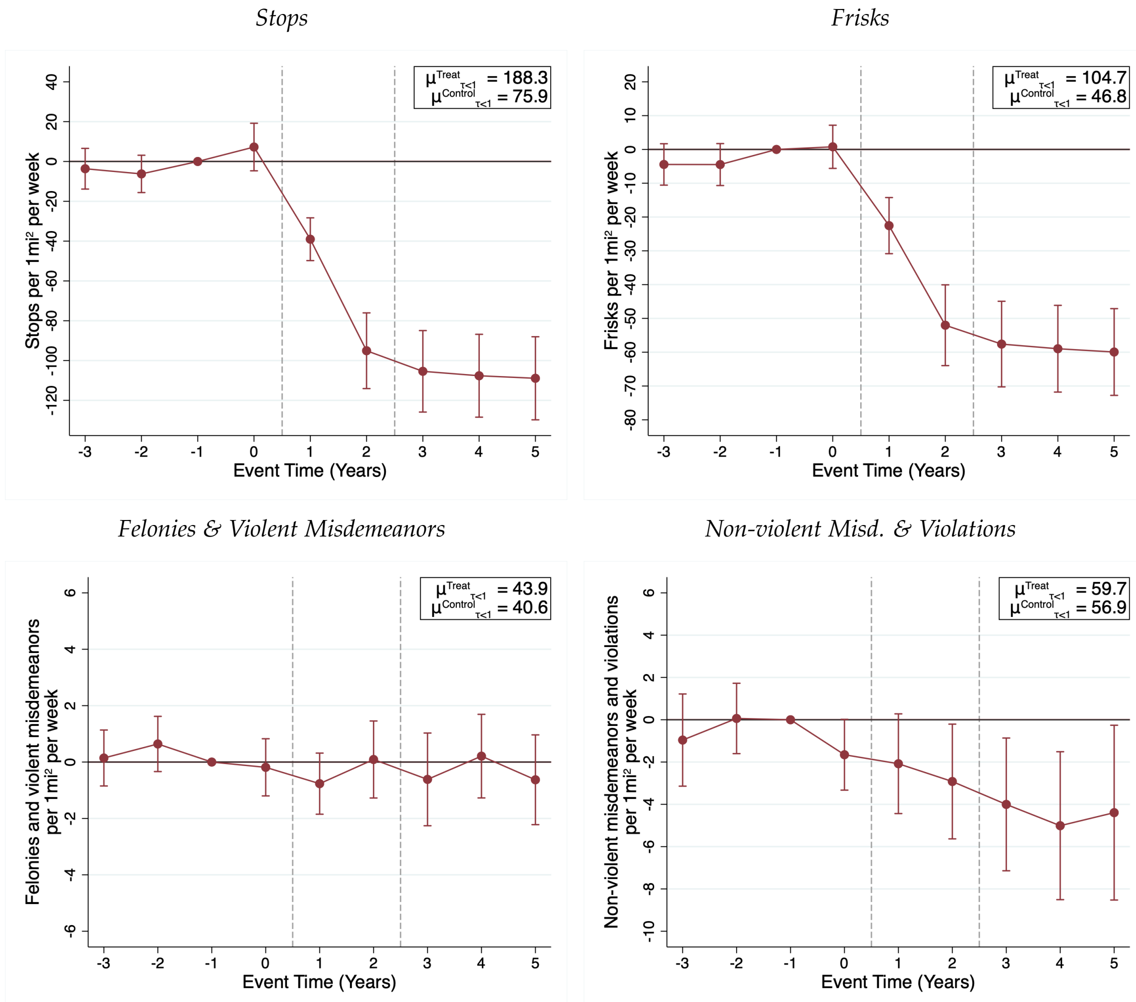
Notes: This figure displays a timeline of important changes to stop and frisk policy. *Floyd* and *Ligon* refer to *Floyd et al. v. The City of New York* and *Ligon et al. v. The City of New York*, respectively. Both lawsuits were filed in the United States District Court for the Southern District of New York and were presided over by Judge Shira A. Scheindlin.

Figure A.2: *Pre-Floyd Stop Rates by Age and Race*



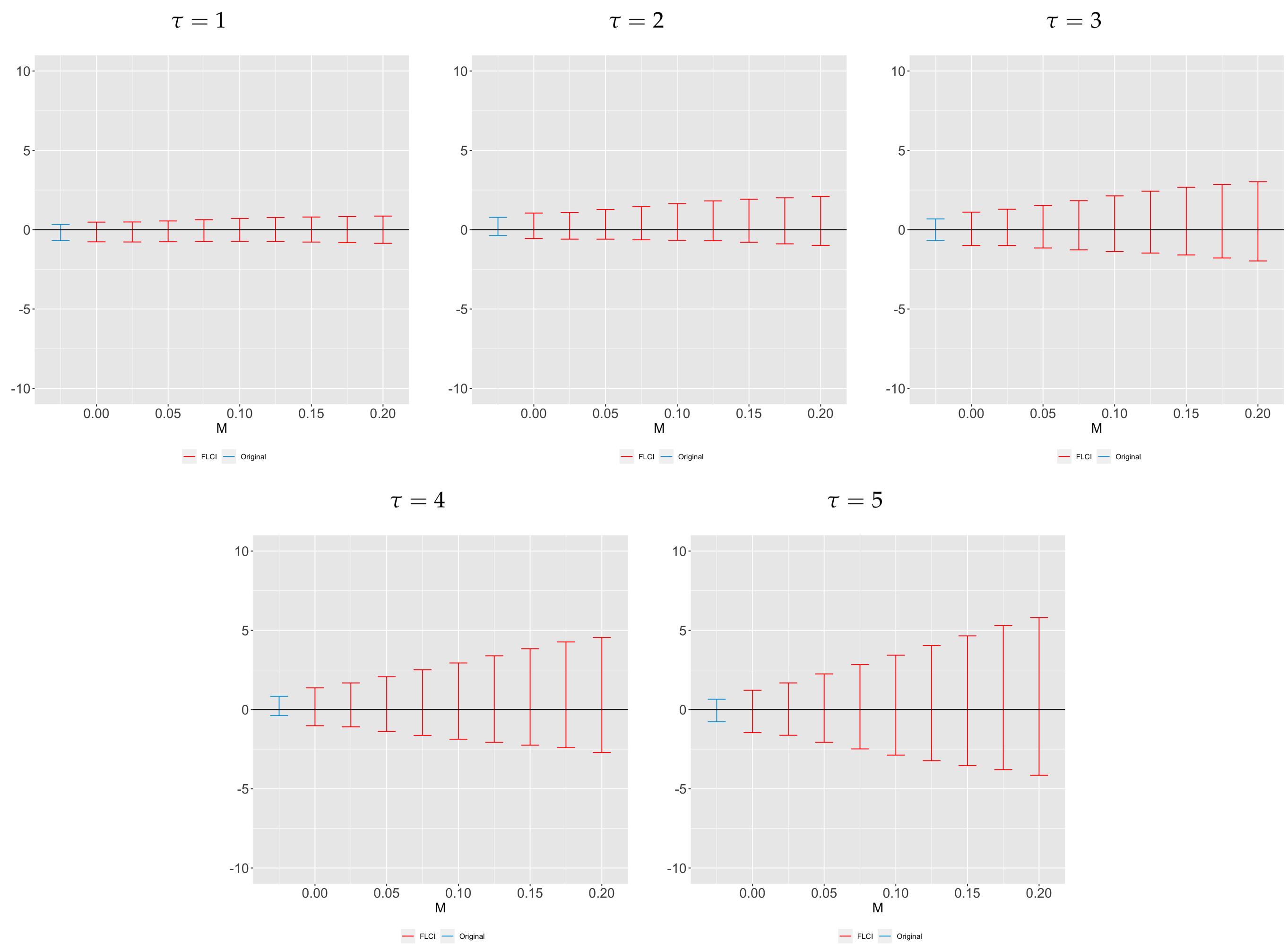
Notes: This figure displays stop rates from 2006-2011 by pedestrian age-by-race sub-groups. We use population estimates for New York City from Table SF1 of the 2010 Census, which includes population estimates for each age (under age 1 to 100) by race sub-group (e.g. All, Black Non-Hispanic, and Hispanic).

Figure A.3: *Floyd Difference-in-differences Estimates in Highest Crime Neighborhoods*



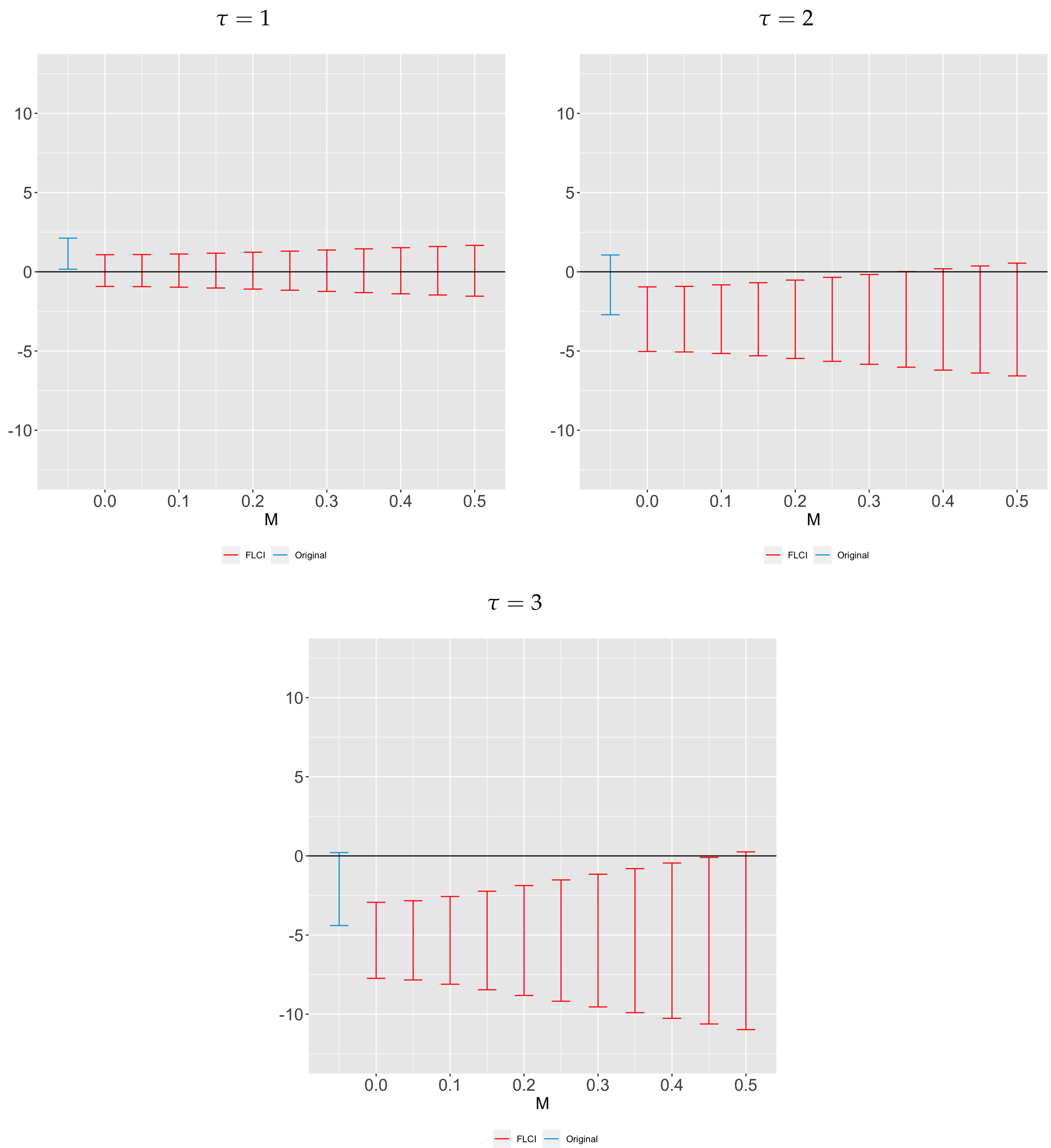
Notes: This figure graphs coefficients from Equation 1.3 on stops, frisks, felonies and violent misdemeanors, and non-violent misdemeanors and violations for neighborhoods that have above-median felony crime rates during the training period. Maroon dots denote point estimates and whiskers show 95% confidence intervals. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling. All standard errors are clustered at the precinct level. Pre-period outcome means are given for treatment and control neighborhoods in the northeast corner of each figure.

Figure A.4: *Sensitivity Analysis for DD Estimates of Floyd on Felonies and Violent Misdemeanors*



Notes: Figures display sensitivity analysis of estimated effects on felony and violent misdemeanor crimes (mean=22.98) to potential violations of the parallel trends assumptions per Rambachan and Roth (2021). The blue bar in each panel represents the 95% confidence interval of the difference-in-differences estimate for relative time τ from estimation of Equation 1.3 using precinct-clustered standard errors, except that we set the reference period to the year directly proceeding *Floyd* ($\tau = 0$). The red bars represent corresponding 95% confidence intervals when allowing for per-period violations of parallel trends of up to $\bar{M}$. That is, red confidence intervals impose that the change in the slope of the underlying trend be no more than M between consecutive periods.

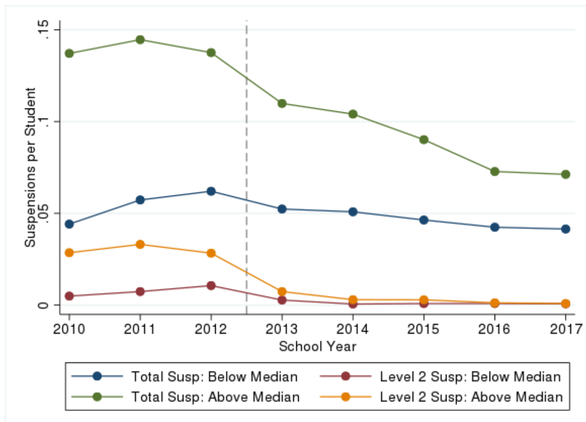
Figure A.5: *Sensitivity Analysis for DD Estimates of Operation Impact on Felonies and Violent Misdemeanors*



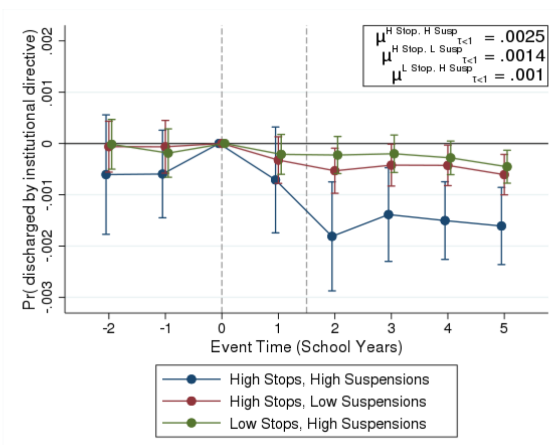
Notes: Figures display sensitivity analysis of estimated effects on felony and violent misdemeanor crimes (mean=49.52) to potential violations of the parallel trends assumptions per Rambachan and Roth (2021). The blue bar in each panel represents the 95% confidence interval of the difference-in-differences estimate for relative time τ from estimation of Equation 1.4 using precinct-clustered standard errors, except that we set the reference period to the year directly proceeding *Floyd* ($\tau = 0$). The red bars represent corresponding 95% confidence intervals when allowing for per-period violations of parallel trends of up to $\bar{M}$. That is, red confidence intervals impose that the change in the slope of the underlying trend be no more than M between consecutive periods.

Figure A.6: *Effect on School Discharges by Changes to School Stop Exposure and Suspension Policy*

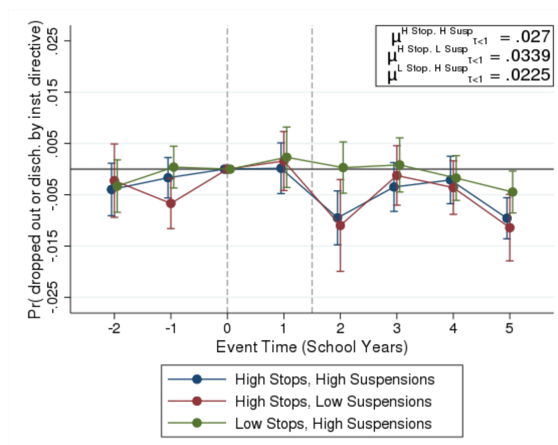
Suspensions per High School Student



Pr(Discharged by Institutional Directive)



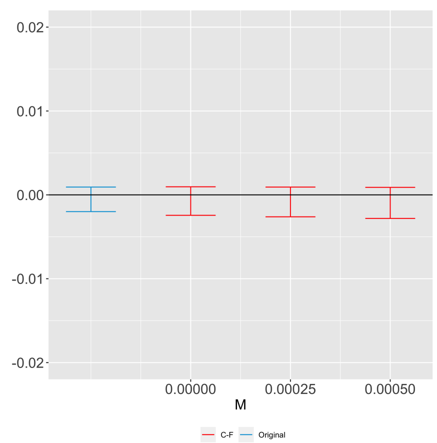
Pr(Dropped Out or Discharged by Inst. Dir.)



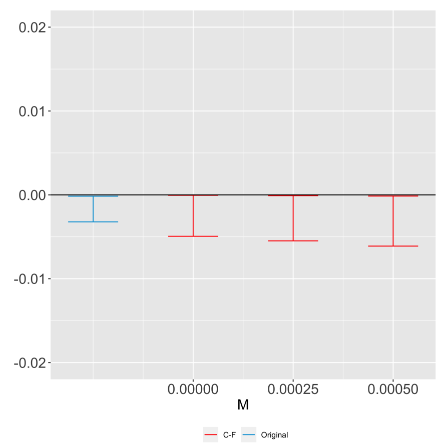
Notes: Panel A displays trends in total and level 2 "disorderly" suspensions for high and low suspension schools. High suspension schools are schools with above-median level 2 suspensions per student during the 2008 and 2009 school years. Panels B and C report coefficients from an adapted Equation 1.6 that collapses third and fourth quartile schools into a singular above-median stop-exposure indicator, and interacts this indicator with an above-median suspension indicator. The reference group is set to schools with below median stop-exposure and below-median suspensions. Whiskers mark 95% confidence intervals. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling. All standard errors are clustered at the school-cohort level.

Figure A.7: *Sensitivity Analysis for DD Estimates of Floyd on Institutional Discharge Rates*

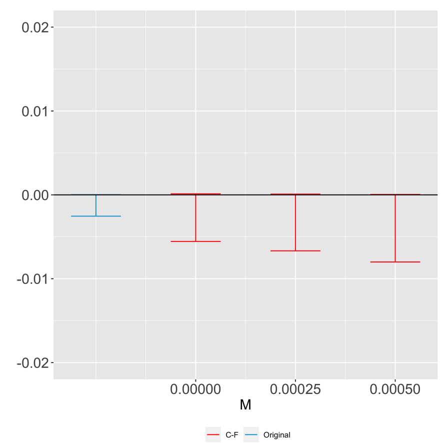
$\tau = 1$



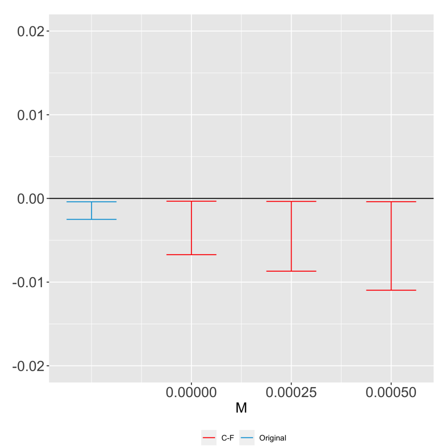
$\tau = 2$



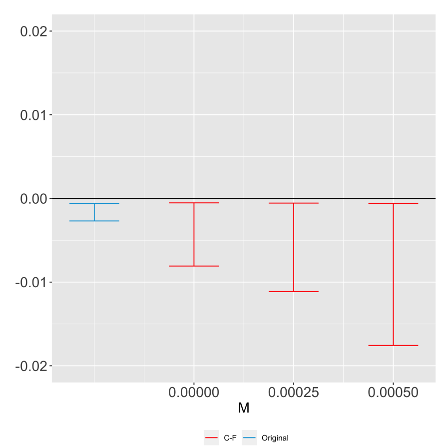
$\tau = 3$



$\tau = 4$

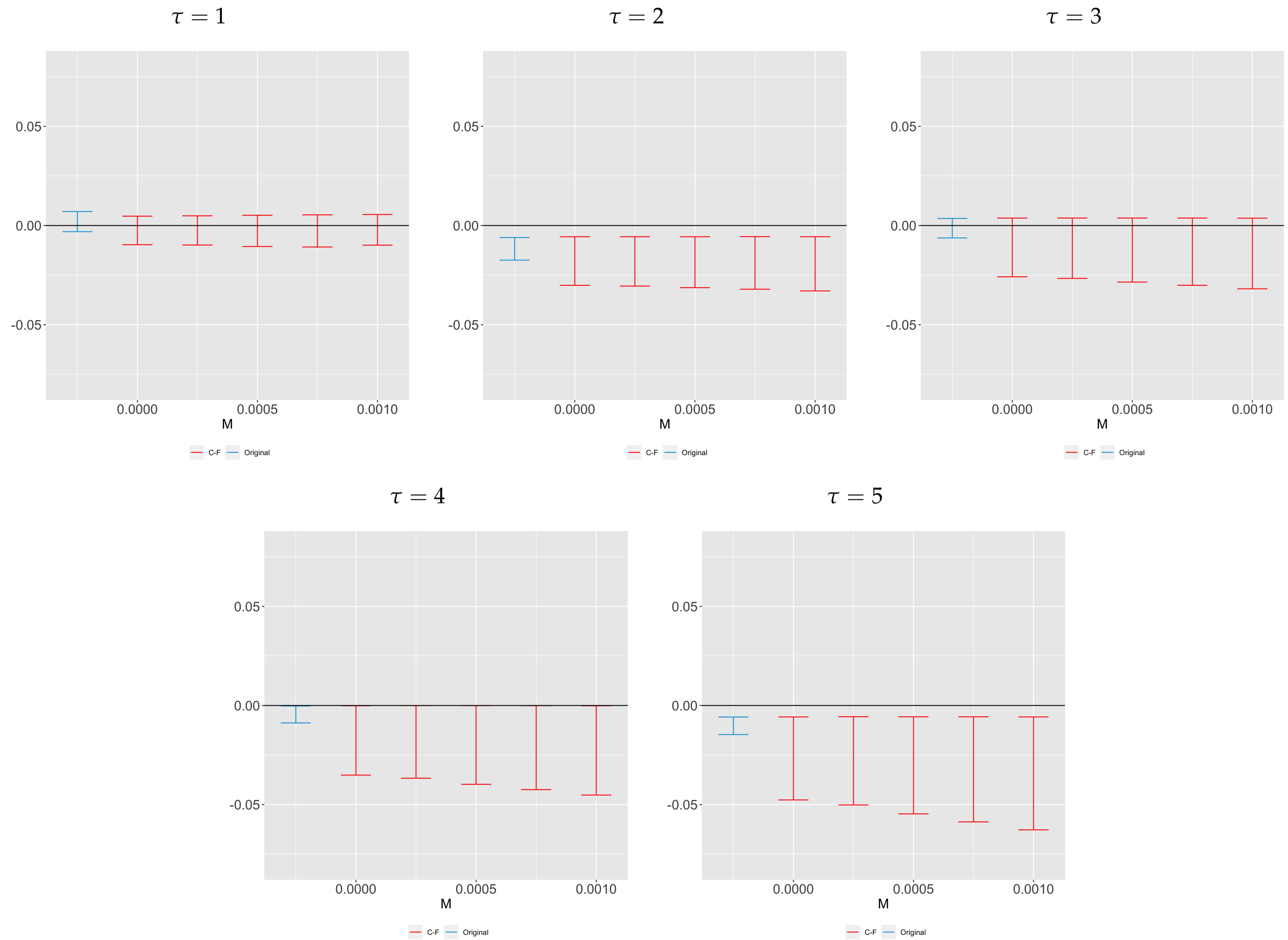


$\tau = 5$



Notes: These figures display sensitivity analysis of estimated effects on the likelihood a student is dicharged by institutional directive. Specifically, we explore potential violations of the parallel trends assumptions per Rambachan and Roth (2021), assuming that the pre-trend differences. The blue bar in each panel represents the 95% confidence interval of the difference-in-differences estimate for relative time τ from estimation of Equation 1.6. Standard errors are clustered at the school-cohort level. Red bars correspond to 95% confidence intervals that allow for per-period violations of parallel trends of up to $\bar{M}$. That is, red confidence intervals impose that the change in the slope of the underlying linear trend be no more than M between consecutive periods.

Figure A.8: *Sensitivity Analysis for DD Estimates of Floyd on Dropout Rates*

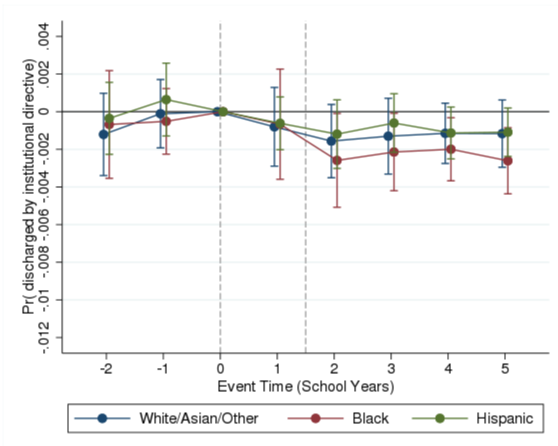


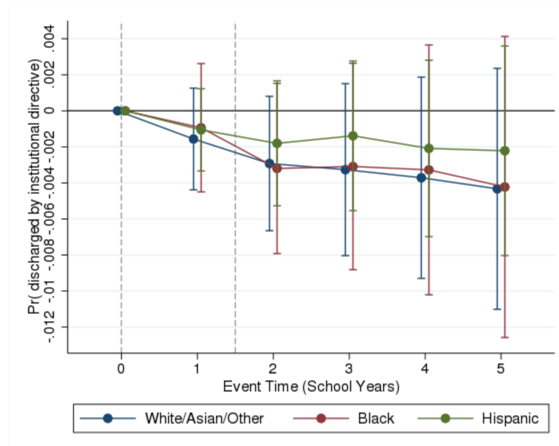
Notes: These figures display sensitivity analysis of estimated effects on the likelihood a student drops out or is dicharged by institutional directive. Specifically, we explore potential violations of the parallel trends assumptions per Rambachan and Roth (2021), assuming that the pre-trend differences lead to positive bias. The blue bar in each panel represents the 95% confidence interval of the difference-in-differences estimate for relative time τ from estimation of Equation 1.6. Standard errors are clustered at the school-cohort level. Red bars correspond to 95% confidence intervals that allow for per-period violations of parallel trends of up to $\bar{M}$. That is, red confidence intervals impose that the change in the slope of the underlying linear trend be no more than M between consecutive periods.

Figure A.9: *Effect on Dropouts and Institutional Discharges by Student Race*

DV = Pr(Discharged by Institutional Directive)

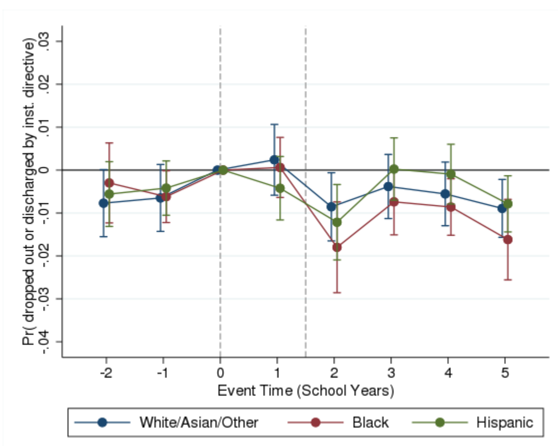
Difference-in-differences | *+ Linear Pre-trend Controls*

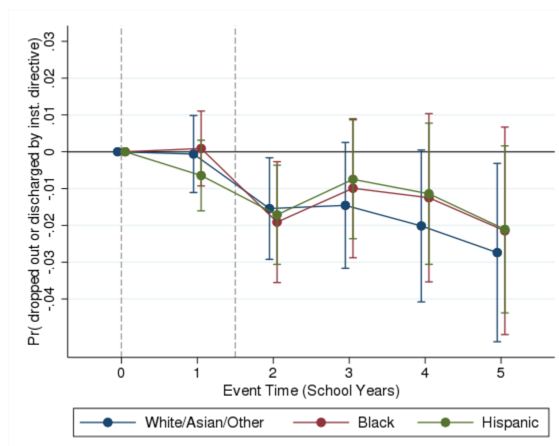




DV = Pr(Dropped Out or Discharged by Inst. Directive)

Difference-in-differences | *+ Linear Pre-trend Controls*

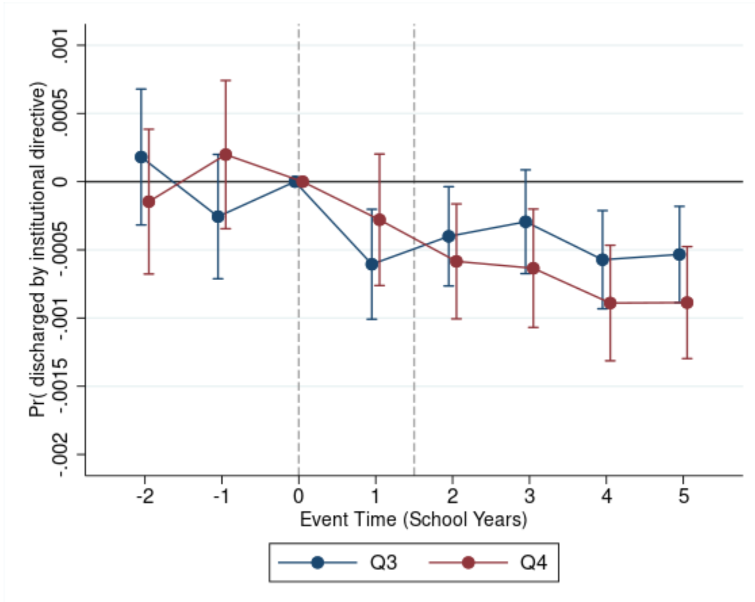




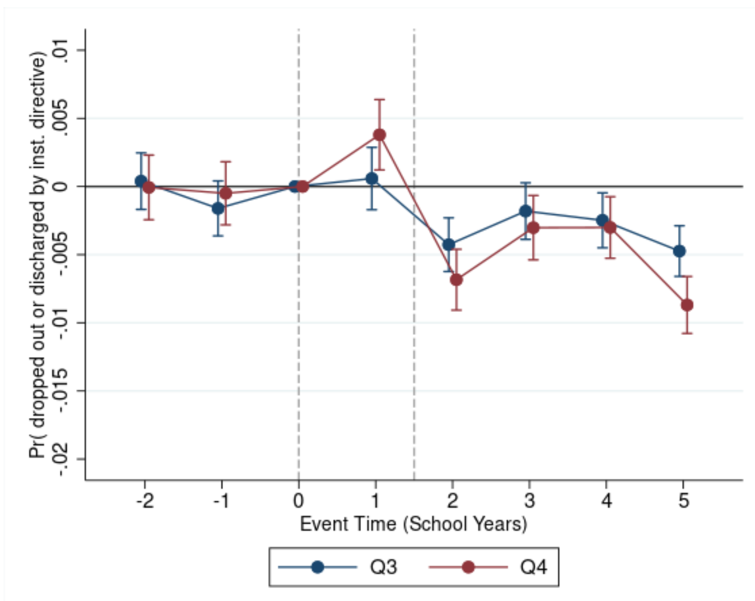
Notes: This figure graphs coefficients from Equation 1.6 in Panels A and C and coefficients from Equation 1.7 in Panels B and D, run separately for each racial sub-group. Only coefficients on Q4 time-to-treat indicators are included. Panels A and B display effects on the likelihood of enrolled students being discharged from high school by institutional directive during expected grades 9 through 12. Panels C and D display impacts on the likelihood of enrolled students dropping out or being discharged by institutional directive. The reference group is set to schools with below median exposure to stops during the training period. Estimates for white/Asian/other-race students, Black students, and Hispanic students are respectively reported in blue, maroon, and green. Whiskers mark 95% confidence intervals. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling. All standard errors are clustered at the school-cohort level.

Figure A.10: *Effect on High School Dropouts using Student-level Variation*

DV = Pr(Discharged by Institutional Directive)

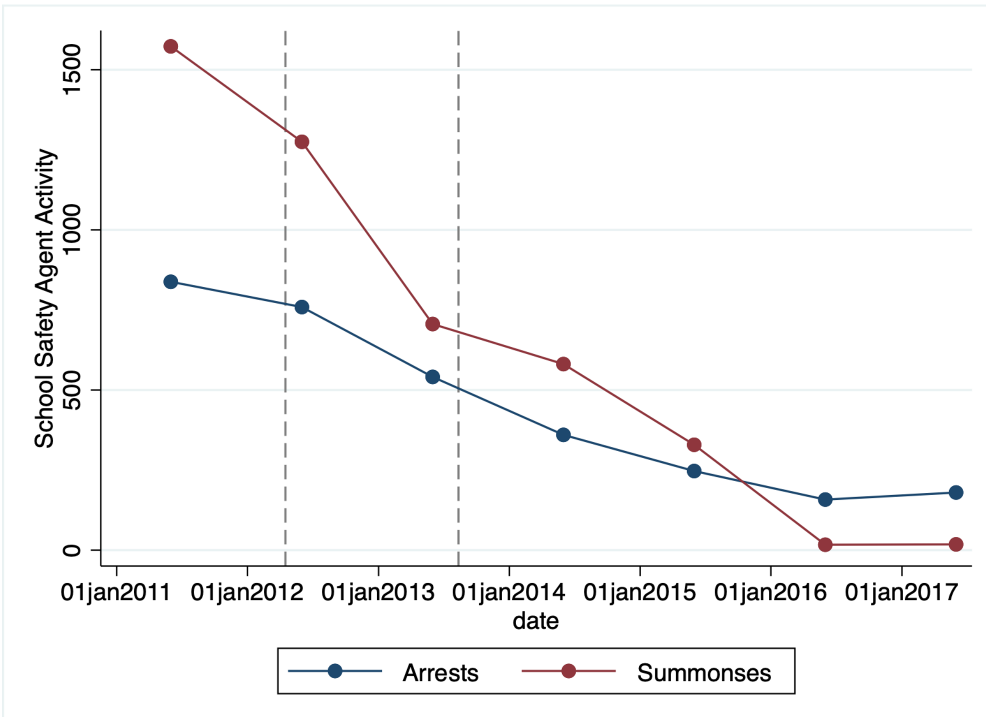


DV = Pr(Dropped Out or Discharged by Inst. Directive)



Notes: This figure graphs coefficients from Equation A.1. Panel A displays effects on the likelihood of enrolled students being discharged from high school by institutional directive during expected grades 9 through 12. Panels B displays impacts on the likelihood of enrolled students dropping out or being discharged by institutional directive. The reference group is set to students from tracts with below median exposure to stops during the training period. Blue dots refer to point estimates for students living in third-quartile neighborhoods, while maroon does denote point estimates for students living in fourth-quartile or the “most exposed" neighborhoods. Whiskers mark 95% confidence intervals. Moving from left to right, the first vertical dashed line marks when police began scaling down stops in response to *Floyd*. The second dashed line marks the final ruling. All standard errors are clustered at the tract-by-cohort level.

Figure A.11: *School Safety Agent Arrests and Court Summonses*



Notes: This table graphs the total number of arrests and court summonses issued by school safety officers in middle schools, high schools, and special education schools under the purview of the New York City Department of Education. Data come from the New York Civil Liberties Union and are available at `www.nyclu.org/en/student-safety-act-data`. Data are missing for the first quarter of 2011, so I impute 2011 totals by dividing the total number of stops observed in quarters two through four of 2011 by the $(1 - q_1)$, where q_1 is the mean fraction of yearly arrests (summonses) observed in the first quarter during the other five years with complete data.

A.4 Supplemental Tables

Table A.1: *Effect of Impact Zone Assignment using Years from First Assignment Event*

	Pre-period Mean	Year 1	Year 2	Year 3	P-value
Policing	(1)	(2)	(3)	(4)	(5)
Fraction of Tract Covered by Impact Zone	0.002	0.331***	0.247***	0.298***	0.000
		(0.033)	(0.034)	(0.047)	
Stops per 1mi^2/week	140.360	38.004***	33.773**	45.674**	0.003
		(7.964)	(14.576)	(19.393)	
Officers with a Stop in 30-day Span	23.195	5.082***	4.666***	6.015***	0.000
		(0.984)	(1.467)	(2.098)	
Mean Stops per Officer	1.494	0.075***	0.036	0.051	0.109
		(0.025)	(0.033)	(0.049)	
Frisks per 1mi^2/week	71.138	17.951***	22.867***	24.704***	0.001
		(4.724)	(7.906)	(9.047)	
Uses of Force per 1mi^2/week	28.128	6.763***	8.879***	8.347*	0.009
		(1.977)	(3.292)	(4.857)	
Stop-related Arrests per 1mi^2/week	6.736	0.035	0.172	-0.690	0.829
		(0.433)	(0.798)	(1.182)	
Stop-related Court Summonses per 1mi^2/week	9.886	3.509***	4.534***	5.030***	0.002
		(0.850)	(1.708)	(1.851)	
Crime					
Felonies & Violent Misd per 1mi^2/week	43.146	1.343	0.296	-1.004	0.878
		(0.990)	(1.453)	(1.867)	
Major Felonies per 1mi^2/week	20.024	-0.871	-1.010	-1.661	0.277
		(0.760)	(1.077)	(1.463)	
Non-major Fel & Violent Misd per 1mi^2/week	23.123	2.214***	1.306	0.657	0.081
		(0.540)	(0.900)	(1.116)	
Non-violent Misd & Violations per 1mi^2/week	52.885	3.171***	2.584	2.346	0.124
		(1.136)	(1.812)	(2.608)	
Shootings per 1mi^2/week	0.680	0.148*	0.238*	0.362*	0.071
		(0.086)	(0.136)	(0.206)	
Killings per 1mi^2/week	0.134	0.025	0.018	0.023	0.661
		(0.037)	(0.052)	(0.070)	
Observations (tract-weeks)		65,964			
Census Tracts		303			

Notes: This table reports coefficients from an adapted version of Equation 1.5 that defines event-time (τ) as six-month periods relative to a Census tract's initial assignment to an Impact Zone from January 2006 through July 2012. We collapse event-time into yearly indicators to improve precision. "Pre-period mean" refers to the outcome mean in treated neighborhoods in the three years leading up to assignment. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Treatment is defined as $> 5\%$ of Census tract area covered by an Impact Zone.

Table A.2: *Effect of Impact Zone Assignment on Neighborhood Policing and Crime by Distance from Impact Zone*

	$\leq$ 5% IZ Cov. & within 0.25 miles					0.25-1.0 miles from IZ				
	Pre-Mean	Yr 1	Yr 2	Yr 3	P-val	Pre-Mean	Yr 1	Yr 2	Yr 3	P-value
Policing	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Fraction of Tract Covered by Impact Zone	0.000	-0.004	-0.005	-0.010	0.370	0.000	-0.003	-0.002	-0.008	0.049
		(0.004)	(0.009)	(0.011)			(0.002)	(0.003)	(0.004)	
Stops per 1mi^2/week	90.435	-1.779	-3.395	4.530	0.963	52.847	0.535	-6.868	-11.639	0.225
		(2.947)	(5.376)	(7.426)			(3.138)	(5.857)	(6.908)	
Officers with a Stop in 30-day Span	16.491	0.217	0.263	1.965	0.253	10.470	0.343	-0.399	-1.041	0.538
		(0.365)	(0.815)	(1.185)			(0.373)	(0.730)	(0.803)	
Mean Stops per Officer	1.482	-0.010	-0.021	-0.006	0.509	1.406	-0.004	0.008	-0.029	0.648
		(0.011)	(0.022)	(0.029)			(0.014)	(0.025)	(0.023)	
Frisks per 1mi^2/week	52.315	-0.225	-1.395	5.913	0.666	29.073	0.220	-5.078	-5.436*	0.169
		(1.651)	(4.209)	(5.340)			(1.527)	(3.519)	(3.195)	
Uses of Force per 1mi^2/week	21.953	1.813	3.948	6.438	0.131	11.968	-0.686	-5.240*	-3.467**	0.032
		(1.457)	(2.927)	(4.317)			(0.819)	(2.638)	(1.554)	
Stop-related Arrests per 1mi^2/week	5.402	0.352	0.091	0.499	0.398	2.964	-0.145	0.099	0.013	0.966
		(0.330)	(0.459)	(0.562)			(0.191)	(0.289)	(0.463)	
Stop-related Court Summ per 1mi^2/week	6.741	0.420	-0.138	0.786	0.462	3.802	0.141	0.082	0.197	0.767
		(0.352)	(0.535)	(0.874)			(0.323)	(0.563)	(0.713)	
Crime										
Felonies & Violent Misd per 1mi^2/week	36.099	0.056	1.240	0.834	0.158	22.219	0.243	0.083	-0.087	0.854
		(0.334)	(0.731)	(0.905)			(0.286)	(0.505)	(0.667)	
Major Felonies per 1mi^2/week	16.860	-0.374	0.019	0.343	0.989	11.681	0.105	0.087	0.235	0.602
		(0.255)	(0.432)	(0.501)			(0.152)	(0.317)	(0.483)	
Non-major Fel & Violent Misd per 1mi^2/week	19.240	0.431*	1.221**	0.491	0.056	10.538	0.138	-0.004	-0.322	0.807
		(0.217)	(0.527)	(0.631)			(0.194)	(0.314)	(0.387)	
Non-violent Misd & Violations per 1mi^2/week	47.393	1.004*	1.056	0.057	0.294	29.907	0.128	-0.263	-0.644	0.704
		(0.552)	(0.660)	(1.419)			(0.495)	(0.821)	(0.969)	
Shootings per 1mi^2/week	0.497	-0.029	-0.018	-0.046	0.765	0.236	-0.041	-0.025	-0.063	0.591
		(0.066)	(0.115)	(0.157)			(0.047)	(0.093)	(0.110)	
Killings per 1mi^2/week	0.096	-0.002	0.028	-0.008	0.875	0.049	-0.004	-0.012	-0.012	0.769
		(0.022)	(0.041)	(0.058)			(0.020)	(0.034)	(0.044)	
Observations (tract-weeks)			84,417					126,522		
Census Tracts			452					666		

Notes: This table reports coefficients from Equation 1.5 for various policing and crime measures. We collapse event-time into yearly indicators to improve precision. "Pre-per. mean" refers to the outcome mean in treated neighborhoods in the three years leading up to assignment. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. We explore two separate definitions of treatment: tracts with $\leq$ 5% IZ coverage or within 0.25-miles of an active Impact Zone, and tracts within 0.25 to 1.00 miles from an active Impact Zone. SEs are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: *Effect of Impact Zone Assignment on Neighborhood Crime for Tracts within 0.25 Miles of Impact Zone*

	Pre-period Mean	Year 1	Year 2	Year 3	P-value
Policing	(1)	(2)	(3)	(4)	(5)
Fraction of Tract Covered by Impact Zone	0.001	0.116***	0.121***	0.114***	0.000
		(0.014)	(0.019)	(0.030)	
Stops per $1mi^2$/week	113.577	16.826***	22.912**	43.896***	0.000
		(4.122)	(9.585)	(12.137)	
Officers with a Stop in 30-day Span	19.350	2.041***	2.850***	5.108***	0.000
		(0.463)	(1.069)	(1.425)	
Mean Stops per Officer	1.498	0.015	0.013	0.035	0.194
		(0.011)	(0.016)	(0.029)	
Frisks per $1mi^2$/week	62.707	8.327***	16.127**	21.536***	0.002
		(2.841)	(6.144)	(7.540)	
Uses of Force per $1mi^2$/week	25.826	4.919**	11.147**	8.825**	0.028
		(2.329)	(5.091)	(4.426)	
Stop-related Arrests per $1mi^2$/week	6.213	0.478*	0.114	0.640	0.153
		(0.267)	(0.388)	(0.429)	
Stop-related Court Summonses per $1mi^2$/week	8.168	1.298**	2.399***	3.605***	0.000
		(0.532)	(0.872)	(1.287)	
Crime					
Felonies & Violent Misd per $1mi^2$/week	39.960	0.314	0.259	-0.324	0.877
		(0.325)	(0.700)	(0.780)	
Major Felonies per $1mi^2$/week	18.360	-0.675***	-0.832**	-0.899*	0.017
		(0.239)	(0.408)	(0.470)	
Non-major Fel & Violent Misd per $1mi^2$/week	21.600	0.989***	1.090**	0.575	0.030
		(0.269)	(0.521)	(0.606)	
Non-violent Misd & Violations per $1mi^2$/week	51.990	1.276**	0.992	0.165	0.456
		(0.627)	(1.233)	(1.586)	
Shootings per $1mi^2$/week	0.584	0.045	0.066	0.119	0.368
		(0.053)	(0.091)	(0.129)	
Killings per $1mi^2$/week	0.111	0.006	0.050	0.071	0.234
		(0.020)	(0.037)	(0.054)	
Observations (tract-weeks)		149,001			
Census Tracts		640			

Notes: This table reports coefficients from Equation 1.5 for various policing and crime measures. We collapse event-time into yearly indicators to improve precision. "Pre-per. mean" refers to the outcome mean in treated neighborhoods in the three years leading up to assignment. Standard errors are clustered at the precinct-level and reported in parentheses below coefficients. We define treatment as any Census tract within 0.25-miles of an active Impact Zone (including those covered by Impact Zones). SEs are clustered at the precinct-level and reported in parentheses below coefficients. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: *Effect of Floyd on Dropout and Institutional Discharge Rates by Pre-Reform School Suspension Rates*

	Pre-period Treat Mean	Base		+ Linear Pre-trends	
		$\beta_{\tau \geq 2}$	P-value	$E[\beta_{\tau \in [2,5]}]$	P-value
	(1)	(2)	(3)	(4)	(5)
Pr(Discharged by Institutional Directive)					
High Stops, High Suspensions	0.00252	-0.00118***	0.00049	-0.00254*	0.06175
		(0.00034)			
High Stops, Low Suspensions	0.00140	-0.00045***	0.00064	-0.00060	0.30049
		(0.00013)			
LowStops, High Suspensions	0.00096	-0.00022*	0.09561	-0.00026	0.63818
		(0.00013)			
Pr(Dropped Out or Discharged by Inst. Directive)					
High Stops, High Suspensions	0.02699	-0.00435***	0.00281	-0.01321**	0.02724
		(0.00145)			
High Stops, Low Suspensions	0.03386	-0.00386*	0.08457	-0.00872	0.26514
		(0.00224)			
LowStops, High Suspensions	0.02245	-0.00027	0.85216	-0.00788	0.18591
		(0.00142)			

Notes: Adapting Equations 1.6 and 1.7, this table reports post-ruling coefficients on interactions between an above-median stop-exposure indicator with above-median suspension exposure indicator. I report DD estimates for [High Stops, High Suspension] schools, [High Stops, Low Suspension] schools, and [Low Stops, High Suspension] schools relative to [Low Stops, Low Suspension] schools. Both indicators are measured at the school-level. Above-median suspensions are measured as mean Level 2 “disorderly" suspensions per high school student enrolled from 2008 to 2009 school years. Column 1 reports mean outcomes during 2010 to 2012 school years. Column 2 reports post-ruling impacts (i.e. years 2014 through 2017). Standard errors are reported below coefficients in parentheses. Column 4 reports the average across all four post-ruling coefficients, and Column 5 reports the p-value from an F-test that the four post-ruling coefficients are jointly zero. Standard errors are clustered at the school-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: *Effect of Floyd on Dropout Rates using Student-level Variation*

	Pre-period	Base		+ Linear Pre-trends	
	Treat Mean	$\beta_{\tau\geq 2}$	P-value	$E[\beta_{\tau\in[2,5]}]$	P-value
	(1)	(2)	(3)	(4)	(5)
DV = Stops per year					
Third Quartile	408.6499	-257.6284***	0.000	-245.6365***	0.000
		7.4320			
Fourth Quartile	923.8398	-763.1276***	0.000	-829.749***	0.000
		12.2966			
DV = Stops of 14-18 year-old residents per year					
Third Quartile	86.1784	-53.1894***	0.000	-58.9141***	0.000
		1.7777			
Fourth Quartile	194.2284	-159.5676***	0.000	-211.1002***	0.000
		3.4112			
DV = Pr(Discharged by Institutional Directive)					
Third Quartile	0.0015	-0.0004***	0.000	0.0000	0.970
		0.0001			
Fourth Quartile	0.0021	-0.0008***	0.000	-0.0011*	0.072
		0.0001			
DV = Pr(Dropped Out or Discharged by Inst. Directive)					
Third Quartile	0.0267	-0.0029***	0.000	-0.0020	0.403
		0.0005			
Fourth Quartile	0.0322	-0.0052***	0.000	-0.0054*	0.055
		0.0006			

Notes: This table reports coefficients from Equation A.1 and an adapted version that controls for linear differences in pre-trends. Treatment is defined as students from neighborhoods with stop exposure during the training period that ranked between the 50th and 75th percentile – "third quartile" – or 75-100th percentile – fourth quartile. Column 1 reports mean outcomes for each quartile during the 2010 through 2012 school years. Column 2 reports the coefficient on a post-ruling indicator that collapses 2014-2017 treatment indicators in Equation A.1 into a singular indicator. Standard errors are reported below coefficients in parentheses. Column 4 reports the average across all four post-ruling coefficients controlling for linear pre-trend differences, and Column 5 reports the p-value from an F-test that the four post-ruling coefficients are jointly zero. Standard errors are clustered at the tract-by-cohort level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B

Appendix to Chapter 2

B.1 Defining the Counterfactual Officer: A Simple Model of Officer Stopping Decisions

We model officer stopping decisions by adapting Anwar and Fang (2006). After observing pedestrian i, an officer decides whether or not to conduct a stop, $S_i \in \{0, 1\}$. The unobserved potential outcome of the encounter is given by $Y_i^* \in \{0, 1\}$, which denotes whether or not the encounter would lead to an arrest or the discovery of drugs or a weapon. Following the literature, the officer's objective is to stop pedestrians who are actively engaging in criminal activity ($Y_i^* = 1$) and not stop pedestrians who are not engaged in criminal activity ($Y_i^* = 0$).

Consider a continuum of patrol officers and a continuum of potential stops in neighborhood n. Let $R \in \{W, M\}$ denote the race of the pedestrian, where M stands for minorities, and W for whites. Suppose that fraction $\pi^R = E[Y^* = 1|R = r_i] \in (0, 1)$ of potential encounters with pedestrians of race r are engaging in criminal activity. The officer observes pedestrian race R_i and a noisy measure of the pedestrian's contemporaneous criminal activity, given by a uni-dimensional index $\theta_i \in [0, 1]$.

If the pedestrian engages in crime, then the index θ is randomly drawn from a distribution with continuous probability density function (PDF) $f_1^R(.)$ and cumulative distribution function (CDF) $F_1^R(.)$. If they are not, θ is randomly drawn from a continuous PDF $f_0^R(.)$

and CDF $F_0^R(.)$, where we allow signal distributions to vary by race.

We assume that $f_1^R(.)$ and $f_0^R(.)$ satisfy the strict standard monotone likelihood ratio property:[1]

$$\text{MLRP: } \frac{f_1^R(.)}{f_0^R(.)} \text{ is strictly increasing in } \theta. \tag{B.1}$$

Additionally, as in Anwar and Fang (2006), to the extent that there exist pedestrians who are clearly engaging in crime, we assume the unbounded likelihood ratio: $\frac{f_1^R(.)}{f_0^R(.)} \to +\infty$ as $\theta \to 1$. A higher θ informs the officer that encounter i is more likely to result in the detection of criminal activity.

Officers weigh the expected benefits against the costs of conducting a stop. Importantly, officers are only legally allowed to take into account the immediate threat that the civilian poses to society and not potential spillovers of a given encounter on crime or other societal repercussions.[2] Benefits thus include the social benefit of detecting and preventing immediate criminal activity as well as any career benefits for the officer.[3]. Interviews with NYPD officers suggest that career benefits to conducting target numbers of stops and frisks were particularly large.[4] Costs, on the other hand, include effort spent conducting the stop, the value the officer prescribes to the civilian's civil liberties, any anticipated harm to the officer or others due to the encounter, as well as other psychic concerns.

We normalize the potential benefit of the stop to equal one and set stopping costs as a fraction of the benefit, $C_R \in (0, 1)$. Observing θ_i and R_i, the officer forms a posterior risk

[1]Feigenberg and Miller (2021) show that this model can be easily extended to include regions of θ where the signal displays constant risk. The linearity of the search productivity curve at high levels of s_n suggests that cops may be unable to distinguish risk probabilities of low-risk individuals. That is, the MLRP may be constant for $\theta \leq \bar{\theta}$ from some low $\bar{\theta}$.

[2]A known criminal who is visibly not engaging in criminal activity cannot be legally stopped by police. The police must be able to articulate "reasonable suspicion" before making a stop.

[3]Chief Kelly's memo in October 2011 spells out the informal incentives that patrol officers faced throughout our study period: patrol supervisors were to evaluate officers based on their activity numbers, paying particular attention to both the number of stops and stop outcomes, such as court summonses and arrests. Officers whose numbers remained low would suffer serious disciplinary action.

[4]Bronstein (2014) writes that "if you were active (at making stops, frisks, and issuing court summonses), the Captain would pick you for the overtime."

prediction, $p_R(\theta_i)$, using Bayes' rule:

$$p_R(\theta) = P(Y^* = 1|\theta, R = r) = \frac{\mu^R f_{1,R}(\theta)}{\mu^R f_{1,R}(\theta) + (1 - \mu^R) f_{0,R}(\theta)} \tag{B.2}$$

From the MLRP, $\frac{dp_R(\theta)}{d\theta} > 0$. Thus, the officer's problem reduces to a choice of threshold – θ_R^* – for each racial group R, such that officers stop pedestrians if and only if $\theta_R \geq \theta_R^*$. They choose θ_R^* to maximize the expected payoff of stops subject to the capacity constraint.

$$Max_{\{\theta_R^*\}} \sum_{R \in \{M,W\}} \int_{\theta_R^*}^{1} (p_R(\theta) - C_R) f_R(\theta) d\theta \text{ such that } \sum_{R \in \{M,W\}} \int_{\theta_{R,n}^*}^{1} f_R(\theta) d\theta = \frac{s_n}{p_n} \tag{B.3}$$

where $f_R(\theta) = \mu^R f_1(\theta) + (1 - \mu^R) f_0(\theta)$ and $\frac{s_n}{p_n}$ is the stop rate per officer assigned to neighborhood n by management. We assume the constraint holds with equality, since officers were required to maintain stop rates ordered by police management and enforced by mid-level managers.

Officers thus set marginal benefits equal to marginal costs plus the shadow price of the constraint.

$$p_R(\theta_R^*) = C_R + \lambda_n \tag{B.4}$$

and are indifferent to stopping an additional civilian of either race:

$$p_R(\theta_W^*) - C_W = p_R(\theta_M^*) - C_M \tag{B.5}$$

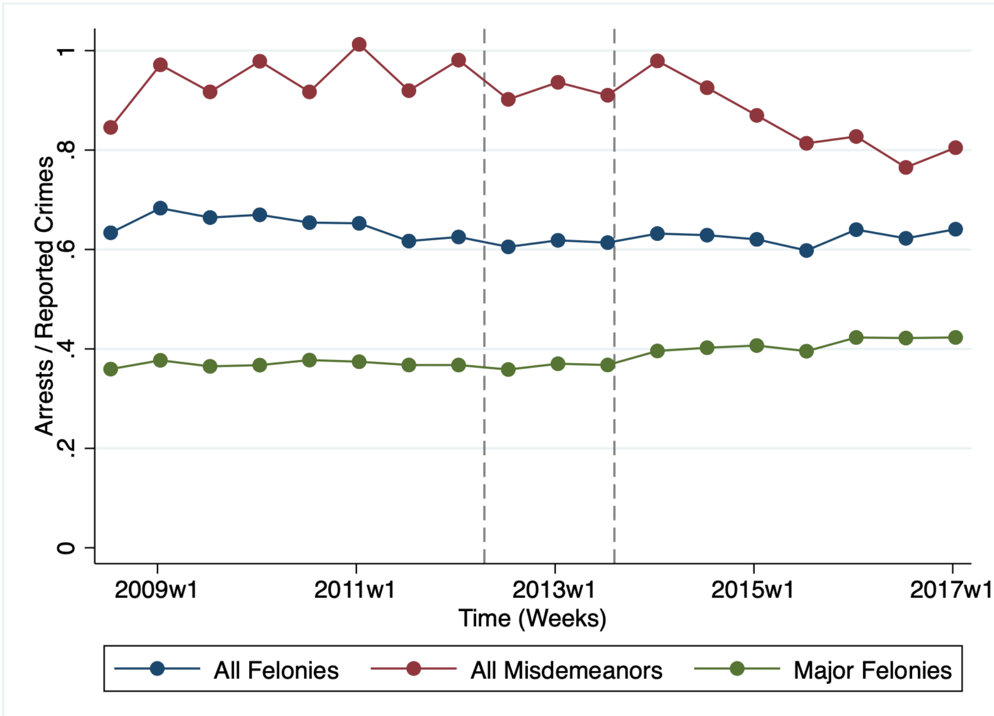
This model differs from Anwar and Fang (2006) in that officers may conduct stops with negative returns so as to carry out the stop rate desired by management in their neighborhood.

Following Becker, we define racial bias as having a preference for stopping one group over another. That is, if an officer is racially-biased against minority pedestrians, then $C_M < C_W$. Then, from Equation 2.2 it is clear that marginal stops of white pedestrians will be more likely to detect criminal behavior than marginal stops of minority pedestrians – $p_R(\theta_W^*) > p_R(\theta_M^*)$. This is the standard outcomes test result.

As Arnold *et al.* (2018) point out, racial bias can also arise from differences in racially-biased prediction errors, where even if $C_M = C_W$, officers that overestimate the likelihood that minority pedestrians are committing crime, will act as if they were taste-based discriminators. Then, their posterior risk predictions about minority pedestrians are given by $\pi_M(\theta) = p_M(\theta) + \tau_M(\theta)$, where $\tau_M(\theta) > 0$ is a racially-biased prediction error. Finally, it is important to note that taste-based discrimination could also arise entirely through top-down orders from police management (and have nothing to do with line officers' preferences) if management requires officers to over-stop pedestrians in higher-crime neighborhoods that are comprised of predominantly racial minorities. Then, marginal hit rates in these neighborhoods would be substantially lower than in other (more white) neighborhoods.

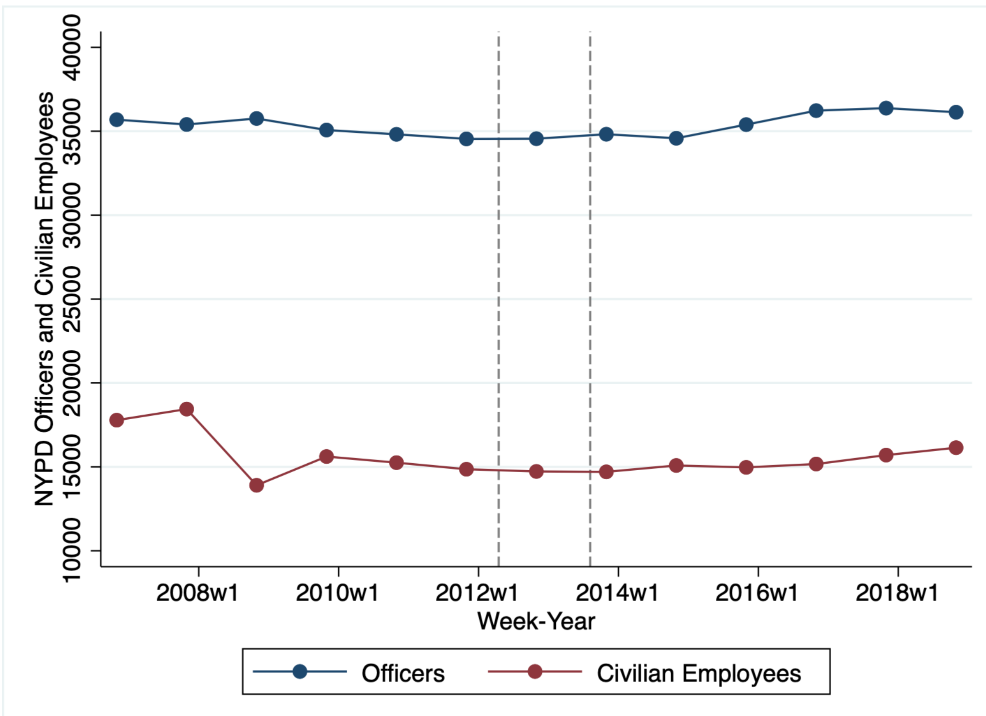
B.2 Supplemental Figures

Figure B.1: *Estimated Clearance Rates by Offense Category*



Notes: This table provides approximates clearance rates for felonies, misdemeanors, and major felonies across New York City. To do so, we divide the total number of arrests in a given offense category by the total number of reported crimes for that same category. Reported crimes and arrest data are not linked, so we use match on date, even though arrests for a given crime likely lag behind reporting dates.

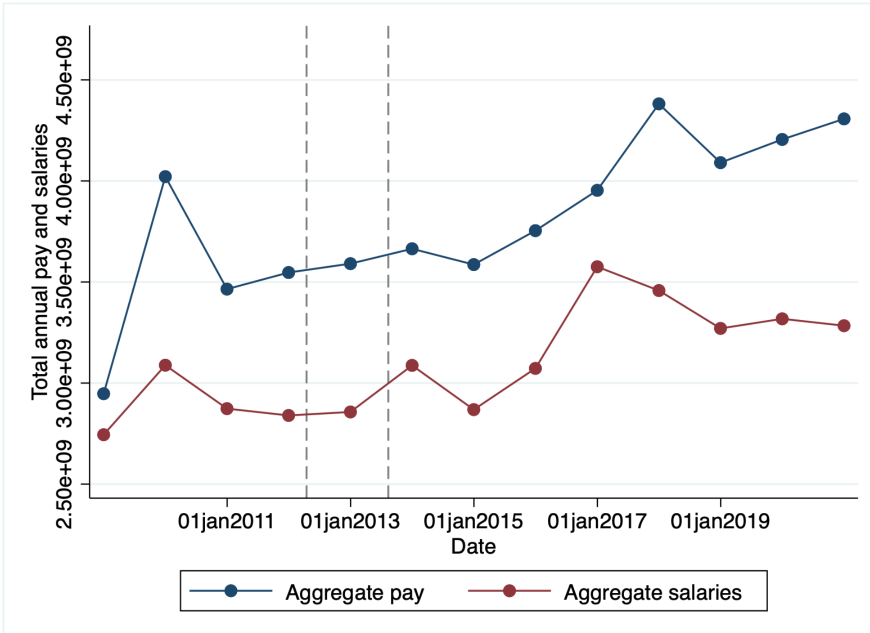
Figure B.2: *NYPD Officers and Civilian Employees by Year*



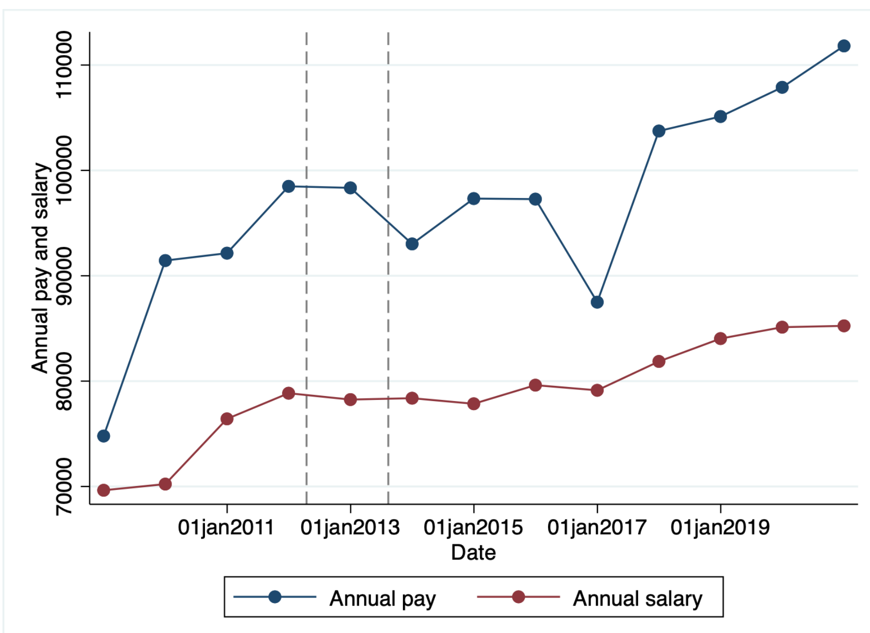
Notes: This table provides yearly counts of the number of officers employed by the NYPD. Data come from the FBI's Uniform Crime Reporting Program data. Specifically, we use data reported by the NYPD for the Law Enforcement Officers Killed and Assaulted (LEOKA) database for year 2006-2018. Data were collected from each agency as of October 31st of a given year.

Figure B.3: *NYPD Payroll*

Total Pay Expenditures



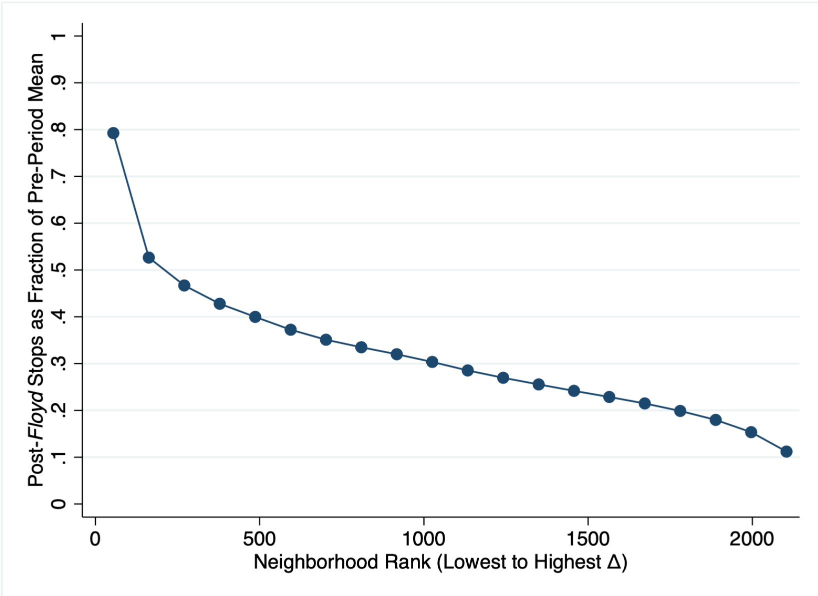
Pay per Officer



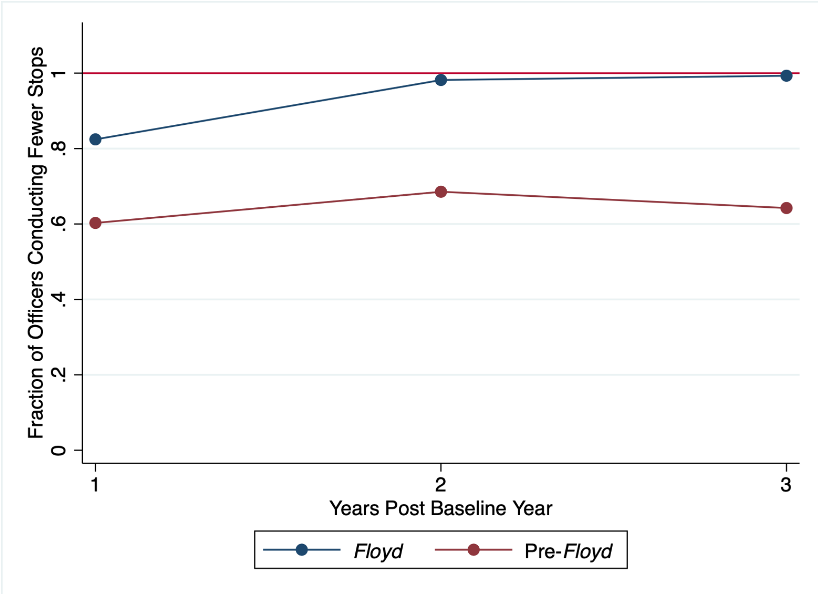
Notes: Data are from NYC payroll records that are publicly available on www.seethroughny.net/payroll. We include all officers affiliated with the NYPD whose titles include "Police Officer, Special Officer, Sergeant, Lieutenant, Captain". This excludes school safety officers, school guards, traffic enforcement agents, police cadets, police chief, police commissioner, deputy commissioners and deputy chiefs.

Figure B.4: *Floyd's Impact on Stop Rates Within-neighborhood & Within-officer*

Neighborhood

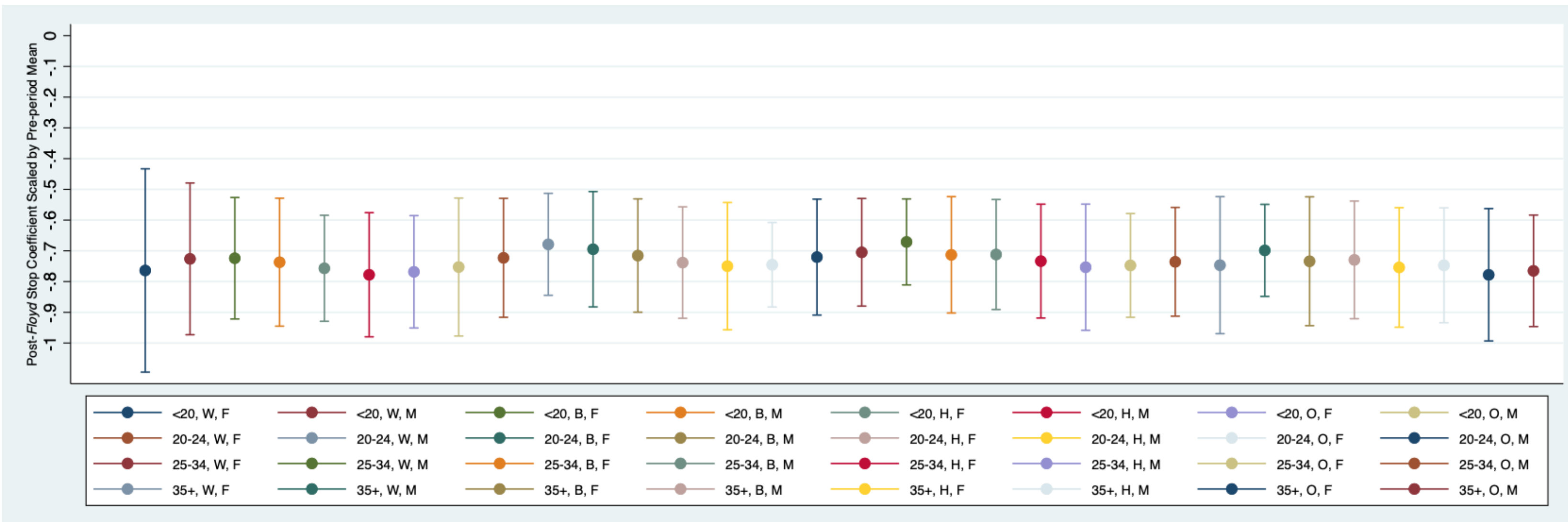


Officers



Notes: Panel A graphs the average change in stops in the three years post-*Floyd* scaled by mean stops in the two years prior to *Floyd* across all neighborhoods in our sample ranked from the lowest to highest changes. The blue line in Panel B depicts the fraction of officers that recorded at least 10 stops in the year prior to *Floyd* that recorded fewer stops in each year post *Floyd*. The red line is a placebo test that graphs the same fraction during the four years prior, such that the placebo reference year spans April 15th, 2007 through April 14th, 2008 and the placebo study period spans April 15th, 2008 through April 14th, 2010.

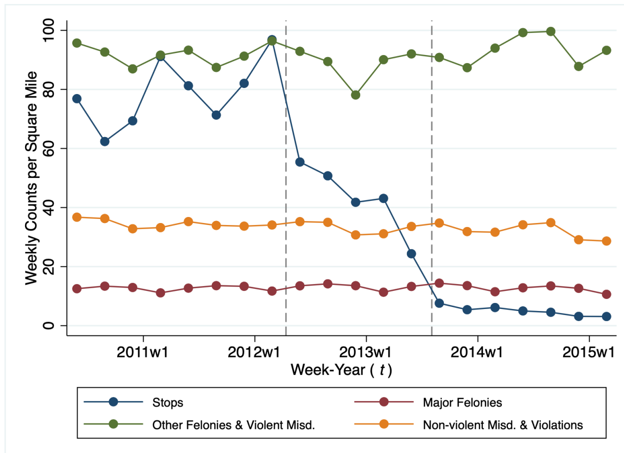
Figure B.5: *First-stage Effect on Stops for Each [Race] x [Age] x [Sex] Sub-group*



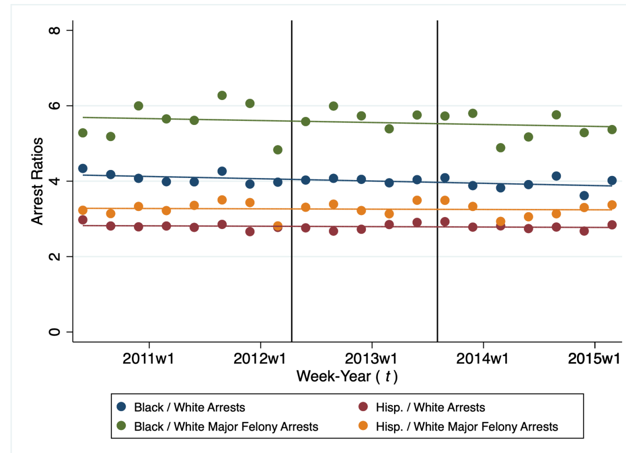
Notes: This figure graphs the coefficients on the post-*Floyd* indicator (Z) from regressing neighborhood stops on Z in the first-stage equivalent of Equation 2.5. We run this regression separately for each race-by-age-by-sex sub-group, where race is split into three mutually exclusive categories – White, Black, Hispanic – and age is split into four quartiles – < 20, 20-24, 25-34, 35+. Whiskers denote 95% confidence interval where standard errors are clustered at the precinct-level. Colors are given in the legend in order from left-to-right and top-to-bottom.

Figure B.6: *Exclusion Restriction Checks*

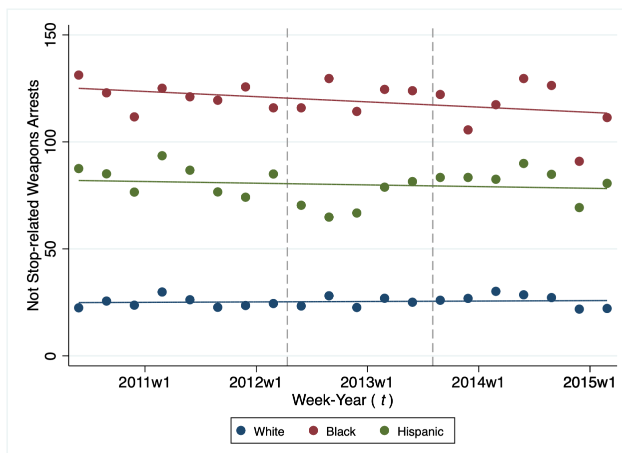
Trends of Stops and Reported Crimes



Black-White and Hispanic-White Arrest Ratios



Weapon-related Arrests Not Made During Stops



Notes: All data are aggregated to the city-week-level and each figure plots raw means in each quarter of the sample period, which spans April 15th, 2010 to April 14th, 2015. Panel A plots trends in weekly stops, reported major felonies, other felonies and violent misdemeanors, and non-violent misdemeanors and violations per square mile. Panel B plots the Black-White and Hispanic-White arrest ratios separately for all arrests and for all major felony arrests. Panel C plots all weapon-related arrests minus weapon-related stop arrests for each racial group.

B.3 Supplemental Tables

Table B.1: *Racial Disparities in Pre-Floyd Stop Outcomes*

Y	Arrest Made	Weapon Found	Gun Found	Drugs Found	Court Summons Issued
	(1)	(2)	(3)	(4)	(5)
Black	0.003***	-0.007***	0.001***	-0.004***	-0.015***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Hispanic	0.001*	-0.004***	0.000	-0.005***	-0.003***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)
Asian, Pacific Islander, or Native American	-0.004***	-0.007***	-0.000	-0.006***	0.006***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
Other or Missing Race	-0.010***	-0.008***	-0.000*	-0.009***	-0.015***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
White Mean	0.0608	0.0136	0.0007	0.0221	0.0615
Observations	3,423,793	3,423,793	3,423,793	3,423,793	3,423,793
R-squared	0.017	0.006	0.003	0.008	0.020

Notes: This table reports race coefficients from Equation 2.3. The reference group are white non-Hispanic pedestrians who are stopped by police. "Other or missing race" cover 3.52% of all stops and include stops where police reported race as "other" or did not provide racial information. Column (1) reports coefficients without any situational controls. All regressions include fixed effects for precinct, tract, day-of-week, male, age quintiles, and eight three-hour time-of-day indicators. Standard errors are clustered at the date-time-location level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

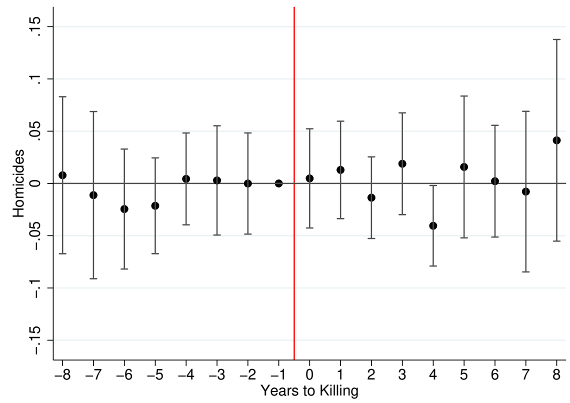
Appendix C

Appendix to Chapter 3

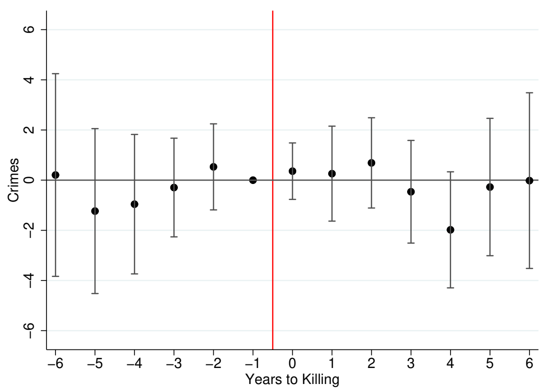
C.1 Supplemental Figures

Figure C.1: *Effects on Local Crime and Arrests*

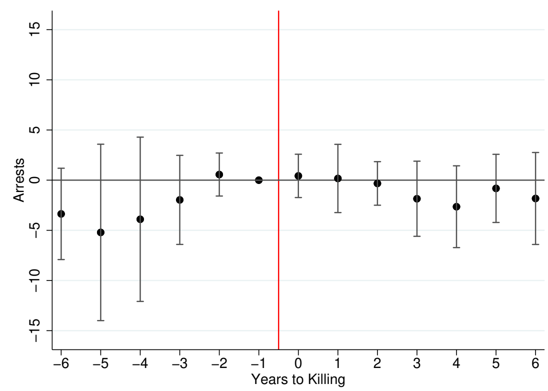
Panel A: Homicides



Panel B: Crimes

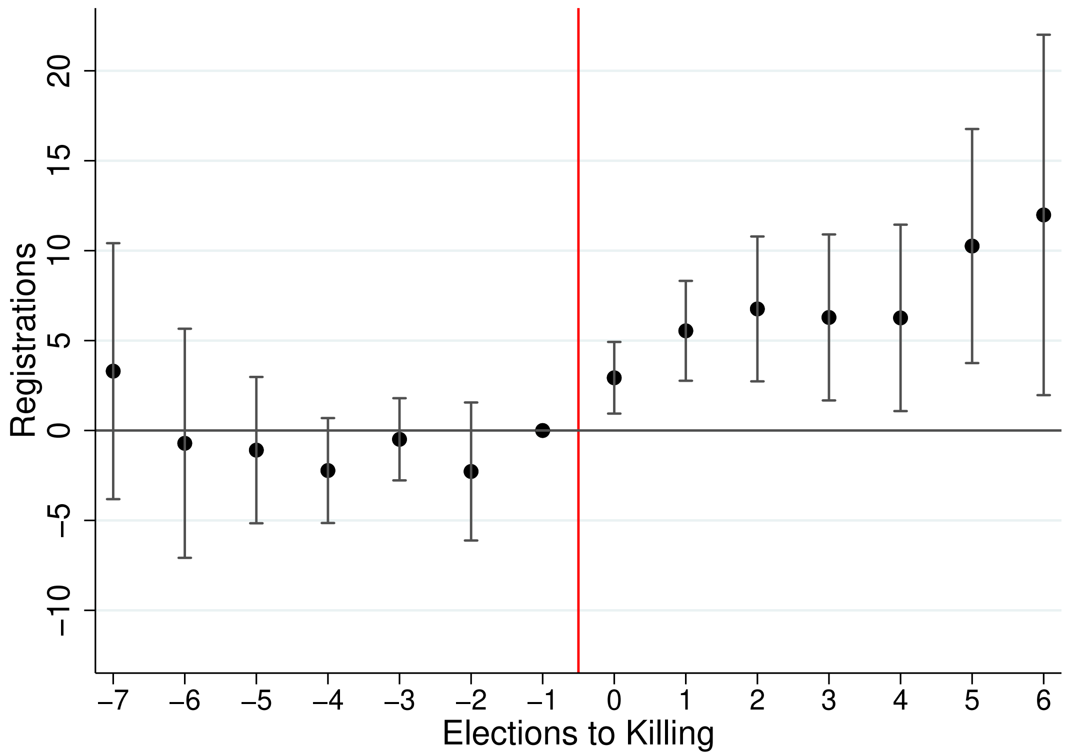


Panel C: Arrests



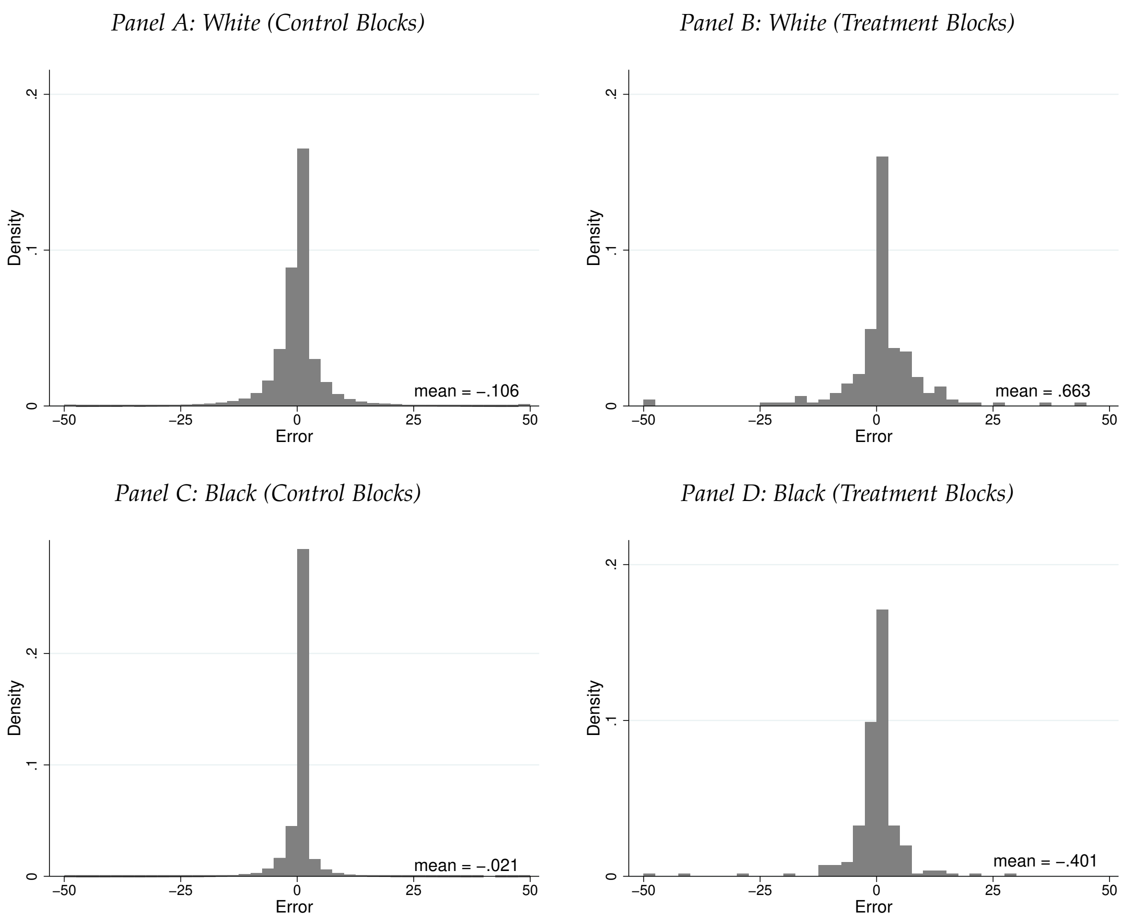
Notes: Figure shows treatment estimates and 95 percent CIs from estimation of Equation 3.2 on homicides, total crimes and arrests. Unit of observation is number of homicides, crimes and arrests in a Census block-year. Standard errors are clustered by Census block group. For Panel A, the sample spans the 2002 to 2010 elections and treatment is defined by blocks where police killings occurred during the sample period. As data on crimes and arrests is only available after 2010, for Panels B and C, the sample spans 2010 to 2016 and treatment is defined by blocks where police killings occurred from 2010 to 2016. Red vertical line represents time of treatment. Full regression results are included in Table C.2.

Figure C.2: *Effects on Voter Registration (2002-2016)*



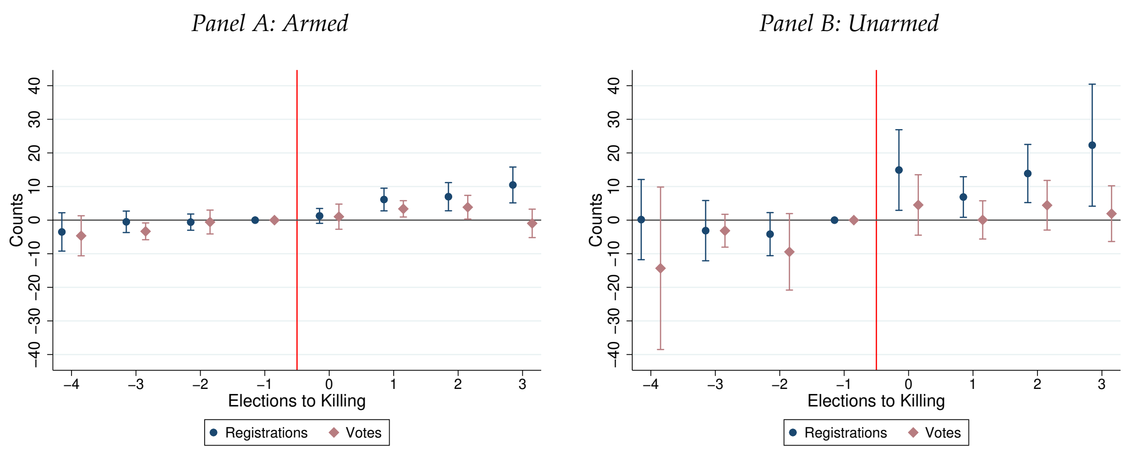
Notes: Figure shows treatment estimates and 95 percent confidence intervals from estimation of Equation 3.2 on registrations (pre-treatment mean = 85.7) on the extended sample. Unit of observation is registrations in a Census block-election. Standard errors are clustered by Census block group. The sample spans general elections and police killings from 2002 to 2016 and treatment is defined by blocks where police killings occurred during the sample period. Red vertical line represents time of treatment. Full regression results are included in Table C.4.

Figure C.3: *Validation of Predicted Race Counts*



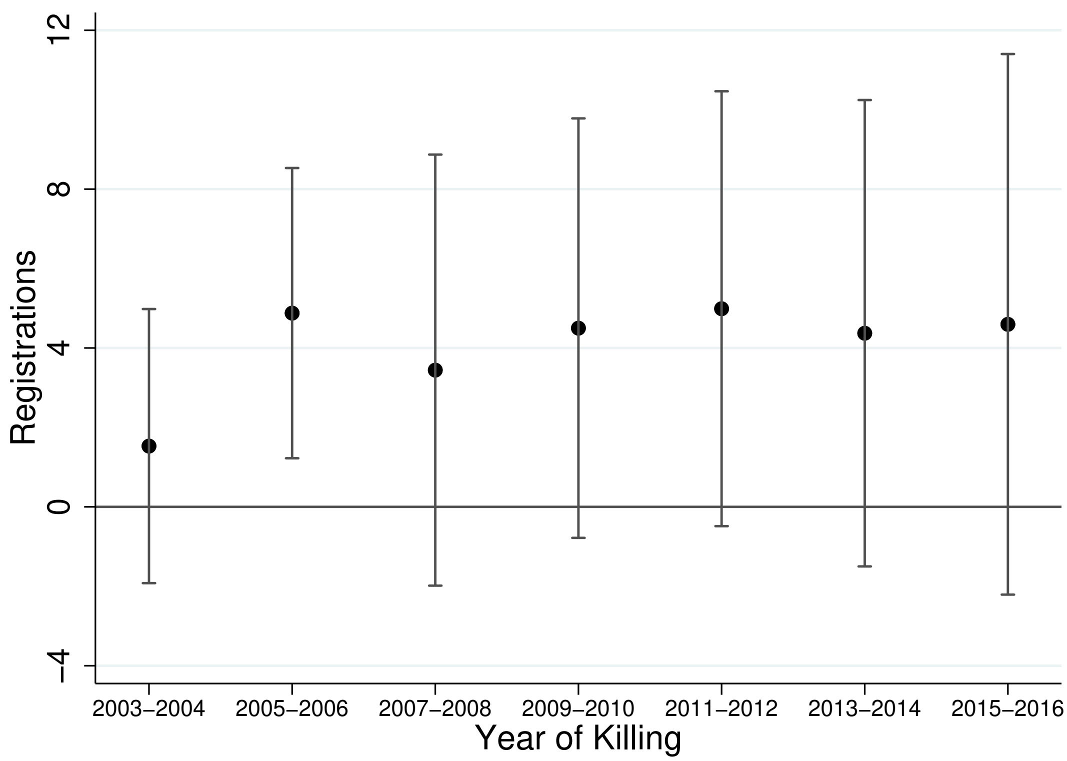
Notes: Figure shows histograms validating estimated vote counts by race obtained from Equation 3.5.2 against estimates predicted from individual-level voter registration file extracted on February 4, 2011. Each voter's race is predicted from the registration file using surname and address based on the Consumer Financial Protection Bureau's Bayesian Improved Surname Geocoding method (Consumer Financial Protection Bureau, 2014). A voter is classified as Black (white) if his/her predicted probability of being Black (white) exceeds that of any other race group. Vote counts are then aggregated to Census block. Histograms show the algebraic difference between 2010 estimates from Equation 3.5.2 and corresponding voter file estimates, separately for white/Black voters and Census blocks that did/not experience a police killing during the sample period.

Figure C.4: *Effects by Deceased Weapon without Population Controls*



Notes: Figure shows treatment estimates and 95 percent confidence intervals from estimation of Equation 3.2 on registrations (pre-treatment mean = 81.6) and votes (pre-treatment mean = 42.9) excluding population decile by election fixed effects. Unit of observation is registrations/votes in a Census block-election. Panel A restricts treatment group to killings of individuals armed with a knife or gun. Panel B restricts treatment group to killings of individuals who were unarmed. Estimates are similar when including incidents with unknown weapon type. Standard errors are clustered by Census block group. The sample spans the 2002 to 2010 general elections and treatment is defined by blocks where police killings occurred during the sample period. Red vertical line represents time of treatment. Full regression results are included in Table C.3, Columns 4 and 5.

Figure C.5: *Effects on Voter Registration by Year of Killing*



Notes: Figure shows effects on voter registration by year of police killing. Each point represents the average of β_τ for $0 \leq \tau \leq 1$ from separate estimations of Equation 3.2, including all control blocks but only those treatment blocks that experienced a killing in the corresponding time period. For example, the point at 2003-2004 represents the average effect of police killings that occurred between the 2002 and 2004 general elections on registrations in 2004 and 2006. The sample spans general elections and police killings from 2002 to 2016 and treatment is defined by blocks where police killings occurred. Standard errors are clustered by Census block group. Full regression results are included in Table C.6.

C.2 Supplemental Tables

Table C.1: *Effects by Distance from Police Killing (Full Results)*

	(1)		(2)		(3)		(4)	
Distance	*DV=Registrations*							
In Block	4.576***	(1.486)	3.052**	(1.480)	9.928**	(4.207)	3.546**	(1.513)
0.1 mi	0.016	(1.090)	0.331	(1.849)	3.261**	(1.485)	1.442	(0.887)
0.2 mi	0.424	(0.448)	-0.214	(0.706)	3.782***	(1.247)	0.539	(0.626)
0.3 mi	0.506	(0.335)	-0.166	(0.544)	2.701***	(0.977)	0.342	(0.493)
0.4 mi	0.207	(0.327)	-0.155	(0.420)	2.388***	(0.861)	0.537	(0.448)
0.5 mi	0.845***	(0.267)	0.392	(0.347)	1.185	(0.810)	0.745*	(0.395)
0.6 mi	0.491	(0.330)	0.238	(0.349)	-0.135	(0.623)	0.550	(0.355)
0.7 mi	-0.206	(0.256)	-0.315	(0.281)	-0.141	(0.523)	0.327	(0.295)
0.8 mi	-	-	-	-	-	-	-	-
0.9 mi	-0.340	(0.249)	-0.454*	(0.273)	0.028	(0.511)	0.481	(0.357)
1.0 mi	-0.233	(0.328)	-0.239	(0.391)	-0.195	(0.558)	0.387	(0.403)
1.1 mi	0.013	(0.280)	-0.159	(0.370)	0.065	(0.639)	0.415	(0.459)
1.2 mi	-0.049	(0.297)	-0.379	(0.412)	0.605	(0.650)	0.173	(0.488)
1.3 mi	-0.207	(0.269)	-0.467	(0.463)	0.872	(0.704)	0.224	(0.551)
1.4 mi	0.360	(0.315)	0.386	(0.530)	0.804	(0.770)	0.777	(0.673)
1.5 mi	0.459	(0.333)	0.433	(0.564)	0.726	(0.791)	0.307	(0.626)
1.6 mi	0.133	(0.341)	0.209	(0.584)	0.479	(0.823)	0.060	(0.664)
1.7 mi	0.158	(0.319)	0.575	(0.623)	-0.248	(0.850)	-0.206	(0.684)
1.8 mi	-0.056	(0.303)	0.230	(0.661)	-0.510	(0.870)	0.265	(0.754)
1.9 mi	0.233	(0.383)	0.310	(0.738)	0.186	(0.889)	-0.445	(0.812)
2.0 mi	-0.173	(0.350)	-0.371	(0.712)	0.370	(0.908)	-0.380	(0.806)
≥2.0 mi	0.619**	(0.267)	-0.566	(0.748)	-0.739	(0.960)	-0.629	(0.827)
Police Killings	All		All		w/ Media		w/o Media	
Fixed Effects	Election		Blk Grp x Election		Blk Grp x Election		Blk Grp x Election	
Obs.	341,630		341,420		341,420		341,420	

Table reports coefficients from estimation of Equation 3.1 on registration counts. Sample spans 2002 to 2010 general elections. Indicators are mutually-exclusive and track a Census block's minimum distance to police killings that occurred prior to a given election."In Block" refers to blocks where killings occurred. Other blocks are partitioned by 0.1-mile bins. Columns 1 and 2 examine all police killings in sample. Column 1 includes election fixed effects, Column 2 includes block group by election fixed effects. Columns 3 and 4 include block group by election fixed effects and examine only police killings that did and did not receive media coverage, respectively. Standard errors clustered at the Census block group-level and shown in parentheses. Estimates from Column 1 and 2 are displayed visually in Figure 3.1 and from Column 3 and 4 in Figure 3.4.

Table C.2: *Effects on Local Crime and Arrests (Full Results)*

	(1)		(2)		(3)	
Years to Treat	*DV=Homicides*		*DV=Crimes*		*DV=Arrests*	
-8	-0.040	(0.041)				
-7	-0.002	(0.037)				
-6	-0.026	(0.028)	0.205	(2.061)	-3.363	(2.322)
-5	-0.012	(0.023)	-1.233	(1.677)	-5.209	(4.484)
-4	0.018	(0.025)	-0.957	(1.418)	-3.899	(4.176)
-3	0.004	(0.026)	-0.294	(1.003)	-1.968	(2.264)
-2	0.009	(0.024)	0.530	(0.875)	0.561	(1.093)
-1	-	-	-	-	-	-
0	0.009	(0.024)	0.358	(0.574)	0.425	(1.101)
1	0.018	(0.024)	0.259	(0.966)	0.169	(1.735)
2	-0.009	(0.020)	0.689	(0.918)	-0.326	(1.108)
3	0.023	(0.025)	-0.463	(1.044)	-1.854	(1.910)
4	-0.036*	(0.019)	-1.980*	(1.180)	-2.645	(2.079)
5	0.020	(0.034)	-0.274	(1.396)	-0.821	(1.732)
6	0.006	(0.027)	-0.018	(1.786)	-1.826	(2.338)
7	-0.004	(0.039)				
8	0.046	(0.049)				
Mean	0.06		13.84		8.28	
Sample	2002-2010		2010-2016		2010-2016	
Obs.	614,556		326,879		326,879	

Table reports results from estimating Equation 3.2 on homicides, crimes and arrests. Unit of observation is number of homicides, crimes and arrests in a Census block-year. Standard errors are clustered by Census block group and shown in parentheses. For Column 1, the sample spans 2002 to 2010 and treatment is defined by blocks where police killings occurred during the sample period. As data on crimes and arrests from the Los Angeles Police Department are only available after 2010, Columns 2 and 3 span LAPD jurisdiction areas from 2010 to 2016 with treatment defined as blocks in those areas where police killings occurred from 2010 to 2016. Mean homicides, crimes and arrests for treatment blocks during the year prior to treatment are listed. Estimates are displayed visually in Figure C.1.

Table C.3: *Effects on Civic Engagement (Full Results)*

	(1)		(2)		(3)		(4)		(5)	
Elections to Treat					***Panel A: DV=Registrations***					
-4	-1.460	(2.424)	-2.305	(2.761)	2.939	(6.087)	-3.511	(2.910)	0.157	(6.088)
-3	0.221	(1.548)	0.680	(1.577)	-0.846	(4.452)	-0.503	(1.629)	-3.133	(4.574)
-2	-0.471	(1.146)	0.107	(1.165)	-2.230	(3.169)	-0.598	(1.231)	-4.172	(3.270)
-1	-	-	-	-	-	-	-	-	-	-
0	2.585*	(1.332)	0.606	(1.138)	13.793**	(6.112)	1.258	(1.127)	14.912**	(6.117)
1	4.369***	(1.281)	4.876***	(1.562)	4.945*	(2.654)	6.132***	(1.721)	6.876**	(3.080)
2	4.407***	(1.670)	4.860**	(1.932)	10.502**	(4.134)	6.974***	(2.139)	13.871***	(4.415)
3	3.640	(2.218)	6.725***	(2.438)	18.872**	(8.504)	10.467***	(2.720)	22.305**	(9.261)
Mean	81.59		80.67		101.53		80.67		101.53	
Elections to Treat					***Panel B: DV=Votes***					
-4	-1.294	(2.793)	-0.406	(2.452)	-4.367	(12.364)	-4.658	(3.039)	-14.327	(12.332)
-3	-1.813*	(1.014)	-2.022*	(1.145)	-0.935	(2.266)	-3.334***	(1.275)	-3.153	(2.492)
-2	-1.038	(1.371)	0.257	(1.315)	-6.118	(5.386)	-0.561	(1.811)	-9.436	(5.805)
-1	-	-	-	-	-	-	-	-	-	-
0	0.586	(1.398)	0.181	(1.728)	4.687	(3.821)	1.026	(1.909)	4.516	(4.586)
1	2.187**	(0.882)	3.010***	(1.060)	0.003	(2.578)	3.357***	(1.239)	0.066	(2.901)
2	1.133	(1.426)	1.953	(1.358)	2.945	(2.713)	3.842**	(1.800)	4.428	(3.767)
3	0.240	(1.399)	-0.523	(2.090)	2.372	(4.106)	-0.966	(2.159)	1.928	(4.228)
Mean	42.87		43.49		54.86		43.49		54.86	
Model	Main		Main		Main		w/o Pop Ctrls		w/o Pop Ctrls	
Police Killings	All		Armed		Unarmed		Armed		Unarmed	
Obs.	341,420		341,000		340,200		341,000		340,200	

Table reports results from estimating Equation 3.2 on registrations (Panel A) and votes (Panel B). Unit of observation is registrations/votes in a Census block-election. Standard errors are clustered by Census block group and shown in parentheses. Column 1 examines all police killings in sample. Columns 2 and 4 restrict treatment group to killings of individuals armed with a knife or gun. Column 3 and 5 restricts treatment group to killings of individuals who were unarmed. Columns 2 and 3 include population decile by election fixed effects, Columns 4 and 5 drop those controls. The sample spans the 2002 to 2010 general elections and treatment is defined by blocks where police killings of a given type occurred during the sample period. Mean registrations and votes for treatment blocks during the election prior to treatment are listed. Estimates from Column 1 are displayed visually in Figure 3.2, from Columns 2 and 3 in Figure 3.5, and from Columns 4 and 5 in Figure C.4.

Table C.4: *Effects on Voter Registration for 2002-2016 (Full Results)*

	(1)	
Elections to Treat	*DV=Registrations*	
-7	3.299	(3.630)
-6	-0.711	(3.249)
-5	-1.092	(2.075)
-4	-2.227	(1.488)
-3	-0.489	(1.165)
-2	-2.279	(1.958)
-1	-	-
0	2.931***	(1.016)
1	5.543***	(1.416)
2	6.760***	(2.054)
3	6.287***	(2.354)
4	6.262**	(2.644)
5	10.257***	(3.320)
6	11.986**	(5.111)
Mean	93.53	
Obs.	546,272	

Table reports results from estimating Equation 3.2 on registrations. Unit of observation is registrations in a Census block-election. Standard errors are clustered by Census block group and shown in parentheses. The sample spans general elections and police killings from 2002 to 2016 and treatment is defined by blocks where police killings occurred during the sample period. Mean registrations for treatment blocks during the election prior to treatment are listed. Estimates are displayed visually in Figure C.2.

Table C.5: *Heterogeneous Effects (Full Results)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Panel A: DV=Registrations															
	Voter Race			Voter Age			Years Registered				Party			Voter Race		
	Black	Hisp	Wht/Asn	18-34	35-54	55+	0-3	4-7	8-11	12-15	Dem	Ind	Rep	Blk/Hsp	Wht/Asn	All
Treat x Post	1.751***	1.418**	0.338	1.740***	1.043*	0.746	2.823***	0.456	0.354	0.223	2.499***	0.544	0.592			
	(0.589)	(0.717)	(0.609)	(0.633)	(0.568)	(0.501)	(0.843)	(0.444)	(0.317)	(0.222)	(0.787)	(0.353)	(0.446)			
x Blk/Hisp Killing														3.497***	-0.197	
														(1.164)	(0.643)	
x Wht/Asn Killing														1.362	2.251	
														(1.069)	(1.632)	
x Near Home																5.459**
																(2.724)
x Not Near Home																3.875**
																(1.659)
Mean	20.22	31.08	27.58	26.8	30.36	22.45	36.83	19.31	9.24	4.85	48.48	14.75	18.35	51.3	27.58	81.59
	Panel B: DV=Votes															
	Voter Race			Voter Age			Years Registered				Party			Voter Race		
	Black	Hisp	Wht/Asn	18-34	35-54	55+	0-3	4-7	8-11	12-15	Dem	Ind	Rep	Blk/Hsp	Wht/Asn	All
Treat x Post	1.021**	0.973**	-0.305	0.971***	0.594	0.080	1.460***	0.403	0.219	0.058	1.240**	0.235	0.267			
	(0.415)	(0.435)	(0.509)	(0.345)	(0.390)	(0.388)	(0.521)	(0.300)	(0.208)	(0.159)	(0.568)	(0.214)	(0.257)			
x Blk/Hisp Killing														2.108***	-0.600	
														(0.731)	(0.552)	
x Wht/Asn Killing														1.289	0.686	
														(0.817)	(1.191)	
x Near Home																3.668*
																(2.065)
x Not Near Home																1.689
																(1.043)
Mean	9.22	16.99	15.34	10.42	16.66	14.52	17.43	9.44	5.27	3.02	26.9	7.99	7.98	26.21	15.34	42.87
Obs.	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420	341,420

Table reports results from estimating Equation 3.2 separately for each voter group (i.e., by race, age, registration length and party affiliation), replacing the time to treatment indicators with a single post-treatment indicator. Columns 14 to 16 include distinct post-treatment indicators corresponding to each incident type (i.e., by deceased race and by proximity to deceased home). Standard errors are clustered by Census block group. Unit of observation is the Census block-election. The sample spans the 2002 to 2010 general elections and treatment is defined by blocks where police killings occurred during the sample period. Mean registrations and votes for treatment blocks during the election prior to treatment are listed. Estimates are displayed visually in Figure 3.3.

Table C.6: *Effects on Voter Registration by Year of Killing (Full Results)*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Elections to Treat				*DV=Registrations*			
-7							5.128
							(4.568)
-6						-4.626	4.260
						(3.868)	(6.078)
-5					-3.172	-0.726	2.900
					(3.816)	(2.518)	(3.919)
-4				-0.911	-2.335	-2.286	-0.014
				(2.707)	(2.642)	(2.579)	(2.704)
-3			-0.269	1.115	-3.327	2.879	-0.720
			(1.700)	(2.282)	(2.398)	(2.399)	(1.941)
-2		-0.127	-0.254	0.527	0.146	-8.440	-1.593
		(1.901)	(1.364)	(2.022)	(2.262)	(10.227)	(1.011)
-1	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
0 to 1	1.530	4.878***	3.443	4.500*	4.989*	4.372	4.596
	(1.760)	(1.864)	(2.768)	(2.694)	(2.793)	(2.995)	(3.472)
2 to 3	1.459	7.708***	2.235	8.694*	16.449**		
	(2.055)	(2.921)	(3.458)	(5.243)	(7.372)		
4 to 5	5.372	11.809***	1.361				
	(3.649)	(4.448)	(6.882)				
6	8.702						
	(5.423)						
Mean	74.40	78.65	97.21	77.36	101.19	123.57	99.73
Police Killings	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014	2015-2016
Obs.	542,736	542,672	542,640	542,640	542,736	542,728	542,728

Table reports results from estimating Equation 3.2 on registrations. Unit of observation is registrations in a Census block-election. Standard errors are clustered by Census block group and shown in parentheses. The sample spans general elections from 2002 to 2016. In each column, treatment group is restricted to blocks where police killings occurred during the respective period. For example, Column 1 examines police killings that occurred between the 2002 and 2004 general elections. Mean registrations for treatment blocks during the election prior to treatment are listed. Estimates for β_τ for $0 \leq \tau \leq 1$ (i.e. Elections to Treat for "0 to 1") are displayed visually in Figure C.5.