Title

Essays in the Economics of Crime and Health Economics

Author

Premkumar, Deepak

Publication Date

2020

Peer reviewed|Thesis/dissertation

Essays in the Economics of Crime and Health Economics

by

Deepak Premkumar

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Michael Anderson, Chair
Associate Professor Jeremy Magruder
Professor Steven Raphael

Summer 2020

Essays in the Economics of Crime and Health Economics

Abstract

Essays in the Economics of Crime and Health Economics

by

Deepak Premkumar

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Associate Professor Michael Anderson, Chair

This dissertation explores policy-relevant issues in the economics of crime and health economics. In the first study, I explore the impact of public scrutiny generated by high-profile, officer-involved fatalities on police officer effort. The second chapter provides a synthesis of the evidence on whether police reduce crime, a timely review during a national discussion on policing—prompted by the "Defund the Police" movement. Finally, the last chapter explores the impacts of rural hospital closings on rural patient welfare, an increasing likely phenomenon as the COVID-19 virus has rendered numerous healthcare facilities insolvent.

The first chapter explores whether public scrutiny of police officers reduces their level of effort—empirically testing the *Ferguson Effect*. Leveraging the quasi-random timing of high-profile, officer-involved fatalities (OFs), this paper provides the first national analysis of this question with department-level data on arrests and crime. To address the simultaneous effects that could occur after an OF—(1) greater scrutiny of police, (2) reduced community cooperation with identifying and locating suspects, (3) reduced civilian crime reporting, and (4) changes in offending behavior—I develop a theoretical model of policing behavior to provide empirical predictions of the changes in arrests due to each of the four possible channels. Following a high-profile OF, theft arrests drop by 4–13%, while arrests for the least serious offenses (e.g., marijuana possession and disorderly conduct) see sharp declines of up to 33%. Notably, arrest rates do not change for violent crime or more serious property crimes. These findings are consistent with scrutiny as the causal channel for the reduction in arrests. While the decline in arrests for theft is temporary, it persists for the least serious offenses, representing a sustained transition to a lower equilibrium effort. Reductions in arrests occur for both black and white suspects, but reductions for black suspects are suggestively larger in cases of theft and marijuana possession.

The second chapter, co-authored with Justin McCrary, discusses the role that the police have in deterring and reducing crimes. After a brief overview of deterrence theory, we discuss the empirical evidence on the efficacy of police staffing and various policing strategies on crime reduction. Using a framework developed in Weisburd and Eck (2004), we quickly evaluate the model of standard policing and then mainly focus on evidence behind three current policing practices: hot spots, problem-oriented, and proactive. Finally, we use the empirical evidence of police staffing to provide a basis for a theoretical model on the optimal

level of policing. Using the Chalfin and McCrary (2018) framework, we discuss how one could estimate how much crime could be reduced if additional funds were directed to hire more law enforcement officers, and if crime-reduction were the sole policing objective, how many cities are in fact underpoliced. We conclude by postulating whether we could implement additional policing without resulting in unwarranted and excessive social costs for the community as discussed by Manski and Nagin (2017).

The third chapter, co-authored with Dave Jones and Peter Orazem, estimates a model of rural patient hospital choice between the nearest rural hospital, the nearest urban hospital, or the nearest research hospital. We present separate estimates for inpatient and outpatient visits, for different diagnoses, and for emergency and nonemergency admissions. The analyses illustrate the tradeoffs between hospital quality and distance in deciding whether to choose the nearest hospital or to travel farther for an alternative. The model parameters are used to simulate two hospital closing scenarios for both outpatient and inpatient data: 1) closing 25% of lowest quality rural hospitals and 2) closing 15% of the least used rural hospitals. Closing 25% of the lowest quality rural hospitals results in a 20.7% increase in expected distance and a 7.7% increase in expected hospital quality for those with inpatient ailments. Closing the least used hospitals modestly increases average distance but lowers average quality. We conclude that closing the lowest quality rural hospitals is a better policy prescription than closing the least used hospitals since closing low quality hospitals results in a substantial increase in average quality of hospital with only a slight increase in distance traveled for chosen hospitals.

Contents

List of Figures

List of Tables

Acknowledgments

I am acutely aware that I did not get here by myself. This section contains a non-exhaustive list of the multitudes of people that helped me get to where I am today. Before I take time to reflect on that village, it is important to recognize that many people do not have the same support system that I did. Perhaps, in some small way, these little doses of introspection and gratitude will eventually help to dismantle systems of privilege, leading toward a more equitable society.

I am grateful for the incisive comments and advice I have received from my advisor, Michael Anderson. The first chapter of this dissertation has been years in the making, almost since the very beginning of graduate school. I appreciate his patience in providing feedback on the same topic, as it slowly evolved to something more refined and structured. Michael's detailed comments on my writing, especially during the job market, were instructive on how to take a smattering of tables and figures and make it into a research product. I hope the end product is something he can take pride in, especially after comparing it to where it started.

I am also thankful for the support and guidance of Jeremy Magruder and Steve Raphael. Jeremy provided helpful feedback, particularly in the early stages of the project, on how to approach theoretical modeling. Steve's research interests greatly overlap with mine, so he was able to work through my ideas using both economic rigor and institutional detail. Every time I finished a meeting with Steve, I left feeling motivated and inspired to keep working on issues in policing.

I would be remiss if I did not mention the help that Rebecca Goldstein and Justin McCrary provided. They were part of a team, along with Michael and Steve, that provided recommendation letters and guidance on the job market. From their comments on my job market paper to collaborating with them on respective research projects, my work with them has been a highlight of my graduate student career. I truly cannot imagine the last year without the personal, intellectual, and monetary support that Rebecca provided. Her consistent meetings, filled with both life and work updates, embodied the culture I hope to create as a future research mentor. It is also important to mention the incredible support that Yotam Shem-Tov provided. He had no incentive to do so, but out of the kindness of his heart, Yotam repeatedly met with me during the job market, providing encouragement when I was at my most stressed and sharing with me invaluable practical advice on how to succeed.

My favorite part of graduate school were the fellow students and close friends I met along the way. I will always have fond memories of the 2014 Cohort in Agricultural and Resource Economics (ARE). I hope we all remain in touch and support each other going forward.

My experience in ARE would not have been the same without Diana Lazo and Carmen Karahalios. They are a strong foundation that provide immense warmth in the department, particularly during parts of graduate school when the road ahead seemed daunting.

Before UC Berkeley, there was Iowa State University (ISU). There are so many people at ISU that provided me with immeasurable opportunities. I had a tremendous experience there, and I absolutely would not have been able to pursue my PhD at Berkeley without that support. In particular, I would like to thank Peter Orazem, since I may not have been an economist if it was not for his openness to work with a freshman with a passion for policy, but no experience with economics or statistics. It is strange to look back on a 10-year journey and realize it was the result of a research-matching program gone awry, but all's well that ends well.

Finally, and most importantly, I'd like to thank my mom (Shoba), my brother (Mukund), and my partner (Sarah). They, along with many friends (obligatory Northside shout out), have cultivated a support system that weathered many ups and downs. They may not have always understood what I was talking about, but they were always willing to listen. My mom comforted me during a rocky start to graduate school, when I wondered whether I was in the right place. My brother overwhelmingly praised me for the smallest things, never taking any credit for his foundational role in my success. Sarah managed to sneak in later on, but she got a front-row seat to the whole job market roller coaster, including several 11 p.m. practice presentations. Somehow, with all of their encouragement, I made it to the other side, and it seems my schooling journey is coming to an end (for now).

I'd like to dedicate this dissertation, and my future research career, to my dad (Prem). He sadly passed away far too soon when he was far too young. But, from what I've gathered and been told, he was a brilliant researcher and professor. I desperately would have loved to chat with him about graduate school, the job market, and frustrations with publishing. I know finishing my PhD would have meant a great deal to him, and I would like to dedicate my degree to his memory.

Chapter 1

Intensified Scrutiny and Bureaucratic Effort: Evidence from Policing After High-Profile, Officer-Involved Fatalities[1]

1.1 Introduction

How can law enforcement agencies provide safety without compromising trust or imposing undue burden on communities? Recent calls to "defund the police" have popularized one answer to this question—that they cannot. Theoretically, though, police departments allocate services to minimize the cost of crime in their jurisdiction, subject to a constraint imposed by the police budget and operational costs. In the United States, the aggregate annual expenditure on policing services is about $115 billion, while the cost of crime victimization is estimated to be over $310 billion each year (Chalfin, 2015; Auxier, 2020). Several studies (Evans and Owens, 2007; Chalfin and McCrary, 2018; Mello, 2019) find that US cities are underpoliced, in the sense that the additional hiring of police officers would result in welfare gains from reduced crime costs. However, many of these studies do not account for the social costs of policing, broader costs imposed upon on a community from an increase in police presence or the confrontational nature of a policing strategy (McCrary and Premkumar, 2019).

Police departments may not internalize these social costs either: First, these social costs may not be broadly understood across the community if they are concentrated on a subset

[1]I would like to thank the Oakland Police Department for providing institutional knowledge, their time for informal interviews, and a ride-along. I really appreciate Michael Anderson, Cyndi Berck, Rebecca Goldstein, Jeff Perloff, Steve Raphael, and Yotam Shem-Tov for their incisive comments and careful read-throughs. I thank Abhay Aneja, Bocar Ba, Aaron Chalfin, Julien Lafortune, Justin McCrary, Steve Mello, and Matt Pecenco for their comments, as well as the seminar participants at UC Berkeley's Agricultural and Resource Economics department, Cornerstone Research, US Treasury, Government Accountability Office, Federal Trade Commission, Yale's Public Health Modeling Unit, and the Public Policy Institute of California. Jacob Kaplan's formatted datasets on the FBI's Uniform Crime Report were incredibly helpful for this study. This project would not have been possible without the data collection efforts of Fatal Encounters. Finally, this research has been financially supported by the Horowitz Foundation and the Berkeley Empirical Legal Studies Fellowship.

of the population, particularly in the presence of a common set of laws that inhibit public access to complaint, disciplinary, and use of force records for police officers (Bies, 2017). Even if they were understood, municipalities do not have market-based mechanisms to elicit their optimal allocation, since policing is a public good. Furthermore, the bureaucratic structure of police departments limits electoral tools to signal and enact local preferences (Friedman and Ponomarenko, 2015). Thus, communities may have to rely on alternative means—such as increased scrutiny through protests, media coverage, or additional monitoring (e.g., cop-watch groups)—to raise awareness and engender their preferred policing allocation (Battaglini, Morton, and Patacchini, 2020; Ouss and Rappaport, 2020). In the wake of highly-publicized, officer-involved fatalities, it has been suggested that police effort has been reduced due to intensified scrutiny from the community and media, widely known as the *Ferguson Effect* (Mac Donald, 2015).

This paper provides the first empirical examination of a key question of policy interest: how does a high-profile, officer-involved fatality in your jurisdiction affect local policing behavior and crime? Using a national analysis that exploits the plausibly exogenous timing of these fatalities to credibly identify causal effects, I explore how police officer effort is affected by public scrutiny, measured through community awareness, media coverage, and local protests. To determine what constitutes "high profile," I create a novel dataset by merging crowdsourced information on officer-involved fatalities (hereafter, OFs) with web-scraped data on local search engine patterns and the amount of media coverage on each incident. I combine this with the FBI Uniform Crime Report, which contains data from the near universe of police departments. I use monthly arrest counts, disaggregated by race and crime type, to analyze policing effort in 2,740 police departments with 53 municipalities experiencing at least one high-profile, officer-involved fatality between 2005–2016. I assess whether there are changes in arresting patterns in these treated departments after a fatality relative to control departments—who experience no OFs or less publicized ones—controlling for changes in reported crime, population, national variation in arrests, time invariant department-level characteristics, and county-specific linear trends.

To include the variation when departments have multiple high-profile OFs, I utilize an event study regression design which allows for numerous "events" per observational unit. I subset the "events" to a sample of fatalities which have over 1,000 articles of news coverage. I restrict fatalities to this "high-profile" subset, because they must be large enough shocks to plausibly affect department-level metrics of officer behavior—which is unlikely for the far majority of OFs, given the common nature of occurrence (Krishnan, McCrary, and Premkumar, 2017). I also adjust the timing of these incidents to reflect community awareness by assessing when the community first starts searching for the incident using Google Trends. Further, I demonstrate that the municipalities with high-profile OFs experience larger and more frequent Black Lives Matter protests than control jurisdictions.

I measure changes in policing effort through the standard metric of arrests, also known as clearances (Mas, 2006; Shi, 2009; Heaton, 2010).[2] After an OF, there are numerous potential channels that could affect arrest rates: (1) greater scrutiny of police, (2) reduced community

[2]In this paper, clearing an offense, or clearances, is synonymous with an arrest, and they are used interchangeably for the remainder of the paper.

cooperation in identifying and locating suspects, (3) reduced civilian reporting of crimes, and (4) changes in criminal behavior, both in response to the OF and to subsequent signals of policing effort. I develop a novel theoretical model of a police officer's objective function, and provide a set of empirical predictions by offense type to identify the relevant mechanisms given any changes in arrests. Decoupling which mechanisms are occurring is integral to test the existence of the Ferguson Effect—whether high-profile OFs cause reductions in policing effort from increased scrutiny. To account for fluctuations in offending behavior and crime reporting, the main empirical specification controls for reported crime. As increased scrutiny drives up the marginal cost of officer effort, the model predicts that declines in effort should occur most prominently in arrest types that have low marginal benefit (MB) to the officer, primarily those corresponding to crimes with low social costs, such as drug possession.

Following a high-profile OF, officers curtail arrests for less serious offenses that generate lower marginal benefits. I find that arrests are reduced by about 3.2% for all UCR Part I offenses (murder, robbery, assault, burglary, theft, motor vehicle theft). This result is driven by theft arrests which see the largest decline, nearly 10%. Reductions in arrests are more marked among crime types that have low social costs. Using Part II offenses, a broader set that contain so-called quality of life crimes, I create an index of low MB arrests (e.g., disorderly conduct, liquor law violation, and marijuana possession). These low MB arrests exhibit a sharper decrease of up to 24%, exemplified most prominently by marijuana possession which sees declines of up to 33%. These arrests are the easiest to reduce, as they usually are officer-initiated and officers have the most discretion in determining whether to intervene. Notably, there are no changes in higher return arrests, such as violent crime or more serious property crime clearances. These findings are consistent with scrutiny being the causal channel for the reduction in arrests. An empirical analysis of the case study of Laquan McDonald's death, a high-profile OF which has separation in the timing of the incident and community awareness, bolsters that story. While the decline in effort for theft is temporary, it persists for the least serious offenses, representing a sustained transition to a lower equilibrium effort. Reductions in arrests occur for both black and white suspects, but reductions for black suspects are suggestively larger in theft and marijuana possession.

Building on past studies, this paper most closely relates to a literature that analyzes the determinants of police officer behavior.[3,4] Mas (2006) illustrates how arrest rates and average sentence length declines, and crime reports rise after police unions in New Jersey lose in arbitration for wage negotiations. Chandrasekher (2016) uses evidence from a union work slowdown in New York to find that officers drastically reduced the number of ticket citations but maintained arrests levels for all serious crimes. Heaton (2010) examines the impact of a policy change following a racial profiling incident in New Jersey, which led to

[3] Officer behavior changing as a result of a policy change or high-profile event: Prendergast (2001); Levitt (2002); Evans and Owens (2007); Klick and Tabarrok (2005); McCrary (2007); DeAngelo and Hansen (2014); Morgan and Pally (2016); Krishnan, McCrary, and Premkumar (2017); Shjarback et al. (2017); Cheng and Long (2018); Long (2019); Mello (2019); Rivera and Ba (2019); Rosenfeld and Wallman (2019); Devi and Fryer (2020)

[4] Discrimination/disparities in policing: Donahue and Levitt (2001); Knowles, Persico, and Todd (2001); Persico (2002); Anwar and Fang (2006); Legewie (2016); Legewie and Fagan (2016); Goncalves and Mello (2017); Manski and Nagin (2017); Fryer (2019); Rozema and Schanzenbach (2019)

sizable reductions in motor vehicle theft arrests for minorities, with partially corresponding increases in auto thefts afterward. Finally, Shi (2009) follows the effects of civil unrest in Cincinnati after a white officer killed a young, unarmed black man, and finds that there is a sharp decline in arrests, particularly for less serious crimes and violations. Previous studies have provided evidence from case studies or singular events, but this project provides the first national analysis of police effort, estimating the effects from 73 officer-involved fatalities and 2,740 police departments. It also answers a fundamental question of how communities instigate practical change in bureaucratic behavior when they do not have a direct mechanism to signal or translate their preferences. The particular context of policing is especially timely and has sizable welfare consequences given the significant cost of crime (Chalfin and McCrary, 2018).

The remainder of the paper is organized as follows: section 1.2 provides the background on the Ferguson Effect. section 1.3 presents a stylized model of the police officer's objective function and predictions for empirical analysis. section 1.4 discusses the numerous datasets in the analysis, how they are cleaned, and why each one is chosen. section 1.5 details the empirical specification used to estimate the results, with an explanation of its added value for settings with multiple treatments. section 1.5 also highlights the results, showing how policing behavior changes after a high-profile OF. Finally, section 1.6 discusses the results in the broader context of the model, and section 1.7 concludes.

1.2 Background on the Ferguson Effect

Over the past six years, there has been a renewed focus on policing issues following intense collective action from numerous organizations, highlighting the discrepancy between community preferences and the allocation of policing services. From thousands of Black Lives Matter protests across the nation (Elephrame, 2018) to a recently passed California bill that limits officer use-of-force (Gardiner, 2019), much of that attention is often traced back to a single event: an infamous officer-involved fatality in Ferguson, Missouri.

In August 2014, a white Ferguson police officer fatally shot an unarmed black teen, Michael Brown. Tension from elements of the event compounded upon racial anxieties and community mistrust, instigating protests in Ferguson, as well as national coverage of the officer-involved fatality (Swaine, 2014). The subsequent coverage brought attention to racial inequalities in policing across the country. In response to the rise in scrutiny of certain law enforcement practices, then FBI Director James Comey commented that the increase in violent crime in 2015 was a result of decreased officer activity, largely due to heightened attention and criticism from their local communities: *"I don't know whether that explains it entirely, but I do have a strong sense that some part of the explanation is a chill wind that has blown through American law enforcement over the last year"* (Schmidt and Apuzzo, 2015).

What he was referring to as the *chill wind* became known as "the Ferguson Effect" to law enforcement officers and crime-focused social scientists. These speculations on the Ferguson Effect have found evidence in anecdotes from law enforcement officers, such as a Chicago PD officer who was being assaulted on the job, but did not draw her firearm because of fear of media backlash (Hawkins, 2016). A more general consequence of the Ferguson Effect could be a reduction in the amount of discretionary police activity, where previously, officers may have exited their patrol cars to illustrate their presence in communities, question people they deem suspicious, and arrest those whom they deem are causing crime or disorder.

Though the effect is named for an incident in 2014, Ferguson was far from the first time where an officer-involved fatality generated community consternation toward the local police department. Before Michael Brown died, Eric Garner died on video from a police officer performing an illegal chokehold maneuver in New York in 2014. In an Oakland metro station, Oscar Grant also died on video in 2009 after an officer shot Grant in the back as he laid face down. These events carry a common thread of intense protests and continuous news coverage. The impetus for the unrest and attention is likely related to the perceived unjust nature of the fatality. It also relates to the discrepancy between community preferences for policing services and current practices. However, whether the change in scrutiny has shaped policing behavior, and to what extent, is still an empirical question. The phenomenon has been studied before with various empirical strategies (Morgan and Pally, 2016; Shjarback et al., 2017; Rosenfeld and Wallman, 2019), but the broader implications are harder to take away, because of limitations from study design. Cheng and Long (2018), and Rivera and Ba (2019) most closely resemble the analysis in this study, providing both a national and a city-specific analysis of Ferguson Effect (St. Louis and Chicago, respectively).

Cheng and Long (2018) estimate the spillover effects of the death of Michael Brown on arrests and crime in large US cities. Cheng and Long's national difference-in-difference analysis defines treated jurisdictions as eight predominantly black cities (out of 47 large cities in the sample) after the 2nd quarter of 2014. They find a reduction in arrests, driven by black arrests. Then, using incident-level data in St. Louis, they show reductions in the arrest and use of force rate after the death of Michael Brown in predominantly black census tracts compared to white ones in nearby St. Louis. Similar to the results I present in this paper, Cheng and Long find larger decreases for misdemeanors than felonies, as well as an increase in homicides. This paper complements work from Cheng and Long (2018) by directly estimating how local policing behavior and crime change after a high-profile, officer-involved fatality occurs in that jurisdiction, using a broader set of police departments and OFs in the empirical analysis.

Rivera and Ba (2019) test the effects of two police union warnings about complaints, and public scrutiny from the death of Laquan McDonald in Chicago. Since the public is not aware of police union notices, they have a unique setting where they can estimate changes in police behavior without simultaneity of public scrutiny. After the two police union warnings, they find that officers receive less civilian complaints without changes in crime and arrests, implying that the police can engage in effective self-monitoring without reducing officer effort. Conversely, after McDonald's death, crime and civilian complaints increase without commensurate increases in arrests. Like Rivera and Ba (2019), this study recognizes the potential limitations of the previous literature's use of large scandals as exogenous shocks, since they may simultaneously affect the community and police officers, influencing key variables of interest. This paper mitigates that concern in two ways: (1) I provide one of the first tractable models of officer behavior to understand how those separate channels would cause differing empirical patterns by offense type to ascertain what the salient mechanism is.

(2) In the main results of the paper, I control for reported crime, closing the intermediary channels of offending behavior or crime reporting.[5] Finally, an empirical analysis of the death of Laquan McDonald makes clear that police behavior is largely unchanged until public scrutiny occurs (Figure 1.9).

If the Ferguson Effect exists, it should locally manifest in jurisdictions which experience high-profile OFs, necessitating the inclusion of more comprehensive treatment data from across the nation and an empirical specification that internalizes the time-varying aspect of the scrutiny. This paper combines data from 2,740 police departments, 53 of which are treated, between 2005–2016 to study the impact of high-profile OFs on police behavior.

1.3 Theory and Stylized Model

After a high-profile OF, it is plausible that there are (1) increases in scrutiny of police, (2) reductions in community cooperation in identifying and locating suspects (Sunshine and Tyler, 2003), (3) reductions in civilian reporting of crimes (Desmond, Papachristos, and Kirk, 2016)[6], and (4) changes in criminal behavior—both in response to the OF and to signals of policing effort (Becker, 1968; Cloninger, 1991; Persico, 2002; Kirk and Papachristos, 2011; Rosenfeld and Wallman, 2019). These channels may all affect the number of arrests for the involved police department. To assess whether increased scrutiny from high-profile, officer-involved fatalities lead to reductions in policing effort, I develop a theoretical model of the officer's objective function to provide a set of empirical predictions to identify the mechanisms occurring given any changes in arresting patterns.

The officer's objective is to maximize utility by determining the amount of effort (e) to exert. The level of effort differs by type of crime. For simplification, let there be two types of crime, $c = \{L, H\}$, which represent offenses with low and high social cost of crime victimization respectively. Officers earn a baseline salary (w) and have a reward function (f_c) that monotonically increases in the number of arrests (A_c). There are three general parameters that enter into the utility function parameters: a cost of effort multiplier (ϕ),

[5] Concerns about the endogeneity of this control are assuaged after finding similar effects without it.

[6] A recent replication of Desmond, Papachristos, and Kirk (2016) illustrated that the main findings were not statistically significant after excluding an outlier (Zoorob, 2020). The original authors re-ran the analysis without the outlier—now controlling for temperature—and arguably find that there is still a decline in 911 calls (Desmond, Papachristos, and Kirk, 2020). Regardless, reductions in civilian reporting after an officer-involved fatality is possible and should be considered.

community cooperation (F), and reported crime (RC_c).[7] The objective of the officer is to maximize utility:

$$\begin{aligned}\max_{e_L, e_H} U(e_L, e_H; \phi, F, RC_L, RC_H) &= w + f_L(A_L) - \frac{1}{2}\phi e_L^2 + f_H(A_H) - \frac{1}{2}\phi e_H^2 \\ &= w + f_L(A_L(e_L; F; RC_L)) - \frac{1}{2}\phi e_L^2 + \\ &\quad f_H(A_H(e_H; F, RC_H)) - \frac{1}{2}\phi e_H^2\end{aligned}$$

To reward effort in policing, many police departments determine promotions based on arrests, either "a high volume of them or a string of high-quality ones" (Samaha, 2017). More arrests (A) increase the likelihood of promotion and ancillary payments, since officers are compensated by the number of hours worked and overtime pay for court appearances related to their arrests. The officer's marginal benefit of arrest is generally proportional to the social cost of crime ($\frac{df_H}{dA_H} > \frac{df_L}{dA_L}$). Naturally, when mapping to the real world, there is more heterogeneity in the social cost of crime, with Table 1.1 detailing the specific costs by offense.[8] Violent crimes, such as murder and aggravated assault, are tremendously socially costly and constitute high marginal benefit arrests (H). Property crimes, however, vary by offense type, where motor vehicle theft is quite costly and may be a high marginal benefit arrest, but theft—which has an order of a magnitude less cost—may be closer to a low marginal benefit arrest (L). The least serious offenses—such as marijuana possession or disorderly conduct—have no estimate because they do not have clear and distinct victims, making the cost relative to other crimes negligible. Thus, these offenses are low marginal benefit arrests.

Arrests are a concave function of effort ($\frac{\partial^2 A_c}{\partial e_c^2} < 0$), while the number of reported crimes is a combination of on-view reporting—crimes spotted by officers when patrolling—and crimes

[7] I consider individual officers as small, homogeneous agents. This reference point is consistent with a principal-agent framework, where the community (principal) does not have full information on the amount of effort the officers (agents) use. It also aligns with my informal interviews with and direct observations of police officers, which highlighted the large amount of discretion officers employ when enforcing the law, and the occasional disconnect between the stated policies of the police department and what is being executed on the ground. If the focus was on a change in policy, rather than officer behavior in response to scrutiny, I could reconstitute the model at the department level with the main comparative static results in tact, assuming that departments take (past) crime as given and the objection function is crime minimization.

In the model, city-level crime is not directly responsive to the individual effort of a single officer, but rather the aggregate effort at the department level. Although individual officers are not jointly determining their effort with the resulting amount of crime in the model, they are incentivized to be proactive in their policing through the reward function (f_c), where officers receive employment, monetary, and/or utility returns based on the arrest type they clear, consistent with descriptions of police overtime and promotions (Samaha, 2017; Hughes, 2020; Mastrorocco and Ornaghi, 2020). Officers take reported crime, not the true level of crime, as given because they are unable to police beyond the offenses known.

[8] Table 1.1 uses the cost of crime estimates from Chalfin and McCrary (2018). Reflecting the broader literature, these estimates only account for the direct victimization costs—rather than the broader costs imposed on a community—and consequently, they are an undercount.

reported by the community. Officers also face convex costs related to effort, which is parameterized by the function $\frac{1}{2}\phi e_c^2$.

Taking the derivative of the utility function with respect to e_c (i.e., effort for either low or high social cost crimes), the first-order condition is:

$$\frac{\partial U}{\partial e_c} = \frac{df_c(A_c(e_c; F, RC_c))}{dA_c} \cdot \frac{\partial A_c}{\partial e_c} - \phi e = 0$$

Thus, the equilibrium effort ($e^\star$) for an officer satisfies

$$\text{(1.1)} \qquad \implies e^\star = \frac{1}{\phi}\left[\frac{df_c}{dA_c} \cdot \frac{\partial A_c}{\partial e_c}\right]$$

where equilibrium effort for a specific offense is determined by the marginal benefit of arrest to the officer, marginal arrest generated for an additional work hour of effort, and the cost of effort.[9]

The Ferguson Effect—or reduced officer effort as a result of increased scrutiny from a high-profile, officer-involved fatality—manifests itself through exogenous increases in the cost of effort multiplier, ϕ. Intensified scrutiny from media coverage, community attention, and protests may result in additional psychic costs, as officers experience disutility from protests and interactions with a community that may be apprehensive and critical of them. Additionally, officers may perceive higher costs of mistakes and misconduct, serving a more vigilant community where civilians are recording officers as they make stops and more journalists are attentive to policing issues. This channel is evinced by anecdotes such as the Chicago PD officer who did not draw her firearm while being assaulted because of fear of media coverage (Hawkins, 2016).

In other police officer objective function models, such as Shi (2009), the ϕ parameter most closely resembles increases in expected oversight costs from the likelihood of a complaint being filed, the officer being found guilty, and the penalty imposed as a result. However, the expected monetary or employment penalty is likely to be low, which is why there is no explicit civil, criminal, or administrative punishment mechanism built into the model (Schwartz, 2014; Rushin, 2017; Rushin, 2019; Grunwald and Rappaport, 2020).[10]

Thus, the comparative static of interest is $\frac{\partial e^\star}{\partial \phi}$: how does the rise in the marginal cost of effort—from intensified scrutiny—affect equilibrium effort? Let $e^\star$ be the choice of effort

[9] Aside from the intensive margin of effort within a shift, higher effort exertion may be represented by taking on longer shifts or doing overnight shifts, both of which are physically and mentally taxing.

[10] In the collective bargaining agreements negotiated between police unions and many US cities, there often are clauses in the contract requiring the city to indemnify police officers (Schwartz, 2014; Schanzenbach, 2015). Schwartz (2014) found that "governments paid approximately 99.98% of the dollars that plaintiffs recovered in lawsuits alleging civil rights violations by law enforcement. Law enforcement officers in my study never satisfied a punitive damages award entered against them[...]." These lawsuits include wrongful death claims, which usually range into the millions of dollars. These cases are common in these high-profile OFs, usually resulting in the family settling with the city. Even in cases where the police department admits wrongdoing, criminal charges for officers are rare, and very few result in convictions. This New York Times article provides a breakdown of wrongful death settlement amounts and police officer accountability by incident.

that maximizes utility for the given community cooperation, reported crime, and scrutiny parameters. By the envelope and implicit function theorems,

$$\frac{\partial e_c^\star}{\partial \phi} = -\frac{\dfrac{\partial^2 U}{\partial e_c \partial \phi}}{\dfrac{\partial^2 U}{\partial e_c^2}}$$

Because the scrutiny parameter only affects the cost function, the mixed partial derivative with respect to ϕ is

$$\frac{\partial^2 U}{\partial e_c \partial \phi} = -e_c \implies -\frac{\partial^2 U}{\partial e_c \partial \phi} = e_c > 0$$

and the second-order derivative with respect to e_c is

$$\frac{\partial^2 U}{\partial e_c^2} = \underbrace{\frac{d^2 f_c}{dA_c^2}}_{=0} \cdot \underbrace{\left[\frac{\partial A_c}{\partial e_c}\right]^2}_{>0} + \underbrace{\frac{df_c}{dA_c}}_{>0} \cdot \underbrace{\frac{\partial^2 A_c}{\partial e_c^2}}_{<0} - \underbrace{\phi}_{>0} < 0$$

confirming that the utility function is concave. $\frac{d^2 f_c}{dA_c^2} = 0$ because the marginal reward for each arrest is set proportional to the social cost and does not diminish.[11]

Consequently,

$$\frac{\partial e_c^\star}{\partial \phi} = \frac{e_c}{\dfrac{d^2 f_c}{dA_c^2} \cdot \left[\dfrac{\partial A_c}{\partial e_c}\right]^2 + \dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} = \frac{e_c}{\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} < 0$$

Despite there not being a direct measure of officer effort, because arrests are (weakly) monotonically increasing in effort, decreased effort can be exhibited by reductions in arrests holding all else constant. Explicitly,

$$\frac{\partial A_c^\star}{\partial \phi} = \frac{\partial A_c^\star}{\partial e} \cdot \frac{\partial e_c^\star}{\partial \phi} = \frac{\partial A_c^\star}{\partial e} \cdot \frac{e_c}{\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} < 0 \tag{1.2}$$

Outside their initial probationary period when they first join the force, firing or demoting patrol officers is quite difficult, since they usually are protected through union contracts (Schanzenbach, 2015; Rushin, 2017; Rushin, 2019; Hughes, 2020). Officers who are fired often end up re-hired in nearby jurisdictions, as complaint and disciplinary records of officers are often kept hidden, including from prosecutors, public defenders, and even other police departments (Friedersdorf, 2015; Bies, 2017; Kelly, Lowery, and Rich, 2017; Grunwald and Rappaport, 2020). Schanzenbach (2015) highlights that "the expense and low success rate deter cities from pursuing misconduct." Moreover, legal protections and officer contracts unequivocally encourage proactive policing and high arrest rates, whether through qualified immunity, indemnification, overtime compensation, or future career prospects. The theoretical underpinning for these protections is to shift the risk- and cost-burden from the officer to the city. Consequently, the protection from legal ramifications and costs allow a rational officer to police proactively, thereby reducing crime.

[11] If I relax this assumption, allowing $\frac{d^2 f_c}{dA_c^2} < 0$, then $\frac{\partial^2 U}{\partial e_c^2}$ would still be less than zero.

Additionally, because $\dfrac{\partial A_L^\star}{\partial e_L} > \dfrac{\partial A_H^\star}{\partial e_H}$ and $\dfrac{df_H}{dA_H} \gg \dfrac{df_L}{dA_L}$, where differences between the highest and lowest social cost crimes are several orders of magnitude, then

$$\left|\frac{\partial A_L^\star}{\partial \phi}\right| > \left|\frac{\partial A_H^\star}{\partial \phi}\right| \tag{1.3}$$

Therefore, an exogenous increase in scrutiny, which increases the MC of effort, reduces officer effort and arrests in equilibrium. Further, model predicts that this decrease in arrests will be larger for offenses that produce lower marginal benefit to the officer (i.e., generally low social cost crimes). I explore this mapping empirically in section 1.5 when testing for reductions in arrests after controlling for crime, population, and a number of other department- and municipality-level controls.

However, in addition to policing behavior, a high-profile OF may affect community cooperation (Sunshine and Tyler, 2003), which has its own impact on arrests. First, I need to determine the sign of the second-order mixed partial derivative:

$$\frac{\partial^2 U}{\partial e_c \partial F} = \underbrace{\frac{d^2 f_c}{dA_c^2}}_{=0} \cdot \underbrace{\frac{\partial A_c}{\partial F}}_{>0} \cdot \underbrace{\frac{\partial A_c}{\partial e_c}}_{>0} + \underbrace{\frac{df_c}{dA_c}}_{>0} \cdot \underbrace{\frac{\partial^2 A_c}{\partial e_c \partial F}}_{>0} > 0$$

Thus, the sign of the comparative static is

$$\frac{\partial A_c^\star}{\partial F} = dfrac\partial A_c^\star \partial e \cdot \frac{-\dfrac{\partial^2 U}{\partial e_c \partial F}}{\dfrac{\partial^2 U}{\partial e_c^2}} = \frac{\partial A_c}{\partial e_c} \cdot \frac{-\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial F}}{\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} = \frac{-\dfrac{\partial A_c}{\partial e_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial F}}{\dfrac{\partial^2 A_c}{\partial e_c^2} - \phi \Big/ \dfrac{df_c}{dA_c}} > 0 \tag{1.4}$$

suggesting that decreases in community cooperation from a high-profile OF result in reductions in equilibrium arrests. Scrutiny and community cooperation are separate mechanisms that both result in declines in arrests, and therefore the latter could confound the Ferguson Effect estimate. However, for the least serious offenses (L), such as marijuana possession or disorderly conduct, the arrests are largely on-view, meaning that those clearances require the little to no community cooperation (i.e., the officer does not have a previous incident report or a warrant, but rather detains the suspect based on something they witnessed), whereas the opposite is true for more serious offenses (e.g., homicide)[12]. Because $\dfrac{\partial^2 A_H}{\partial e_H \partial F} \gg \dfrac{\partial^2 A_L}{\partial e_L \partial F} \approx 0$, then

$$\frac{\partial A_H^\star}{\partial F} > \frac{\partial A_L^\star}{\partial F} \approx 0 \tag{1.5}$$

Therefore, if empirical findings show reductions in arrests and they are concentrated in more serious offenses (H), then the reductions are a result of a decline in community cooperation after a high-profile OF. Conversely, if the reductions are primarily in low marginal

[12]In fact, in California, police officers are currently not allowed to arrest a suspect for a misdemeanor crime, unless they have probably cause to believe they witnessed the crime in their presence ("on-view") or they have a warrant. Here is the relevant legal section.

benefit arrests (L), then the resulting mechanism is a decline in officer effort from additional scrutiny.

Finally, since arrests are based on police officers' knowledge of reported crime, it is important to understand how RC_c is affected following a high-profile OF. Offending, one part of RC_c, could decrease based on a perceived rise in the cost of engaging in crime (Cloninger, 1991), or it could increase due to reductions in apprehension risk (Becker, 1968) or reactions to the high-profile fatality itself, where diminished perceptions of police legitimacy and procedural justice lead to declines in legal compliance (Persico, 2002; Kirk and Papachristos, 2011; Rosenfeld and Wallman, 2019). Similarly, civilian reporting of crime, the other component of RC_c, may change after a fatality (Desmond, Papachristos, and Kirk, 2016). Since I am interested in changes in ϕ holding all else constant, in the main empirical specification, I directly control for changes in RC_c where possible, closing an intermediary channel where arrests may fluctuate from changes in offending or reporting. It is the appropriate control because officers can only plausibly police offenses they aware of. Consequently, using the predictions in Equation 1.3 and Equation 1.5, the focus is on the two remaining channels: community cooperation and scrutiny.

Nevertheless, to understand how changes in reported crime after a high-profile OF may affect police effort and arrests, I provide the comparative static analysis below.

$$\frac{\partial^2 U}{\partial e_c \partial RC_c} = \underbrace{\frac{d^2 f_c}{dA_c^2}}_{=0} \cdot \underbrace{\frac{\partial A_c}{\partial RC_c}}_{>0} \cdot \underbrace{\frac{\partial A_c}{\partial e_c}}_{>0} + \underbrace{\frac{df_c}{dA_c}}_{>0} \cdot \underbrace{\frac{\partial^2 A_c}{\partial e_c \partial RC_c}}_{>0} > 0$$

The sign of the comparative static is

$$\frac{\partial A_c^\star}{\partial RC_c} = dfrac\partial A_c^\star \partial e \cdot \frac{-\dfrac{\partial^2 U}{\partial e_c \partial RC_c}}{\dfrac{\partial^2 U}{\partial e_c^2}} = \frac{\partial A_c}{\partial e_c} \cdot \frac{-\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial RC_c}}{\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} = \frac{-\dfrac{\partial A_c}{\partial e_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial RC_c}}{\dfrac{\partial^2 A_c}{\partial e_c^2} - \phi \Big/ \dfrac{df_c}{dA_c}} > 0 \tag{1.6}$$

Expectantly, as officers are aware of more crimes, the equilibrium level of arrests increases—or vice versa. But, that maxim does not strictly hold for the least serious offenses (L), where $\dfrac{\partial^2 A_H}{\partial e_H \partial RC_c} \gg \dfrac{\partial^2 A_L}{\partial e_L \partial RC_c} \approx 0$. Given the significant number of L offenses that go without arrest (e.g., marijuana possession), an increase in them is unlikely to change the marginal arrest of effort, especially compared to higher social cost crimes, where police services may be reallocated to address the additional cost burden.[13] As a result,

$$\frac{\partial A_H^\star}{\partial RC_c} > \frac{\partial A_L^\star}{\partial RC_c} \approx 0 \tag{1.7}$$

[13]In fact, section 1.4, I discuss how many L offenses are not even recorded by the FBI, unless they result in an arrest, since reports would be substantially undercounted.

1.4 Data Description

1.4.1 Arrest and Crime Data

The study combines government data on arrests, reported crime, and demographic information with crowd-sourced data on officer-involved fatalities. The backbone of this study uses large administrative datasets from FBI's Uniform Crime Report, which have nearly every police department in the US. For the primary outcome of interest and controls, I use the FBI's Offenses Known and Clearances by Arrest dataset, and the Arrests by Age, Sex, Race dataset (hereafter, UCR and ASR, respectively). Specifically, I use the versions cleaned and organized by Jacob Kaplan (Kaplan, 2018; Kaplan, 2019). The UCR data contains monthly arrests for Part I offenses (murder, robbery, assault, burglary, theft, motor vehicle theft) and known offenses (i.e., reported crime).[14,15]

I use the ASR in the analysis because it has data on Part II arrests, a broader offense classification set than Part I—notably containing the least costly crimes, such as disorderly conduct or marijuana possession.[16] In addition to analyzing some Part II offenses separately, I create an index of low marginal benefit (MB) arrests—Part II clearances that produce little return to the officer in terms of promotion or compensation, and mostly have lower social costs of crime victimization: curfew/loitering, disorderly conduct, public drunkenness, Driving Under the Influence (DUI), liquor violations, marijuana possession, marijuana sale, prostitution, suspicion, vagrancy, and vandalism.[17] Despite the offense having potentially higher social cost, DUI arrests are included in the low MB index because officers have anecdotally indicated that there is less priority placed on their enforcement relative to other offenses. By compiling certain Part II offenses into an index, I reduce the false discovery rate—the likelihood of finding falsely significant, spurious effects (Anderson, 2008). There also may be larger effects on these arrests because officers have more discretion in determining whether to make an arrest.

Importantly, the ASR reports arrests by race. The main results are run using the ASR dataset at the agency-month-race level, after merging over the crime controls from the UCR. I focus on the arrests of black and white suspects. In order to ensure the data maps accurately to the jurisdictions, I clean both the UCR and ASR using rules described in subsection A.1.2, following Evans and Owens (2007) and Mello (2019).

For this study, I use monthly arrest counts, disaggregated by race and crime type, to analyze policing effort in 2,740 police departments with 53 treated municipalities (73 OFs overall) between 2005–2016. The identification strategy relies on the quasi-random timing

[14] The exact rules for how they record crime and arrest data are further explained in subsection A.1.1. It is important to highlight that theft and motor vehicle theft are mutually exclusive categories.

[15] For this study, I exclude rape, since the FBI's definition of rape was modified in 2013. Subsequently, changes in rape offense or arrest around that time could not be separated from the change in definition. I also do not include arson because it's uncommon and provided on a different supplemental dataset.

[16] Part II offenses usually have no distinct victims, and subsequently, they are far less likely to be reported by the community. Thus, the FBI only reports them if they result in arrest.

[17] A weapon arrest is a violation of a local law or ordinance, "prohibiting the manufacture, sale, purchase, transportation, possession, concealment, or use of firearms" (FBI, 2004). Suspicion arrests are detentions that are not tied to any specific offense and the offender is released without formal charges being placed.

of the high-profile OFs to induce a plausibly exogenous shock of increased scrutiny to the local police department. Table 1.2 reports that on average treated jurisdictions have lower arrest rates for whites than control jurisdictions, as well as higher rates for blacks across all types of crime: violent, property, and the low marginal benefit arrests. Across race, the arrest rates for blacks are substantially higher than whites for all crime types. Typically, treated departments are policing larger and more populous cities with an average population of nearly 500,000 compared to nearly 150,000. These treated cities have about 11 percentage points more blacks and fewer whites, while having similar educational attainment and slightly higher poverty rates, driven by black poverty.[18]

1.4.2 Officer-Involved Fatality Data

There is no reliable government database for officer-involved fatalities. The government estimates are from the FBI Supplementary Homicide Report within the category of "Justifiable Homicides." These estimates are extremely undercounted, resulting in non-governmental estimates having up to three times the number of fatalities (Krishnan, McCrary, and Premkumar, 2017). The most comprehensive dataset is from Fatal Encounters (FE). FE has managed to create the most sophisticated collection system: collating data from Freedom of Information Act requests, web-scraping of news sources, and using paid researchers to run additional searches and data checks from public sources.[19] The dataset contains individual information on victim age, race, gender, location of death, cause of death, and a brief description of the incident. With over 1,100 incidents per year, officer-involved fatalities are quite common, comprising 7–9% of annual homicides.

For this study, I focus on the subset of fatal encounters that generate comparatively high amounts of media coverage, ones that could plausibly engender enough local scrutiny toward officers that they change behavior. [20]

I create this high-profile subset by scraping the number of news articles written about each incident. In subsection A.1.3, I describe the scraping procedure I use in more detail. Figure 1.1 provides a histogram of the distribution of news articles in bins of 500 articles.[21] The two obvious break points in the distribution are OFs that are covered more than 1,000 times and those covered more than 2,500 times. After manual inspection of each OF above 1,000 articles, many of them report on a community action after the death, often a protest, in line with the mechanism of increased scrutiny. While choosing the 2,500 threshold may

[18]These summary statistics required an imperfect merge of the 2014 American Community Survey, and as a result, it uses only a subset of the observations used in the main analysis.

[19]To find out more information, see here on the Fatal Encounters website.

[20]Many fatalities do not result in significant media coverage and subsequently produce limited community awareness. For instance, any empirical design including the largest US cities and a before-after analysis timing would be hampered, as they occur

In largest cities in the US, they occur so frequently that they hamper any empirical designs that rely on before-after shocks.

[21]The bin from 0-500 articles is excluded since it contained over 99.9% of officer-involved fatalities, rendering the rest of the histogram invisible. Similarly, the distribution is heavily right skewed and unable to depict some of the most high-profile OFs. There are nine OFs for which there are more than 20,000 news articles, which are inputted into the farthest right bin.

lead to larger treatment effects, there is reduced statistical power due to the small number of OFs. Therefore, I define "high-profile" as having at least 1,000 news articles written about the incident.[22]

To see how the fatalities vary across space, Figure 1.2 provides a map of the US where city dot size corresponds to the number of events that occurred during the study frame. The geographical dispersion of the treated jurisdictions is ideal for conducting a national analysis of the Ferguson Effect. As expected, there are a few clusters in population centers like Chicago, New York, and Los Angeles. However, the salient takeaway is the numerous incidents in relatively non-populous localities, particularly in the southern region of the US. This finding suggests that incident-specific characteristics may explain how high-profile an OF is, such as whether it happens to be captured on video or seen by witnesses, as demonstrated in the case of Walter Scott in North Charleston—one of the highest profile OFs in the dataset.

The proliferation of cellphone cameras and the advent of social media has played a factor in how much attention fatalities receive (Lacoe and Stein, 2018; Battaglini, Morton, and Patacchini, 2020), with many of the fatalities toward the end of sample frame (Figure A.1). One concern is the advent of social media may have changed the locality of these incidents, dispersing public scrutiny to many jurisdictions. To assess this issue, I integrate new data on Black Lives Matter (BLM) protests from August 2014 to August 2015.[23] I find that there is distinctly more protest activity in treated police department municipalities than control municipalities, with an average of 4.74 protests versus 0.78 respectively (Table 1.2). There are sizable differences in attendance as well, with control jurisdictions having an average turnout of just over 200 people versus over 1,200 people in treated jurisdictions. I discuss this issue further in subsection 1.5.3, illustrating through empirical tests that certain types of spillovers do not seem to be present. However, if there is a spillover effect, it would likely attenuate the main findings, resulting in them being a lower bound of the true effect.

Finally, I use data from Google Trends to modify the timing of each fatality to align with when community members started searching for the incident to reflect a plausible beginning of scrutiny. This modification is necessary for OFs such as Laquan McDonald, who died in October 2014 but the Chicago community did not become fully aware until November 2015, when a video of the shooting was released. I discuss this specific case in more detail in section 1.6. Overall, there are 11 events in the analysis that have timing changed.

The main analysis uses 73 high-profile OFs spread across 53 distinct jurisdictions. The maximum number of OFs in a single jurisdiction is six (in Los Angeles). Table 1.3 illustrates that high-profile OFs typically involve young to middle-aged black men who die from gunshot wounds. The victims are predominantly unarmed and not suffering from mental impairment,

[22]Because the focus is on high-profile fatal encounters, I am not concerned with missing data in FE. Conducting manual checks for victim names mentioned in prominent articles on fatal encounters with police, I confirm full reporting of these OFs.

[23]I use replication data from Trump, Williamson, and Einstein (2018), who source their information from the online platform Elephrame (2018). Elephrame has collated data on over 2,700 BLM protests across the country, from protests of over 50,000 to gatherings of fewer than ten, with related news article links, subjects of the protest, attendance estimates, and location.

and thus plausibly consistent with a public perception of injustice, which could be reflected in the intensified scrutiny from additional media coverage of the incident.

1.5 Empirical Strategy and Results

In this section, I describe the empirical specifications and the corresponding results related to whether increased scrutiny affects arresting behavior. The empirical strategy leverages aspects of scrutiny directly by only including incidents that reach over 1,000 articles of media coverage and adjusting treatment time to when the local community first searches for the incident.[24] Thus, control jurisdictions include departments who experience no OFs or less publicized ones. subsection 1.5.1 discusses difference-in-difference (DD) and triple difference (DDD) estimation using only the highest-profile, officer-involved fatality per jurisdiction. Then, subsection 1.5.2 outlines why the "integrated" event study design is the preferred specification, providing the empirical model and the respective results. Both sections discuss heterogeneity in effects by offender race.

1.5.1 Difference-in-Difference and Triple Difference

1.5.1.1 Difference-in-Difference and Triple Difference Econometric Model

Using only the highest profile OF per police department, I estimate the effect on arrests, where the identifying variation is from the plausibly exogenous timing of fatalities. The DD estimating equation is

$$Y_{it}^{c} = \mu_i + \lambda_t + \rho(t)_i + X_{it}^{c}\beta + \theta(Treat * Post_{it}) + \nu NegBin_{it} + \alpha PosBin_{it} + e_{it}^{c} \tag{1.8}$$

where Y_{it}^{c} represents log arrests of crime type c in department i in month t.[25] Crime type c consists of a variety of violent, property, and low marginal benefit offenses, as well as those respective crime categories. The coefficient of interest is θ, which represents the approximate percent change in arrests in the 16 months after a fatality in the treated department relative to control departments in the pre-period. To have θ compare periods just before and after the event, I include binned endpoints outside the event window, where $NegBin_{it}$ and $PosBin_{it}$ control for 8 months before and 16 months after an OF respectively (McCrary, 2007).[26] I

[24] A concern from creating a news article cutoff for treatment is that media coverage may be correlated to other time-varying characteristics that affect the timing of treatment, such as the population of a city or the time period of analysis, such as post-Ferguson (Figure A.1). The empirical analysis addresses the concern by including department and month-of-sample fixed effects, while controlling for population, reported crime, and county-specific linear trends.

[25] To not lose data where arrests and crime counts are zero, I transform the distribution to log(variable+1). Running the specification using the inverse hyperbolic sine transformation is nearly indistinguishable in terms of coefficients and standard errors. For simplicity, I focus only on the log+1 transformation.

[26] With heterogeneous event dates, balanced panel data becomes unbalanced in event time. Subsequently, each municipality would have a varying amount of leads and lags. Since I prefer θ to compare periods just before and after the event, introducing binned endpoints controls periods before and after the broader event window, while also limiting changes in estimates from sample composition.

use department and month-of-sample fixed effects (μ_i, λ_t), and county-specific linear trends ($\rho(t)_i$).[27] Standard errors are clustered at the police department level.

X_{it}^c is a vector of controls for yearly log population, log reported violent and property crime, and log reported crime c if c is a Part I offense.[28] I use these covariates because I am interested in the effects of scrutiny on policing effort, which I measure through changes in arrests after controlling for other intermediary channels such as civilian crime reporting and offending behavior. Assuming arrests are (weakly) monotonically increasing with the amount of reported crime, it would not be a decline in policing effort if arrests are reduced mechanically because of a drop in crime after an OF, which could occur if criminals perceive higher risk associated with altercations with police (Becker, 1968; Cloninger, 1991).[29] I discuss the potential theoretical mechanisms that may induce changes in arrests after an OF, as well as their predicted empirical effect by offense type, more thoroughly in section 1.3. However, as discussed in section 1.7, the main results are generally not sensitive to the inclusion of the crime control.

Table 1.3 provides evidence that black OFs are higher profile with an average of over 20,500 news articles written about each incident compared to the next highest of white OFs at 3,000. It is possible that community scrutiny can operate along racial dimensions, increasing the marginal cost of effort for arresting black suspects without significant changes for white suspects. To test for that, the DDD specification is quite similar to the DD model

$$\begin{aligned} Y_{it}^c &= Black * \mu_i + Black * \lambda_t + Black * \rho(t)_i + X_{it}\beta + \omega(Black * Treat * Post_{it}) \\ &\quad + \delta(Treat * Post_{it}) + \nu(Black * NegBin_{it}) + \alpha(Black * PosBin_{it}) + e_{it}^c \end{aligned} \tag{1.9}$$

with the interaction of black arrests ($Black$) on the Treat*Post variable, the binned ends of the event window, and the fixed effects. For the DDD, the coefficient of interest is ω on the Black*Treat*Post variable, identifying racial differences in the Treat*Post coefficient.

1.5.1.2 DD and DDD Results

In the 16 months after the highest profile OF per jurisdiction, statistically significant drops in log arrests—after controlling for changes in log crime—represent evidence of the effect of scrutiny on policing effort. Using the ASR sample, in Figure 1.3, I present the DD coefficients from separate regressions of individual offense types in descending order of the social cost of crime. None of these regressions have significant differences between the pre-treatment trends of control and treated departments, providing a partial test of the

[27] I employ county trends to control for the steady decline in arresting patterns during my sample frame. For example, certain municipalities gradually reduced how strictly they enforce certain drug laws. County-specific trends reflect the upstream effects prosecutors have by choosing what crimes to charge. If I use department-specific trends, the results are similar, but computationally more intensive.

[28] I do not have the adequate control of reported crime for each Part II arrest type. However, for each specification, I still control for the amount of violent and property crime.

[29] I do not set Y_{it}^c to be the arrest rate—the ratio of arrests over crime—because that presumes that the increase in arrests from a rise in crime is equivalent across offense types. Judging from the differing coefficient estimates on the crime control in these regressions, that is not the case (e.g., an increase in homicides does not result in the same increase in arrests as an increase in thefts).

identifying assumption: treated police departments would trend similarly to control police departments in the absence of the high-profile OF. I first focus on the changes in arrests of more serious crimes that cause significant social costs. After a high-profile, officer-involved fatality, the involved police department experiences no change in violent crime arrests (murder, aggravated assault, robbery) with the murder and aggravated assault coefficients at near zero.[30] However, property crime arrests (motor vehicle theft, burglary, and theft) have a statistically significant decrease of 7.1%, driven by theft (-9.2%)—the least socially costly property offense (Table 1.1). There is also a more modest and suggestive reduction of burglary arrests (-5.2%).

Shifting to least serious offenses, I see much sharper reductions in arrests for disorderly conduct (14.7%) and marijuana possession (14.2%). These decreases help drive the overall decline in low marginal benefit arrests of 9.8%, relative to the eight months before an OF (Table A.3).[31] This effect pattern is broadly reflective of the reductions in policing effort occurring along less socially costly crimes.

In Table A.1, I present the DDD race results for violent crime arrests in Column (3), while Column (4) runs the test of racial parallel pretrends. The table highlights the lack of racial differences in the (lack of) change in violent crime arrests. On the other hand, Table A.2 highlights suggestive evidence that black theft arrests fall by 5.8% relative to white theft arrests. The DDD empirical design does not uncover any significant racial differences for low MB clearances (Table A.3). Overall, the changes in arrests following an OF do not appear to vary significantly by race.

1.5.2 Event Study Design

1.5.2.1 Event Study Econometric Model

Since the effects on arrests could vary over the event window, the DD estimates are not as informative as results from an event study. The canonical models of event studies, such as Jacobson, LaLonde, and Sullivan (1993) and McCrary (2007), replace the Treat*Post coefficient with event time dummies, where there is a dummy for each unit of time away from the event within a specified event window. Binned event time dummies pool the estimates outside of the event window, controlling for the periods before and after an OF that are outside of the window of interest. For this study, the preferred event window is -3 to 5 quarters (-9 to 17 months) around the event with a time dummy for each quarter. The dummy coefficients are normalized to the period before the event.

However, unlike in McCrary (2007), the study involves observational units potentially having multiple events (e.g., Los Angeles Police Department, with six high-profile police fatalities), which complicates the design. There are a few proposed solutions to accommodate

[30] The actual DD tables, such as Table A.1, are in the appendix. Column (1) provides the DD estimate, while Column (2) runs the test of parallel pretrends. The "Pretrend Test" variable is generated by splitting the pre-treatment period, the eight months before an OF, in half and testing for statistically significant differences in arrests between 5-8 months versus 1-4 month before an OF. If the pretrend test is significant, then the identifying assumption of the DD is not met.

[31] For more explanation on the low MB index, refer to section 1.4.

a multiple treatment design. One suggestion is narrowing the event sample to the first or largest treatment per observational unit, where the argument is that all other treatments afterward could be potentially correlated to that one. However, by using only the first treatment, the statistical power from additional events is lost—from 73 to 53 events—taking away the variation in timing of later treatments. Moreover, drops in arrests from future events would be incorrectly attributed to later event dummies.

The second solution comes from the finance literature and it is used more regularly in recent economics papers (Lafortune, Rothstein, and Schanzenbach, 2018). Each observational unit gets expanded by the number of events for that unit. For example, since I have data on Los Angeles from 2005–2016 and it has six events, I would expand the dataset to include six Los Angeles, each with their own unique event date. I would then cluster at the observational unit level, allowing for correlated shocks across expansion sets. Similar to the first solution, there are unaccounted treatments in the pre- or post-trend of each expanded observational unit. Both designs could result in biased estimates of pre- or post-trends (Sandler and Sandler, 2014).

To address limitations in previous designs, I implement an "integrated" approach, which adapts aspects of designs in McCrary (2007), and Sandler and Sandler (2014). Unlike in the "first event" or "expansion" designs, multiple time dummies can be turned on (e.g., if an observation is 1 quarter past an event in 2014 and 5 quarters past an event in 2013, then the +1 and +5 dummies should equal one). The preferred event window is -3 to 5 quarters (-9 to 17 months) around the event with a time dummy for each quarter. Because certain departments experience multiple OFs, the binned endpoints of the event window need to be the sum of pooled time dummies outside of the event window, taking on values up to six in the dataset for Los Angeles. Consequently, I define a flexible "dummy" setup of D_{it}^{j}[32]:

$$D_{it}^{j} = \begin{cases} \sum_{j=-J}^{a-1} \sum_{m=1}^{n_i} \mathbb{1}(t = \tau_{im} + j) \text{ for } j < a \\ \sum_{m=1}^{n_i} \mathbb{1}(t = \tau_{im} + j) \text{ for } a \le j \le b \\ \sum_{j=b+1}^{J} \sum_{m=1}^{n_i} \mathbb{1}(t = \tau_{im} + j) \text{ for } j > b \end{cases} \tag{1.10}$$

For the function above, τ_{im} is the date of the event m in jurisdiction i, a and b define the event window of choice, t is the month-of-sample, and $j \in \{-143, ..., 143\}$ represents the lead/lag counter, where J is the maximum value the lead/lag counter takes for the sample of 2005–2016. $\mathbb{1}$ is an event indicator function. If there is no event date (i.e., police department i did not experience a high-profile OF), then τ_{im} is missing and then the function always returns zero. Department i experiences n_i events.

Figure 1.4 shows a visual example of the dummy structure in months, using simulated data of a department with two officer-involved fatalities: one in January 2006 and the other

[32]For simplicity, the shown stepwise function is a representation of the month event time dummies, which I pool to create the quarter time dummies. Technically, the function is not a dummy function since it outputs numbers aside from zero and one.

in May 2006—the highlighted months. Thus, the event time dummy zero is turned on for those respective months (the highlighted cells). Since May 2006 is also four months after an OF occurred in the municipality, it needs have the fourth lag dummy turned on. However, since the event window is between -3 to 3 months around the event, all events outside three months get pooled together in a negative and positive bin (-4+ and 4+), so the positive bin is turned on for May 2006 (and each month afterward). Only the negative and positive bins can take on values greater than one. Since the bins only serve as controls for time periods outside the preferred event window, the coefficients of interest in the event window bear the same interpretation as standard designs. As shown in Figure 1.4, there are missing values for a few event time dummies, at the beginning of the simulated data frame, because of data limitations. For example, if I only have data from January 2005 to October 2006, I have no information on whether an OF happened in 2004, so the corresponding lags (event time dummies 1 to 3) need to be set to missing. Thus, the remaining data which the regression uses for its estimates is denoted by the green rectangle. In practice, I use a wider event window of -9 to 17 months around an event, and I pool those month dummies into quarter dummies (-3 to 5 quarters) to increase precision. The identifying assumption for the empirical strategy is the same as the DD: the timing of treatment is plausibly random, and absent experiencing a high-profile OF, control and treated departments would trend similarly.

To understand how police officers' arrests shift following an OF, the event study design can non-parametrically estimate changes in arrests after a high-profile OF by month, crime type, and race. Given parallel pre-treatment trends between treated and control departments, any change in arrests after an OF is reflective of the impact of the event on officer behavior. Following the DD model, the preferred estimating equation is

$$Y_{it}^{c} = \mu_i + \lambda_t + \rho(t)_i + X_{it}\beta + \sum_{j=a-1, j\neq -1}^{b+1} \theta_j D_{iq}^{j} + e_{it}^{c} \tag{1.11}$$

where the dummy structure (Equation 1.10) replaces the Treat*Post variable. The θ_j's are the coefficients of interest on event time dummies, D_{iq}^{j}, where I pool the month dummies into quarter dummies (q) to increase precision. The quarter before the event is dropped, and the coefficient estimates are normalized such that $\theta_{-1} = 0$. The other θ_j's represent the percent change in the outcome in the quarter before the event to j quarters after, relative to changes in the control jurisdictions, after controlling for yearly population and monthly reported crime. With all of the specifications, I cluster the standard errors at the department level, and I include department and month-of-sample fixed effects, as well as county-specific linear trends.

To estimate changes in arrests by race, I create a separate dummy set for both black and white arrests by interacting a black and a white dummy ($Black$, $White$) with the original dummy set, using the same integrated approach detailed above.

$$\begin{aligned} Y_{it}^{cr} &= Black * \mu_i + Black * \lambda_t + Black * \rho(t)_i + X_{it}\beta \\ &\quad + \sum_{j=a-1, j\neq -1}^{b+1} \chi_j * Black * D_{iq}^{j} + \sum_{j=a-1, j\neq -1}^{b+1} \gamma_j * White * D_{iq}^{j} + e_{it}^{c} \end{aligned} \tag{1.12}$$

The new outcome (Y_{it}^{cr}) is race-specific (r) log arrests. Both sets of coefficients of interest, χ_j and γ_j, are identifying the percent change in race-specific arrests from their race-specific quarter before the high-profile OF to j quarters after. Further, I interact the black dummy with the police department and month-of-sample fixed effects, as well as the county-specific linear trends, allowing for the ability to control for race-specific unobservables and trends.

1.5.2.2 Event Study Results

Unlike the DD estimates, the event study graphs integrate all of the high-profile OFs, including cases in which there are multiple OFs per police department, providing more events (73 in main analysis sample) and richer variation. Event study graphs have the coefficient on log arrests on the vertical axis and event time dummies as the horizontal axis, where period 0 is when a police department is involved in a high-profile OF. Figure 1.5 shows a reduction in all Part I arrests (violent and property crime plus simple assault and negligent manslaughter) that builds over the first few quarters in a downward arc shape, reaching statistically significance in Q3 and then returning to the initial equilibrium in Q5. The effects range from about -2% to -8%, which are being driven by a similar pattern of reductions of 3–10% in property crime arrests. On the contrary, violent crime arrests are not affected. There is a persistent and continual decline in low MB arrests, where the sharpest drop of 24% is over twice that of property crime arrests.

Consistent with the DD results, murder and aggravated assault exhibit no change in arrests, while there is a suggestive decline of around 5% in burglary arrests in Q1 (Figure 1.6). Theft arrests, again, show a significant decline of 4–13%, instigating the effects witnessed in property crime arrests. These changes are dwarfed by those of marijuana possession, and disorderly conduct, which contribute to the descending pattern in low MB arrests (Figure 1.7).[33] Marijuana possession arrests experiences the most prominent change, falling by 14–33% in Q1–Q5. Disorderly conduct arrests follow a similar pattern with reductions between 13–21% for Q1–Q5. These patterns are emblematic of sharp reductions in other low MB arrests after Q0, such as liquor violations (20–33%) or marijuana sale arrests (10–21%).

Since race-specific arrests tend to move jointly, as evidenced by the DDD, an event study can be more informative in displaying the effect heterogeneity. For most offense types, there is not racial heterogeneity in arrests. However, Figure 1.8 shows that theft arrests suggestively fall more for black offenders (4-18%) compared to white ones (3-9%). While the same trend does not occur in all low MB arrests, black arrests for marijuana possession suggestive drop more than white arrests. Although none of the individual coefficients were significantly different in either offense type, black theft arrests are jointly different than white arrests, but that is not the case for marijuana possession.

1.5.3 Threats to Identification

The integrated event study design controls for past local OFs–as well as department and month-of-sample fixed effects, and county-specific linear trends in arrests—to isolate the

[33] In a few graphs, the vertical axis of the log point scale may be shifted to accommodate larger confidence intervals from larger effects.

change in arrests from a current OF, after adjusting for changes in population and reported crime. Provided that treated and control jurisdictions exhibit parallel pre-treatment trends, then any change in arrest patterns would represent the impact of increased scrutiny from a high-profile, officer-involved fatality.

One remaining threat to the identification assumption of the analysis is time-varying, idiosyncratic shocks at the department level whose timing aligns with the OF. Specifically, if rapid local expansions of the police force occur around the same time as an OF, the effects from the OF may be confounded by the changes in police staffing. In 2009, the federal government substantially increased the funding for the Community Oriented Policing Services (COPS) program, which provided many more hiring grants for local police departments, so they could reduce crime while remaining financially solvent. If certain police departments rapidly expanded their police staffing and the timing coincided with a high-profile OF, crime and arrest levels may be partially endogenous to these expansions.

The amount of funding for the COPS program is particularly high in 2009 and only provides significant funds for the few years after the 2009 grant allocation (Mello, 2019). Assuming departments hire quickly after receiving the grant, there are only six municipalities with a high-profile OF from 2009–2011, and only one of the treated departments during that time period received a COPS grant (the Oakland Police Department). Provided that there are similar effects in Oakland as found on average in Mello (2019), there would be little to no effect on arrests, and there would be a decrease in crime, attributed to deterrence from increased police presence. Thus, there would be an upward bias on the coefficients of interest, suggesting any reductions in arrests would be a lower bound. Again, that is only if the variation is not absorbed by the controls used in the regression, such as reported crime, some of which determine whether departments got the grant in the first place. Nevertheless, it could be a concern that the pure control departments have problematic trends because a subset of them receive COPS grants to increase police staffing. Or conversely, the intensified scrutiny from an OF could result in a reduction in police staffing for the involved department.

Another concern is that the treatment exposure may not be department-specific, potentially leading to spillover effects in other cities (Morgan and Pally, 2016; Cheng and Long, 2019). If these spillovers occur nationally, they are already being controlling for by the month-of-sample fixed effects. Additionally, the results are robust to spillovers that occur across jurisdictions with similar populations or within a state, since the size and significance of effects are largely unchanged after interacting the population group or state with the month of sample. Understanding that the most likely scenario is adjacent geographic shocks, I directly test for changes in policing effort in nearby municipalities in subsection A.2.3, where I expand the geographic definition of treatment to the county, rather than just the involved police department. Separating the effects between the involved department and the spillover departments that reside in the same county, I reassuringly find that treated agencies exhibit the same arrest patterns for Part I offenses, without consistent patterns in the spillover jurisdictions aside from an occasional drop in the fifth quarter after an OF (Q5). If spillovers exhibit some other idiosyncratic pattern that is not being controlled for or has not been examined, then the coefficients of interest would be attenuated, assuming the spillover effect has the same sign as the effect on the treated city. In that case, the magnitude of the results would be a lower bound of the true effect. Overall, the findings largely substantiate the hypothesis that police departments that experience a high-profile OF experience much greater scrutiny (Table 1.2), with nearly two-thirds of them facing protests. Consequently, there are sharp drops in policing effort of less serious offenses, evinced by reductions in low marginal benefit arrests of up to 24%.

1.6 Discussion

These results broadly show reductions in arrests for less serious offenses without changes in arrests for more serious ones. Because I control for reported crime, I am able to rule out changes in offending behavior and civilian reporting as alternative channels. Given that reductions exclusively occur in less serious arrests, predominantly the low MB ones, the findings are consistent with scrutiny being the causal mechanism (Equation 1.3). The opposite pattern of reductions in only the more serious arrests (e.g., homicides) is not present, suggesting that changes in community cooperation are not driving the estimates. The model illustrates that the intensified scrutiny after a high-profile OF increases the marginal cost of effort for officers, potentially manifesting through increases in psychic costs or perceived costs of mistakes. As a result, they reduce effort, and in turn arrests, for offenses that have less monetary, personal, and/or professional return.

Given the recent decriminalization of the possession and (regulated) sale of marijuana in a handful of municipalities, these accentuated estimates are internally consistent with the less reward to the officer because of less cost to society. Importantly, the results are not being driven by changes in state law, as the findings are unchanged after interacting the month-of-sample fixed effects with the state. Further, there are other low MB arrests—such liquor violations and disorderly conduct—that did not have major changes in law, yet still exhibited sizable drops after high-profile fatalities. These clearances are low social cost crimes, where officers face the lowest marginal benefit from an arrest. Naturally, these are the cases where officers also have the most discretion in determining whether to make an arrest and they usually are on-view (i.e., an officer observes the offense and hence does not require much community cooperation). The lack of community cooperation necessary for low MB arrests—where the effects are most concentrated—is instrumental in disentangling the mechanism as scrutiny.

Moderate reductions in theft arrests—the least costly of Part I crimes (Table 1.1)—are still consistent with the theoretical prediction that the declines occur for lower marginal benefit arrests. I find additional assurance in that there are only small and suggestive drops in burglary arrests, while none for robbery or motor vehicle theft (Figure A.2)—both of which have over a magnitude more social cost per crime. Moreover, there is no evidence of officers significantly reducing arrests for weapon violations or the sale of heroine/cocaine—two Part II offenses that have a higher return and social cost. The results in Part II arrests, in concert with the findings on Part I arrests, are empirically consistent with the model's prediction of scrutiny being the relevant mechanism.

For robustness, subsection A.2.2 replicates the general Part I results using the UCR sample, where the changes in arrests are concentrated in theft.[34] Additionally, the results are qualitatively similar if I change the high-profile threshold to 2,500 articles (not shown).

The declines in effort are present across white and black arrests for the most part, suggesting that the scrutiny increases the marginal cost of effort for policing generally. But, among the categories where the steepest declines in effort are observed—theft for Part I, marijuana possession for Part II—the effects on arrest for black offenders are suggestively larger than white offenders. The motivating anecdote of the Ferguson Effect may explain the race results: Officers may not want to make public, street-level arrests of blacks, potentially substituting their time to patrolling in their vehicle. Racial divergence in arrests may be found by exploring this story more carefully with the additional incident-level data.

The focus on street-level clearances may also provide clarity in understanding the persistence of the decline in certain low MB arrests, such as marijuana possession or liquor violations (Figure 1.7). These arrests show no return to the original equilibrium effort for at least 1.5 years afterward, contrasting declines in arrests such as theft (Figure 1.6). The sustained transition to a lower equilibrium effort may be a result of officers enduring perpetual psychic costs or perceived error costs in street-level interactions for these low MB clearances. When examining arrest types that involve less public interactions (e.g., vehicle stops, incident reports, or warrants)—such as DUI (not shown), marijuana sale, or theft—the low point in officer effort is Q3 or Q4 before the effect abates toward the original equilibrium. This abatement may partially represent a direct shift from pedestrian encounters to traffic stops, as was the case in Chicago where pedestrian stops declined by over 80% in late 2015 and early 2016 while traffic stops increased by 200% (Hausman and Kronick, 2019). The low point in these arrests at Q3 or Q4 may align with latent scrutiny coming from the court proceedings of the involved officer(s) or released video evidence. In a handful of high profile cases, it took months to have an announcement from the prosecutor or the grand jury decision whether to charge the officer. Although it is rare that grand juries are convened to consider charges against an officer (and even rarer for them to be charged), the event sample is unique in that I subset to only high-profile events.

To further substantiate the impact of scrutiny from media coverage, community attention, and protests, consider the case study of Laquan McDonald's death: In October 2014, he was fatally shot in Chicago, where police incident report incorrectly described McDonald lunging at an officer with a knife. Given the misinformation, there was a subsequent lack of awareness and scrutiny, resulting in a marginal change in misdemeanor arrests (Figure 1.9). However, in February 2015, a journalist informed the public that the police department had a video of the incident but was not releasing it. He also included the official autopsy report, which raised additional questions since McDonald was shot 16 times—with some of bullets entering from his back (Kalven, 2015). Misdemeanor arrests decline by 12% in the month afterward, but subsequently return to the original arrest level. In November 2015, the Chicago Police Department were forced by a court order to release the video, which directly contradicted the initial incident report. Consequently, low-level arrests sharply declined by 7–25%, coinciding

[34] It should be noted that the analysis with the UCR has more treated and control departments, and does not include the county-specific trends.

with the spike in scrutiny—measured by local Google searches of the incident, protests, and over 71,000 news articles. The separability between the incident and community awareness, as well as the delayed reductions in low MB arrests, help negate other potential explanations such as a police response to the incident itself.

In the aftermath of these high-profile, officer-involved fatalities, I find substantial increases in *offending*: For violent crimes, Figure A.7 documents a statistically significant rise of 10–16% in robberies and murders for Q0–Q4. There is also evidence of smaller increases in property crime, driven by theft (Figure A.8). Importantly, the main results of this paper related to arrests are robust to dropping the crime controls (subsection A.2.5). The additional crimes, especially the rise in murder, cause a tremendous loss of welfare (Table 1.1). These results are consistent with Evans and Owens (2007), Chalfin and McCrary (2018), and Mello (2019), which find that violent crime is more sensitive to fluctuations in aggregate policing effort than property crime. However, because community awareness of the incident and reductions of officer effort are usually simultaneous, it is difficult to identify the cause of the offending response.

One important direction for future research is investigating whether these crime increases are the result of the reduction in policing effort for primarily lower-level offenses, where lower apprehension risk increases the expected payout of crime (Becker, 1968)—or alternatively, whether marginal offenders are reacting to the high-profile fatality itself and subsequent (lack of) legal proceedings, to the extent that diminished perceptions of police legitimacy and procedural justice lead to declines in legal compliance (Persico, 2002; Kirk and Papachristos, 2011; Rosenfeld and Wallman, 2019).

1.7 Conclusion

In this study, I seek to explore how increased scrutiny of a bureaucratic agency, one with tremendous discretion in its actions, affects its effort level. Specifically, I provide the first estimates of how high-profile, officer-involved fatalities affect the arresting patterns for the involved police department and crime in that jurisdiction. This phenomenon has been dubbed the Ferguson Effect, named after a high-profile fatality in Ferguson, Missouri. The research question is complicated by the possibility that, after an OF, there may be multifaceted mechanisms affecting arrest rates, such as (1) greater scrutiny of police, (2) reduced community cooperation in the clearance of the crime, (3) reduced civilian reporting of crime, and (4) changes in offending behavior. I develop a novel theoretical model of an officer's objective function that uses insights into the institutional details around policing to provide model predictions that are empirically testable. I use these predictions to guide the analysis, tracing broader patterns in the changes of arrests by offense type to determine whether effort declines and what the causal mechanism is.

To answer this question, I utilize an event study design to estimate effects when there are multiple "events" per observational unit, combining strategies from McCrary (2007) and Sandler and Sandler (2014). I run this integrated design on a large administrative dataset from the FBI that is merged with novel data on OFs. The empirical strategy directly leverages aspects of scrutiny by only including incidents that reach over 1,000 articles of media coverage and adjusting treatment timing to when the local community first searches for the incident, as measured by Google Trends. Further, I show that treated municipalities experience larger and more frequent protests.

This paper explores a largely unanswered question, and these findings suggest that officers *do* reduce their effort following a highly publicized OF, but not evenly across crime types. Consistent with the model prediction, in the presence of intensified scrutiny, officers curtail arrests for offense types that generate lower marginal benefit for themselves (whether monetary, personal, or professional). Theft arrests experience reductions of 4–13%, while there are declines of up to 24% in low MB arrests. The most marked change is in marijuana possession arrests, which drop 14–33% after Q0. Similar trends can be seen across a handful of low MB arrests, including disorderly conduct, liquor violation, and marijuana sale arrests. These low-level arrests are the easiest to reduce, as they usually are officer-initiated stops, where they have the most discretion in determining whether to intervene. Notably, these reductions are not present with higher return arrests, demonstrated by the insignificant changes in violent crime or more serious property crime clearances. Further, the results are robust to numerous specifications, including the exclusion of reported crime as a control.

These results suggest that intensified scrutiny leads to reductions in police effort for arrests of both black and white offenders, suggesting that the ϕ parameter in section 1.3 is not race-specific overall. However, among the margins where the steepest declines in effort are observed—theft for Part I, marijuana possession for Part II—there is suggestive evidence of larger effects for black offenders.

To the extent that high-profile, officer-involved fatalities represent exogenous increases in scrutiny and the marginal cost of effort, they are comparable the results to other policies that increase the marginal cost for officers, such as the implementation of a thorough and cogent civilian review board. In labor economics, shirking typically arises from incomplete contracts because of imperfect monitoring, and is taken as a socially costly. However, these reductions in policing effort may reflect an allocation of services that is more closely in line with the preferences of the community, who are unable to clearly signal and translate their preferences through voting and campaigning because of the bureaucratic structure of police departments (Battaglini, Morton, and Patacchini, 2020). Given that the sharpest and most persistent drops in arrests are from low social cost offenses such as marijuana possession, liquor violations, and disorderly conduct, these reductions in arrests could be socially beneficial, after accounting for the associated social costs of policing and incarceration. However, if the drop in the least serious arrests does lead to increases in more serious offenses, as proponents of "broken windows" theory argue, then the crime increases could offset the welfare gains.

Figures

Figure 1.1: Histogram of News Articles on Officer-Involved Fatalities (2005-2016)

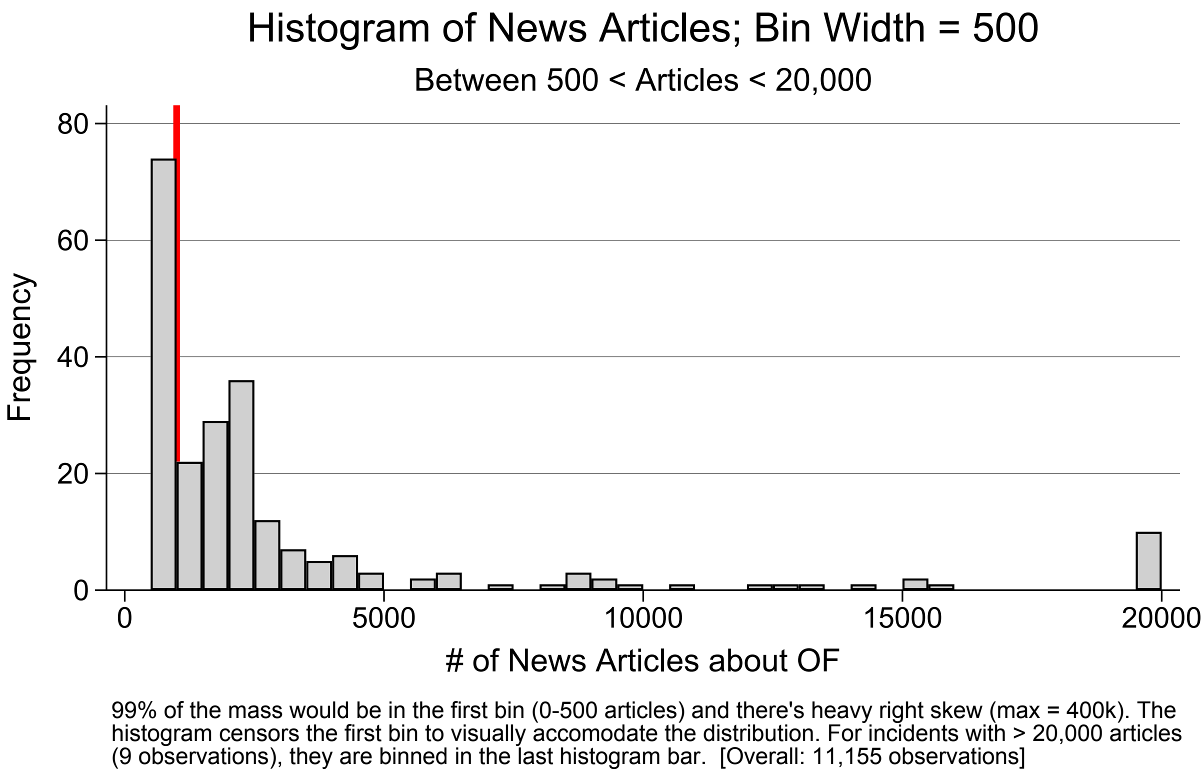
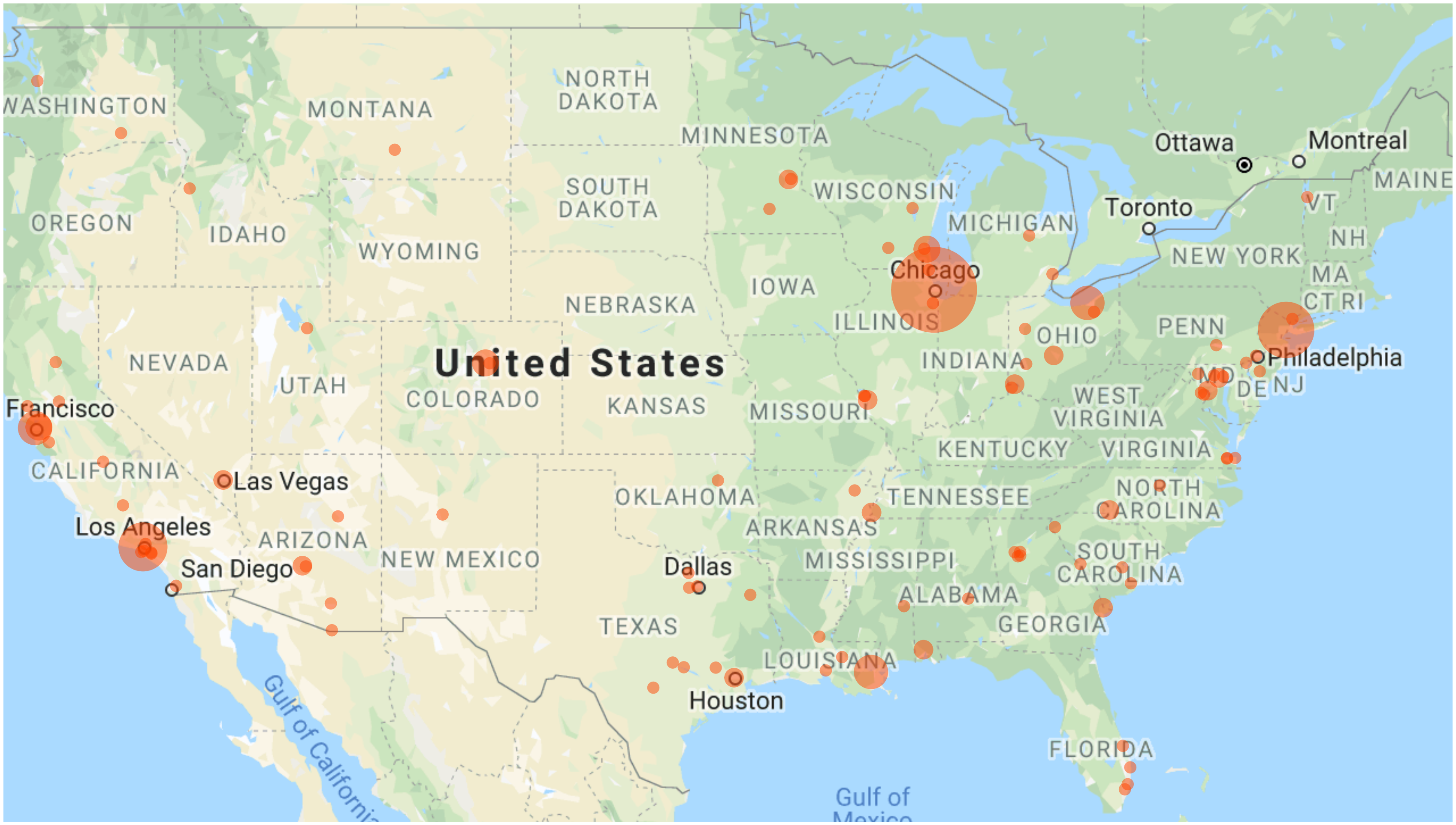
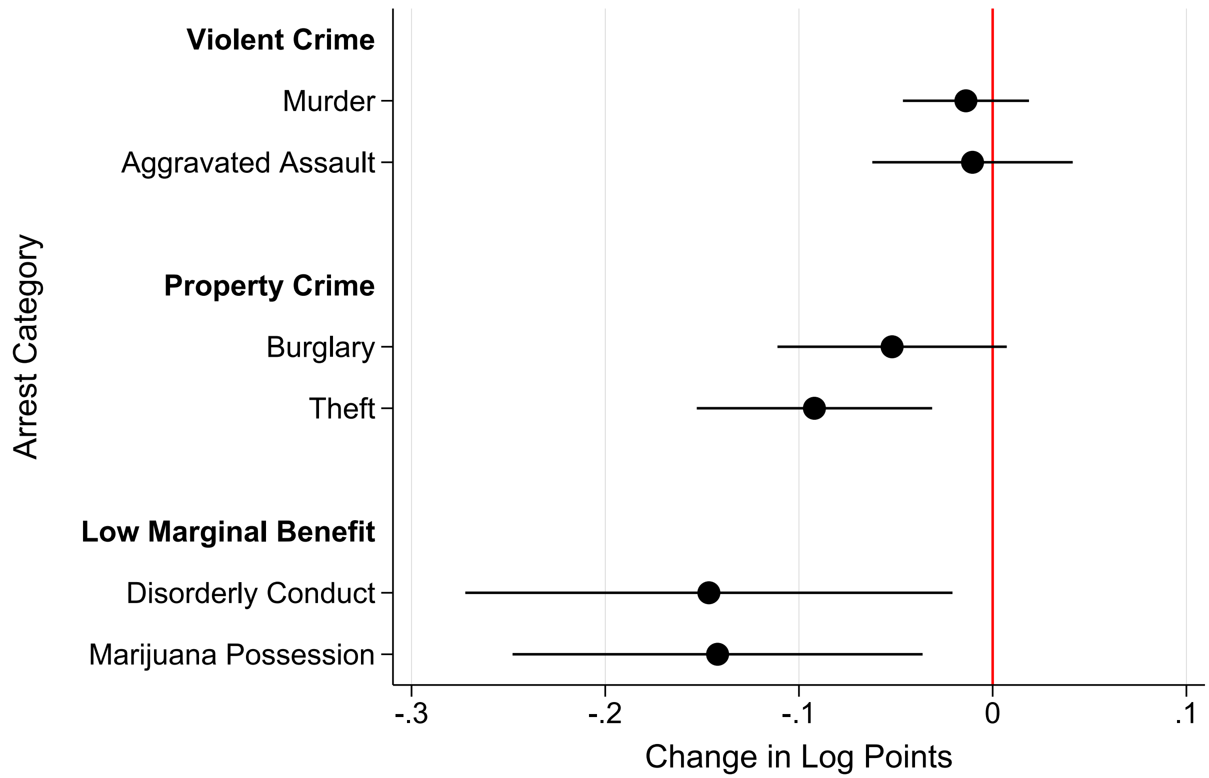


Figure 1.2: Map of High-Profile, Officer-Involved Fatalities (2005–2016)



The map illustrates the geographic variation in high-profile, officer-involved fatalities. The size of the dot corresponds to the number of fatalities in that city during 2005–2016. The map includes fatalities that will not be used in the main analysis sample, likely because the involved department poorly reported data to the FBI. Figure A.1 shows the distribution of timing of high-profile OFs.

Figure 1.3: Difference-in-differences (DD) Coefficient Estimates by Arrest Category

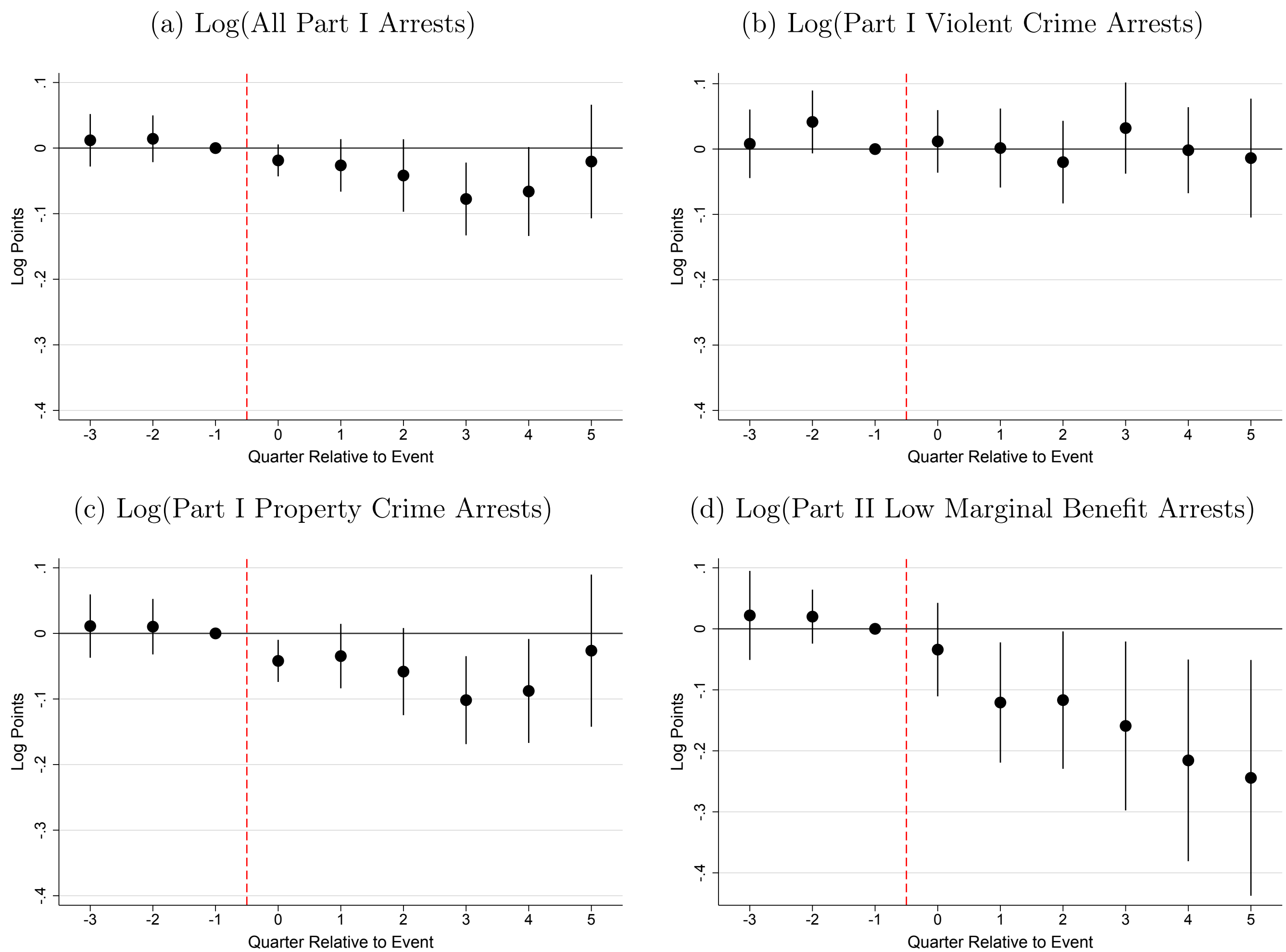


Circles display DD coefficients from separate regressions—in descending order of the social cost of crime—using a sample of city police departments with fewer than 9 outliers and a population greater 10,000. There are no significant differences between the pre-treatment trends of control and treated departments in any these regressions. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and violent and property crime, while Part I arrests control for their specific offense type. All regressions have month-of-sample and department fixed effects, as well as linear county trends. The actual DD tables begin with Table A.1. [740,980 observations; 53 treated; 2,686 control agencies]

Figure 1.4: Example Integrated Dummy Setup; Event window: -3 to 3 months around event with binned endpoints

Month Relative to Event	**-4+**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4+**	**Total**
Ferguson: 2005M1	2	0	0	0	.	.	.	.	.	2
Ferguson: 2005M2	2	0	0	0	0	.	.	.	.	2
Ferguson: 2005M3	2	0	0	0	0	0	.	.	.	2
Ferguson: 2005M4	2	0	0	0	0	0	0	.	.	2
Ferguson: 2005M5	2	0	0	0	0	0	0	0	.	2
Ferguson: 2005M6	2	0	0	0	0	0	0	0	0	2
Ferguson: 2005M7	2	0	0	0	0	0	0	0	0	2
Ferguson: 2005M8	2	0	0	0	0	0	0	0	0	2
Ferguson: 2005M9	2	0	0	0	0	0	0	0	0	2
Ferguson: 2005M10	1	1	0	0	0	0	0	0	0	2
Ferguson: 2005M11	1	0	1	0	0	0	0	0	0	2
Ferguson: 2005M12	1	0	0	1	0	0	0	0	0	2
Ferguson: 2006M1	1	0	0	0	1	0	0	0	0	2
Ferguson: 2006M2	0	1	0	0	0	1	0	0	0	2
Ferguson: 2006M3	0	0	1	0	0	0	1	0	0	2
Ferguson: 2006M4	0	0	0	1	0	0	0	1	0	2
Ferguson: 2006M5	0	0	0	0	1	0	0	0	1	2
Ferguson: 2006M6	.	0	0	0	0	1	0	0	1	2
Ferguson: 2006M7	.	.	0	0	0	0	1	0	1	2
Ferguson: 2006M8	.	.	.	0	0	0	0	1	1	2
Ferguson: 2006M9	.	.	.	.	0	0	0	0	2	2
Ferguson: 2006M10	.	.	.	.	.	0	0	0	2	2

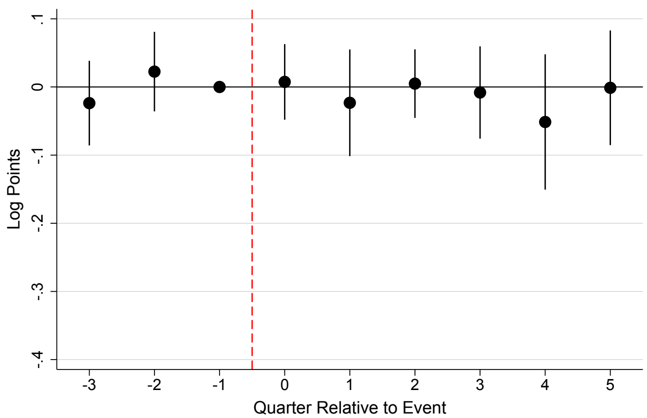
Figure 1.5: Effect of High-Profile, Officer-Involved Fatality on Arrests



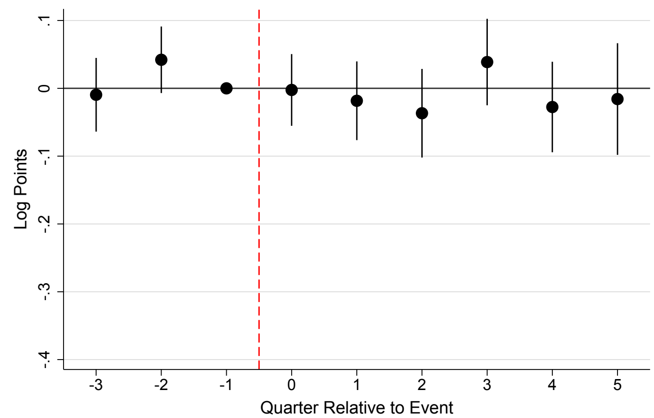
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines are the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, violent and property crime, and where possible, the specific offense. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Figure 1.6: Effect of High-Profile, Officer-Involved Fatality on Violent and Property Crime Arrests

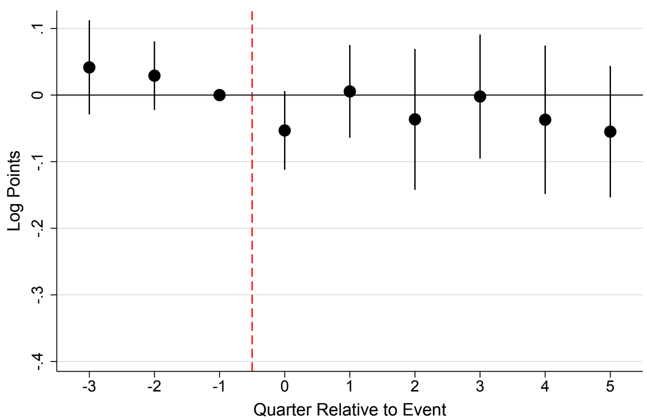
(a) Log(Murder Arrests)



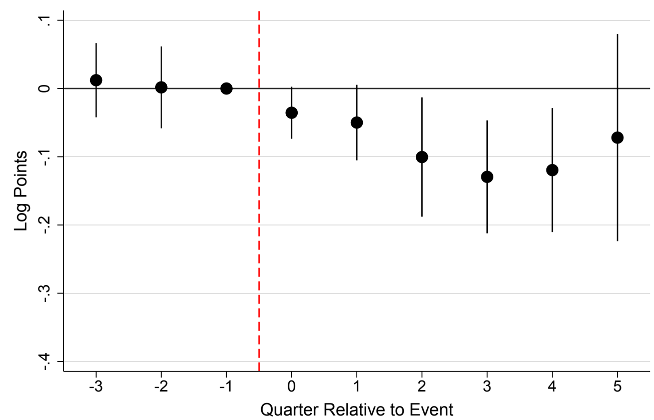
(b) Log(Aggravated Assault Arrests)



(c) Log(Burglary Arrests)



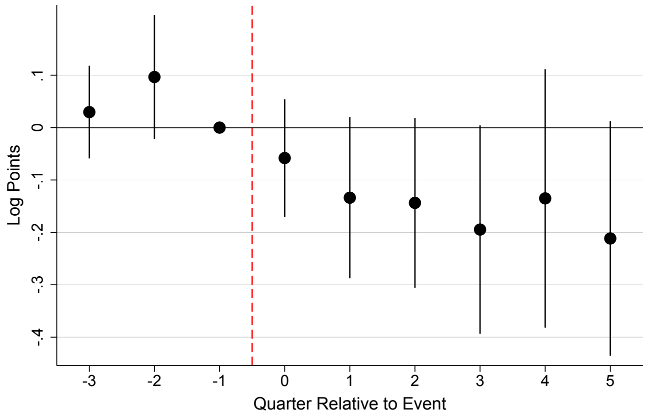
(d) Log(Theft Arrests)



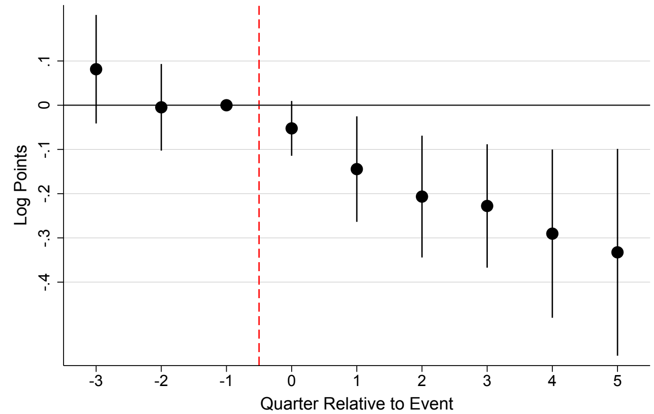
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, violent and property crime, and their specific offense type. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Figure 1.7: Effect of High-Profile, Officer-Involved Fatality on Low Marginal Benefit Arrests

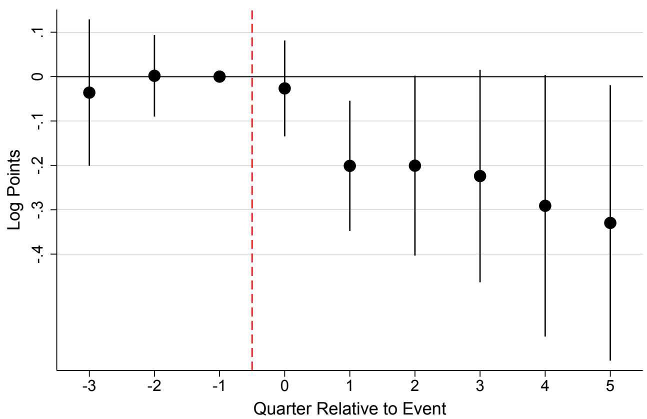
(a) Log(Disorderly Conduct Arrests)



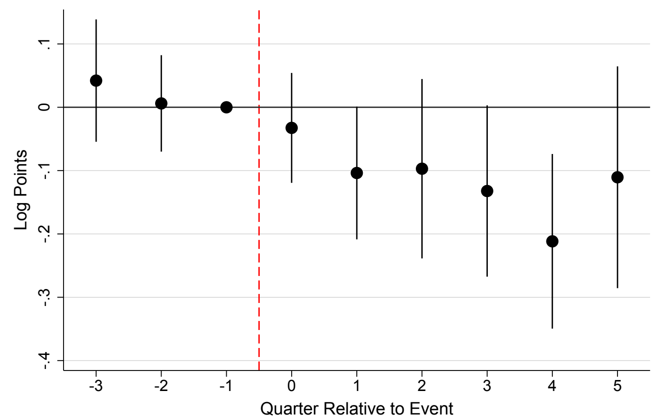
(b) Log(Marijuana Possession Arrests)



(c) Log(Liquor Law Violation Arrests)



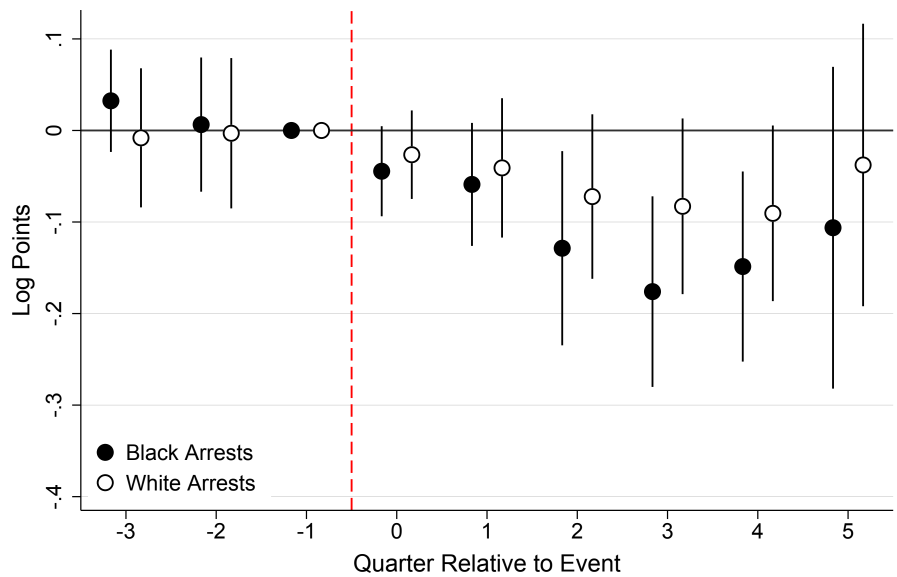
(d) Log(Marijuana Sale Arrests)



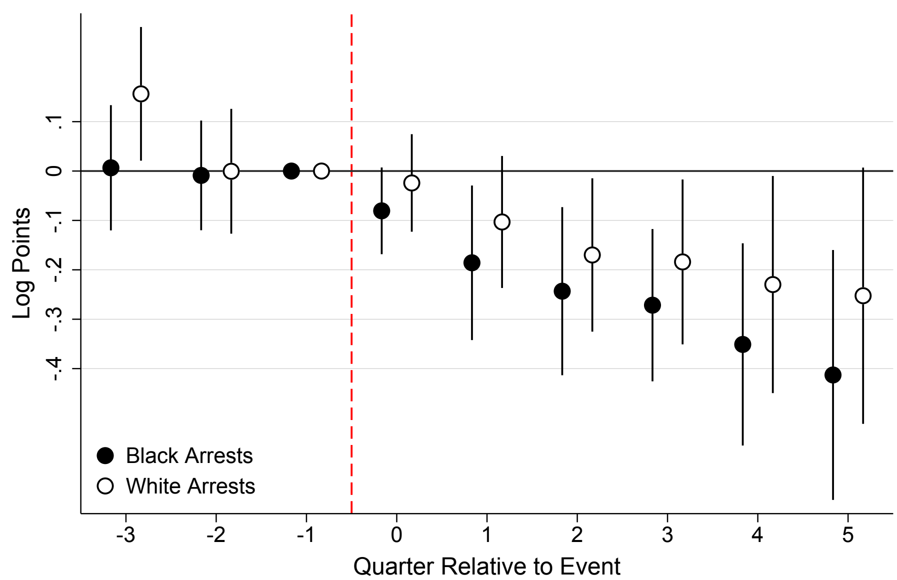
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and violent and property crime. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [599,088 observations; 53 treated; 2,687 control agencies]

Figure 1.8: Effect of High-Profile, Officer-Involved Fatality on Race-Specific Arrests

(a) Log(Theft Arrests)

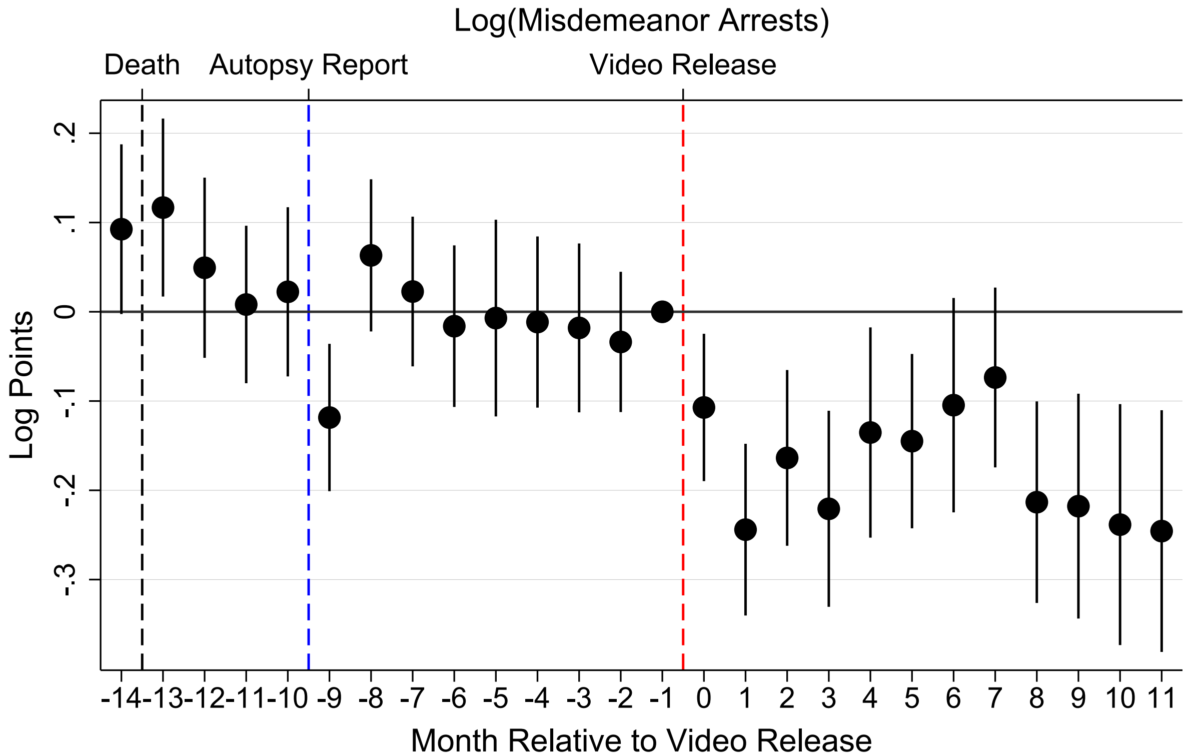


(b) Log(Marijuana Possession Arrests)



Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and violent and property crime, while Part I arrests control for their specific offense type. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Figure 1.9: Case Study of Laquan McDonald and the Chicago Police Department (2014–2017)



The dotted black line (left) depicts the month of death for Laquan McDonald (October 2014). The dotted blue line (middle) denotes when an article with the associated autopsy report was released (February 2015). The dotted red line (right) highlights when the video of his death was released to the public and the community became aware of the incident (November 2015). Circles display monthly event time coefficients for a regression of log misdemeanor arrests at the beat level in Chicago from 2014–2017. Lines represent the 95% confidence interval using standard errors clustered at the district level. I control for district-level fixed events and linear trends. [19,080 observations; 22 districts; 274 beats]

Tables

Table 1.1: Chalfin and McCrary (2018): Cost of Crime and Police

	Cost per Officer	Officers per 100K Population	Annual Cost per Capita
Sworn police	$130,000	262.7	$341

	Cost per Crime	Crimes per 100K Population	Annual Expected Cost per Capita
Murder	$7,000,000	9.9	$693
Aggravated Assault	$38,924	418.9	$163
Robbery	$12,624	286.4	$36
Motor vehicle theft	$5,786	454.3	$26
Burglary	$2,104	976.2	$21
Theft	$473	2,623.30	$12
		Grand Total:	$995
		Income per Capita:	$26,267

Table 1.2: Mean (S.D.) of Municipality Arrest, Characteristics, and Protest Data

	Pure Controls	Treated PDs	Overall
Arrest Rate (per 100,000 pop)			
White Violent Arrests	37.1	27.8	36.7
	(200.3)	(34.0)	(196.1)
Black Violent Arrests	50.9	72.7	51.9
	(111.2)	(62.5)	(109.6)
White Property Arrests	84.7	71.0	84.1
	(275.4)	(54.9)	(269.7)
Black Property Arrests	166.1	181.8	166.8
	(345.6)	(294.8)	(343.5)
White Low MB Arrests	188.2	184.7	188.1
	(489.6)	(168.8)	(480.2)
Black Low MB Arrests	246.8	269.6	247.8
	(651.7)	(252.5)	(639.6)
Local Characteristics (ACS 2014)			
Population	143,208.4	487,497.1	158,057.7
	(353,407.5)	(629,482.6)	(376,151.6)
% White	56.9	44.4	56.4
	(23.5)	(17.5)	(23.4)
% Black	11.9	23.3	12.4
	(14.9)	(19.9)	(15.3)
Poverty Rate	17.1	20.7	17.3
	(8.4)	(6.5)	(8.3)
White Poverty Rate	12.4	12.5	12.4
	(7.2)	(5.8)	(7.1)
Black Poverty Rate	26.1	28.8	26.3
	(13.9)	(8.4)	(13.7)
% Bachelor's	32.2	32.4	32.2
	(14.8)	(11.9)	(14.7)
Square Miles	55.0	124.0	58.0
	(134.3)	(128.9)	(134.8)
Population Density	3,715.7	4,340.5	3,742.8
	(3,813.3)	(2,948.8)	(3,782.1)
BLM Protests (Aug. 2014–Aug. 2015)			
Number of BLM Protests	0.78	4.74	0.95
	(2.83)	(6.98)	(3.23)
Attendance	210.62	1,237.47	254.91
	(2,093.86)	(2,496.65)	(2,123.08)
% Protesting	21.24	66.01	23.17
	(40.90)	(47.37)	(42.19)
Observations	162,132	7,308	169,440

Pure control departments have had no high-profile OFs. Treated PDs have had at least one OF. There are 105 high-profile OFs with a maximum of 6 in one jurisdiction, Los Angeles.

Table 1.3: Fatalities by Victim & Incident Characteristics with Mean News Articles

	Share	**Mean News Articles**	**Number**
Gender			
Female	14%	2,234	10
Male	86%	17,793	63
Race/Ethnicity			
Asian	1%	1,610	1
Black	73%	20,575	53
Hispanic	11%	2,740	8
Native American	4%	1,657	3
White	11%	3,034	8
Cause of Death			
Asphyxiation	3%	2,160	2
Beaten	7%	52,442	5
Gunshot	90%	13,284	66
Armed?			
Allegedly Armed	40%	4,230	21
Unarmed	60%	27,309	32
Mentally Impaired?			
No	75%	22,485	44
Yes, Alcohol/Drug Use	3%	4,270	2
Yes, Mental Illness	22%	8,012	13

Source: Fatal Encounters (2005–2016); 73 observations; ASR Sample

Chapter 2

Evidence on Policing Effectiveness[1]

2.1 Introduction

This chapter discusses the essential role that the police have in deterring and reducing crimes, particularly the most violent and costly ones to society, such as murder. We begin by providing a brief overview of deterrence theory before discussing the empirical evidence on the efficacy of police staffing and various policing strategies on crime reduction. Using a framework developed in Weisburd and Eck (2004), we quickly evaluate the model of standard policing and then mainly focus on evidence behind three current policing practices: hot spots, problem-oriented, and proactive. Finally, we use the empirical evidence of police staffing to provide a basis for a theoretical model on the optimal level of policing. Using the Chalfin and McCrary (2018) framework, we discuss how one could estimate how much crime could be reduced if additional funds were directed to hire more law enforcement officers, and if crime-reduction were the sole policing objective, how many cities are in fact underpoliced. We conclude by postulating whether we could implement additional policing without resulting in unwarranted and excessive social costs for the community as discussed by Manski and Nagin (2017).

2.2 Theories of Deterrence

Deterrence has been articulated in academic writing at least as far back as eighteenth-century treatises by Adam Smith (1776), Jeremy Bentham (1789), and Cesare Beccaria (1764). There are three core concepts embedded in theories of deterrence: individuals respond to changes in the certainty, severity, and celerity (or immediacy) of punishment. Criminology bifurcates deterrence into either general or specific, with general deterrence referring to the idea that individuals respond to the *threat* of punishment and specific deterrence referring to the idea that individuals are responsive to the actual experience of punishment.

---

[1] This chapter is coauthored with Justin McCrary. The ungated, published version can be found online here, while the formal version can be found here. We thank Aaron Chalfin, Eric Miller, and Tamara Lave for their support with the chapter content, suggestions, and editing assistance. We also appreciate the discussion of the chapter at the Law and Society Association.

Economics prefers different terminology, reserving the term deterrence for what the criminologist calls general deterrence (i.e., any changes in the three core concepts outlined above) and then describing specific deterrence as a change in information or a change in preferences themselves.[2] In the next subsection, we briefly characterize the way economists have formalized these concepts. In general, economic theories of deterrence have focused more heavily on certainty and severity. However, recent writing has increasingly focused on celerity and characterized deterrence as part of a dynamic framework in which offender behavior is sensitive to their time preferences (Polinsky and Shavell, 1999; Lee and McCrary, 2017). Naturally, the policy findings of these models are implicated by the discount rate chosen for a given population—or how much potential offenders "value the future."

2.2.1 Economic Models of Crime

The earliest formal model of criminal offending in economics is Becker's seminal 1968 paper, "Crime and Punishment: An Economic Approach." The crux of Becker's model is the idea that a rational potential offender faces a gamble. He can either choose to commit a crime and thus receive a criminal benefit (albeit with an associated risk of apprehension and subsequent punishment) <u>or</u> not to commit a crime (which yields no criminal benefit but is risk free). In Becker's static (i.e., time invariant) model, the expected cost of committing a crime is a function of the offender's probability of apprehension, and the severity of the sanction that he will face upon apprehension.

To be more specific, the individual can face three potential outcomes, each of which delivers a different level of utility: (1) the utility associated with the choice to abstain from crime; (2a) the utility associated with choosing to commit a crime that does not result in an apprehension; and (2b) the utility associated with choosing to commit a crime that results in apprehension and punishment. In this framework, the individual chooses to commit a crime if, and only if, their expected utility exceeds the utility from abstention. Or, put more plainly, the rational potential offender commits the crime if the *expected* or *perceived* benefits outweigh the *expected* costs.

2.3 Police and Crime

Becker's prediction that the aggregate supply of crime will be sensitive to society's investment in police arises from the idea that an increase in police presence, whether it is operationalized through increased staffing or more visible patrols, raises the probability that an individual is apprehended for having committed a particular offense. To the extent that potential offenders are able to observe an increase in police resources and perceive a correspondingly higher risk to criminal participation, crime is expected to decline through the deterrence channel. Increased investment can also lead to reductions in crime through an

[2] Since we are both economists, we will stick to the framework and language utilized in much of the economics of crime literature. However, where they are significant deviations in terminology or perspectives, we will do our best to add the appropriate caveats.

incapacitation effect: rises in arrest rates result in a mechanical reduction in crime since pretrial detention and incarceration limit additional violations from detained offenders. Much of the literature prioritizes and focuses on deterrence since it is the preferred channel of crime reduction, however disentangling which effect is empirically salient in a study is still a significant hurdle. Unlike incapacitation, deterrence is the optimal mechanism for crime reduction since there are no direct costs on a victim from a crime or an offender through lost compensation from incarceration, there are no societal costs in terms of funding a system of jails and prisons, and there is no judicial congestion from going through a lengthy legal process.

A large literature has used city- or state-level panel data analysis and, recently, a variety of quasi-experimental designs to estimate the elasticity of crime with respect to police force size (hereafter, police staffing). (An elasticity is the percent change of one variable, such as crime, in response to a single percent change in another variable—for example, hiring of additional police officers.) This literature is ably summarized by Cameron (1988), Nagin (1998), Eck and Maguire (2000), Skogan and Frydl (2004), Levitt and Miles (2006), and Chalfin and McCrary (2017) all of whom provide extensive references.

The early panel-data literature tended to report small elasticity estimates that were rarely distinguishable from zero and sometimes even positive, suggesting perversely that police increase crime. The ensuing discussion in the literature was whether police reduce crime at all. Beginning with Levitt (1997), an emerging quasi-experimental literature has argued that simultaneity bias is the culprit for the small elasticities in the panel-data literature. (Simultaneity bias occurs when original independent variable—police staffing—is in fact partially determined by the dependent variable—crime. The bias on the regression coefficient is determined by the underlying relationship in the variables, but it suggests the coefficient should be strongly questioned.) The specific concern articulated is that if police are hired in anticipation of an upswing in crime, then there will be a positive bias associated with regression-based strategies, masking a true negative elasticity. Put differently, rather than estimating the true effect of police staffing on crime, elasticities from simple panel-data methods are confounded by partially capturing proactive public investment in the police force that is positively correlated to the crime rate.

To circumvent the bias that plagues the estimates, researchers needed to discover an exogenous change in police staffing that is uncorrelated with the local crime rate. Levitt (1997) was the first to answer this question by exploiting plausible exogenous variation, using an Instrumental Variables (IV) estimation strategy. Leveraging data on the timing of mayoral and gubernatorial elections, Levitt provides evidence that in the year prior to a municipal or state election, police staffing tends to increase, presumably due to the desire of elected officials to appear to be "tough on crime." The necessary condition for feasibly estimating the true effect is that crime does not vary cyclically with respect to the election cycle, aside from the impact of increases in police staffing.

Using data from fifty-seven cities spanning 1972–97 and the historical panel-data estimation, he finds that the effect of police on crime is small, consistent with the prior literature. However, the IV estimates were substantially larger, but not statistically significant, driven by the fact that the relationship between election cycles and police hiring is weak, complicating both estimation and inference.

The underlying lesson in Levitt (1997) regarding simultaneity bias has catalyzed a series of papers that seek to identify a national effect of police staffing on crime by isolating exogenous within-city variation in police staffing levels that is uncorrelated to local crime by using an instrumental variable. These papers include Levitt (2002), who uses variation in firefighter numbers as an instrument variable for police staffing; Evans and Owens (2007), who use the size of federal Community Oriented Policing Services (COPS) grants awarded to cities to promote police hiring as an instrument; and Lin (2009), who uses the differential exposure of US states to exchange-rate shocks depending on the export intensity of local industry. These strategies consistently demonstrate that police do, in fact, reduce crime.

The estimated elasticities, however, display a wide range, roughly -0.1 to -2, depending on the study and the type of crime. Moreover, relatively few of the estimated elasticities are significant at conventional levels of confidence, reflecting a great deal of sampling variability and the use of relatively weak IVs. In many cases, extremely large elasticities (i.e., those larger than one in magnitude) cannot be differentiated from zero. Overall, Chalfin and McCrary (2017) characterize the pattern of the cross-crime elasticities as, in general, favoring a larger effect of police on violent crimes than on property crimes, with especially large effects of police on murder, robbery, and motor vehicle theft. Additional discrepancies between panel-data and IV estimates—and the lack of significance in the IV—could possibly also be explained by measurement error in either police staffing or index crimes from the FBI's Uniform Crime Report. Attempting to quantify the degree to which poor data could bias the police staffing estimates is dealt with in detail in Chalfin and McCrary (2018).

While the majority of the police staffing literature uses aggregate data, there is a corresponding literature that assesses the impact of police on crime using natural experiments in a particular jurisdiction. An early account of such a natural experiment is found in Andenaes (1974), who documents a large increase in crime in Nazi-occupied Denmark after German soldiers dissolved the entire Danish police force (Durlauf and Nagin, 2011; Nagin, 2013). Modern literature has found similarly large effects. In particular, DeAngelo and Hansen (2014) document an increase in traffic fatalities that occurred in the aftermath of a budget cut in Oregon that resulted in a mass layoff of state troopers. Similarly, Shi (2009) reports an increase in crime in Cincinnati, OH, in the aftermath of an incident in which police used deadly force against an unarmed African American teenager and subsequently reduced their arrest rates. The literature using natural experiments as a lens to cleanly identify the effects of police staffing or deployment has grown substantially. There are obvious limitations in the interpretations of these events and extrapolations on policing more broadly. We will have a more extended discussion of this is in subsection 2.3.2.

2.3.1 Police Deployment and Tactics

The police staffing literature is informative with respect to the aggregate response of crime to increases in police staffing. However, much of that literature leaves the underlying variation in how police reduce crime unanswered. In particular, to what extent do the estimated elasticities reflect deterrence? Moreover, what is the specific mechanism that leads to deterrence? If the mechanism is based on general deterrence—the idea that offenders perceive an increase in police presence and adjust their behavior from the elevated risk of apprehension—then it should be the case that offending is especially sensitive to large and easily observed changes in police deployment and tactics.

To address these questions, a related literature that is found mostly in criminology has studied the impact of various policing strategies on crime. In particular, declines in crime that are not attributable to spatial displacement have been linked to the adoption of **"hot spots" policing** (Sherman and Rogan, 1995; Sherman and Weisburd, 1995; Braga, 2001; Braga 2005; Weisburd, 2005; Braga and Bond, 2008; Berk and MacDonald, 2010), **"problem-oriented" policing** (Braga et al., 1999; Braga et al., 2001; Weisburd and Eck, 2004; Weisburd et al., 2010; Chalfin and McCrary, 2017), and a variety of other **proactive approaches**. We expound on each of these strategies in the sections below.

Similarly, a large research literature that has examined the local impact of police crackdowns has consistently found large and immediate (but typically not lasting) reductions in crime in the aftermath of hyper-intensive policing (Sherman 1990). Such findings are further supported by evidence from several informative natural experiments that have identified plausibly exogenous variation in the intensity of policing. Three prominent examples are Klick and Tabarrok (2005), who study the effect of police redeployments in Washington, D.C. that result from shifts in terror alert levels; Di Tella and Schargrodsky (2004), who study the effect of a shift in the intensity of policing in certain areas of Buenos Aires after the 1994 Jewish Community Center bombing; and Draca, Machin, and Witt (2011), who study police redeployments in the aftermath of the 2005 London tube bombings.

To analyze the diverse set of policing strategies, we employ the framework of Weisburd and Eck (2004), describing each strategy through its *level of focus* on place or offender types and its *diversity of interventions* utilized to diminish crime and disorder. With this model in mind, we will contrast standard policing by reviewing three contemporary methods and the evidence of their efficacy.[3] The recent literature on police deployments and tactics has focused predominantly on three types of interventions, with the caveat that in practice boundaries between policing strategies are often blurred and undefined.

The first is an innovation commonly referred to as "hot-spots" policing, which coincided with the propagation of digital mapping technology. As the moniker suggests, hot-spots policing describes a strategy in which police are disproportionately deployed to areas in a city that appear to attract disproportionate levels of crime.

The second type of intervention is often referred to as "problem-oriented" policing (POP). This term is used broadly and refers to a collection of problem solving steps and focused deterrence strategies that are designed to change the behavior of specific types of offenders or to be successful within specific jurisdictions.

A final intervention that has received attention in the literature is that of "proactive" policing. Proactive policing refers to strategies that are designed to make policing more intensive, holding resources fixed. The implementation can be quite varied, but it can be traced back to the concept of "broken windows" or disorder policing introduced by Wilson

[3]Places to review crime or disorder reduction strategies and the actual programs where they have been implemented are the National Institute of Justice, who hosts a repository of evidence at `www.crimesolutions.gov`, and Lum, Koper, and Telep (2011), who keep an updated visual matrix of numerous interventions at George Mason University's Center for Evidence-Based Crime Policy at `http://cebcp.org/evidence-based-policing/the-matrix/`.

and Kelling (1982). The strategy refers to the notion that, just like fixing a broken window sends a message to would-be vandals that the community cares about maintaining social order, arresting individuals for relatively minor infractions sends a message to potential offenders that the police are watchful.

2.3.1.1 Standard Model of Policing

The standard model of policing, one that was used almost exclusively from the 1970s to the 1990s in the US, is characterized by broad strategies to reduce crime and disorder. Weisburd and Eck (2004) narrow them down to five main concepts: "(1) increasing size of police agencies; (2) random patrol across all parts of the community; (3) rapid response to calls for service; (4) generalized investigations of crime; and (5) generally applied intensive enforcement and arrest policies." The standard model's rigid and broad approach to policing results in low scores for focus and diversity of interventions.

Typified by the low scores, this reactive and incident driven strategy has no evidence to bear witness to its efficacy (Weisburd and Eck, 2004). The standard model often uses a one size fits all approach resulting in success metrics such as the number of the patrol cars on the street or response time to citizen. Although these ostensibly could be useful proxies, it is much more important for police departments to understand the fundamental dynamics that are shaping crime and disorder in their jurisdiction. In the evaluations that have been conducted, rapid response time and random preventive patrols have been ineffectual in their reduction in actual crime or citizens' fear of crime (Kelling et al., 1974; Spelman and Brown, 1981). The crime wave which peaked in the 1990s and the lack of police tools to address it led to the innovation in strategies that will be discussed in the next sections.

2.3.1.2 Hot-Spots Policing

We begin with a discussion of hot-spots policing, which we distinguish from aggregate police staffing research for several reasons. First, since the staffing literature largely uses city-level variation and time dynamics, it identifies the effect of adding police at the expense of some other type of public input or additional revenue stream. In contrast, hot-spots policing involves a reallocation of existing resources across neighborhoods within a city. Such a strategy is advantageous, as it does not require a change in current outlays. However, it leaves open the possibility that moving police around merely displaces, rather than reduces, crime.

In order for hot-spots policing to be a viable crime reduction strategy, two conditions must be met. First, given resource constraints, the feasibility of such a deployment strategy relies on crime being sufficiently concentrated in a relatively small number of hot spots. Second, hot spots must be sufficiently stable such that the spatial distribution of crime in the absence of a change in police deployment can be predicted with a reasonable degree of accuracy. Hence, the adoption of hot-spots policing must begin with an accounting of the geographic concentration of crime, as well as an assessment of the extent to which hot spots are permanent as opposed to transitory. Sherman (1995) captures both of these ideas, characterizing crime hot spots as "small places in which the occurrence of crime is so frequent that it is highly predictable, at least over a one-year period." These concepts are predicated on the theory that either certain city locations are criminogenic (i.e., there are environmental factors that make these areas more conducive to crime committing) or that there are positive spatial spillovers to crime such that geographic concentration of crime allows for what economists describe as agglomeration effects.

A seminal paper by Sherman, Gartin, and Buerger (1989) is the first to provide descriptive data on the degree to which crime is spatially concentrated. Using data from Minneapolis, Sherman and coauthors found that just 3 percent of addresses and inter- sections in Minneapolis produced 50 percent of all calls for service to the police. This finding is echoed by Weisburd, Maher, and Sherman (1992) and in more recent papers by Weisburd et al. (2004) and Weisburd, Morris, and Groff (2009), which report that a very small percentage of street segments in Seattle accounted for 50 percent of crime incidents for each year over a fourteen-year period. With respect to predictability, Weisburd et al. (2004), using the same data from Seattle, used trajectory analysis to establish that hot spots tended to be highly persistent, often for many years.

Naturally, the observation that crime is so highly concentrated in a very small number of places has led to efforts to intensify the focus of police resources on these crime corridors to decrease aggregate crime. Primarily, the hot spots methodology uses more traditional policing tactics in a highly focused, low diversity manner. The National Institute of Justice (2017) provides an extensive list of common hot-spots policing tactics, such as "police raids on suspected drug houses, police crackdowns in high-crime areas, bust-buy operations, street sweeps, saturated patrol, curfew and truancy enforcement, and warrant servicing." These treatments have, in turn, led to a corresponding experimental and quasi-experimental research literature that seeks to evaluate the efficacy of such strategies.

The first-order question that the hot-spots policing literature seeks to address involves the degree to which crime is responsive to a highly localized change in the intensity of policing. Criminologists generally define 'responsive' as crime that declines in local areas after being exposed to more intensive police patrols without inducing equivalent spillovers to untreated adjacent areas. However, while spillovers undermine the viability of hot-spots policing as a crime-reduction strategy, they nevertheless constitute evidence of responsiveness and, as such, are useful in identifying deterrence. Moreover, a particular feature of this research makes it especially relevant for determining the mechanism for crime reduction (Nagin, 2013).

The first test of policing crime hot spots may be found in a 1995 randomized experiment conducted by Sherman and Weisburd in Minneapolis. The experiment tested whether doubling the intensity of police patrols in crime hot spots resulted in a decrease in crime and found that crime declined by approximately 10 percent in experimental places relative to control places. No evidence of crime displacement—that is, spillovers—was found. Findings in Sherman and Weisburd (1995) have, to a large extent, been replicated in other places and contexts including the presence of open-air drug markets and "crack houses" (Hope, 1994; Weisburd and Green, 1995; Sherman et al., 1995), violent crime hot spots (Sherman and Rogan, 1995; Braga et al., 1999; Caeti, 1999), and places associated with substantial social disorder (Braga and Bond, 2008; Berk and MacDonald, 2010). Indeed, a review of the literature by Braga (2001) identified nine experiments or quasi-experiments involving hot-spots policing and noted that seven of the nine studies, including a majority of the randomized experiments, found evidence of significant and large crime reductions. Notably, a majority of the literature finds no evidence of crime displacement to adjacent neighborhoods (Weisburd et al., 2006), while a number of studies have found that the opposite is true—that there tends to be a diffusion of benefits to untreated adjacent places (Sherman and Rogan, 1995; Braga et al., 1999; Caeti, 1999). Both of these findings are perfectly consistent with our conceptualization of deterrence.

2.3.1.3 Problem-Oriented Policing

Intensive policing of hot spots is one way that police potentially deter crime. Another strategy is the multifaceted and deterrence-based problem-oriented policing (POP). Broadly speaking, this strategy entails engaging with community residents to identify the most salient local crime problems, and then designing strategies to deter unwanted behavior. The specifics are highly variable and are intended to leverage local resources to address local concerns. Hence, we view POP as a high focus and high diversity police approach, whose multi-pronged implementation is tailored to each jurisdiction. Despite the empirical evidence suggesting efficacy in crime reduction, POP studies struggle with a standardization of intervention, an extrapolation of effect, a knowledge of mechanisms that instigated the effect, and determining the appropriate control groups since there is endogeneity in adopting such a networked approach.

There is a consistent methodology used in POP developed by Goldstein (1979) and Eck and Spelman (1987) that provides commonality across implementation. The prototypical POP methodology is the SARA Model: **Scanning** to identify systematic issues that are of concern to the community and may provoke disorder and crime; **Analysis**: What does the data tell us about this concern? Is there a specific underlying issue? What are the proximal correlates that could be causing it?; **Response**: What are the appropriate interventions based on the findings of the concern and what could be causing it? Who are the appropriate partners to execute these interventions to implement the desired changes within the community?; **Assessment**: Did we respond accordingly and how effective was our response? This step likely includes a process and impact evaluation.

How the SARA model is integrated in the POP is heterogeneous across cities, but there is a reliable emphasis on research and analysis as well as crime prevention and the engagement of public and private organizations in the reduction of community problems. For example, it would be common for the police department to partner with housing and building inspectors, local businesses, and community members to support crime prevention through building inspection crackdowns, and improved tertiary services like street cleaners or street lights in high-crime areas.[4]

[4] We contrast these examples with prototypical community or third-party policing, which has a high diversity of techniques but a low focus in scope. Community policing similarly involves partnerships with the police, community, and other service providers, but the focus is on the general neighborhood rather than a few problematic issues. The motive is to improve police-community relations, build neighborhood cohesion, increase trust in police, and reduce legal cynicism. Additionally, community policing stresses preventative

One component of POP is frequently referred to as "focused deterrence" strategies, where law enforcement deter through directed and personalized information campaigns (Zimring and Hawkins, 1973). The idea is to create deterrence by making potential offenders explicitly aware of the risks of serious criminal involvement and increasing the perceived risk of apprehension. Often, the information is specifically tailored to the potential offender, increasing its poignancy.

Undoubtedly, the most well-known evaluation of a problem-oriented policing approach is that of Boston's Operation Ceasefire by Kennedy et al. (2001). The stated purpose of Ceasefire was to reduce youth gun violence in Boston. The intervention involved a multifaceted approach and included efforts to disrupt the supply of illegal weapons to Massachusetts. It also included the police communicating directly to gang members that authorities would use every available "lever" to punish gangs collectively for violent acts committed by individual gang members. In particular, police indicated that the stringency of drug enforcement would hinge on the degree to which gangs used violence to settle business disputes. The result of the intervention was that youth violence fell considerably in Boston relative to other US cities included in the study. Indeed, the perception of Ceasefire has been overwhelmingly positive and accordingly it has given rise to a number of similarly motivated strategies that are collectively referred to as "pulling levers."

Prominent evaluations of pulling-levers interventions include research carried out in Richmond, VA (Raphael and Ludwig, 2003), Indianapolis (McGarrell et al., 2006), Chicago (Papachristos, Meares, and Fagan, 2007), Stockton, CA (Braga, 2008b), Lowell, MA (Braga et al., 2008), High Point, NC (Corsaro et al., 2012), Nashville (Corsaro and McGarrell, 2010), Cincinnati (Engel, Tillyer, and Corsaro, 2010), and Rockford, IL (Corsaro, Brunson, and McGarrell, 2010). Researchers have also evaluated a multi-city pulling-levers strategy known as Project Safe Neighborhoods (PSN), which enlisted the cooperation of federal prosecutors to crack down on gun violence. A 2012 review of the literature by Braga and Weisburd suggests that pulling-levers strategies have been effective in reducing serious violent crime, with all reviewed studies finding negative point estimates, the majority of which were significant. With respect to individual evaluations, reductions in crime have been found in High Point, Chicago, Indianapolis, Stockton, Lowell, Nashville, and Rockford and null findings have been found in Richmond and Cincinnati. With respect to Project Safe Neighborhoods, the research is promising but certainly not definitive. McGarrell et al. (2010) report that declines in crime were greater in PSN cities than in non-PSN cities. However, there is a great deal of heterogeneity among cities, making it difficult to draw clear inferences.

On the whole, evaluations of pulling-levers strategies produce promising results, though inference is invariably complicated by a lack of randomized experiments and the inherent difficulty in identifying appropriate comparison cities. Identification problems are additionally compounded by the difficulty in identifying mechanisms, as each pulling-levers strategy is complex, multifaceted, and situation dependent, often involving changes in both the intensity of law enforcement as well as sentencing (e.g., Project Exile in Richmond, VA, as well as Project Safe Neighborhoods). Accordingly, it is easy to imagine that as additional

---

and de-escalation measures for police, while POP can be highly confrontational for the subset it is attempting to deter.

resources are brought to bear, some of the effects of pulling-levers strategies might accrue via incapacitation effects. Concerns regarding identification have led Skogan and Frydl (2004) to conclude that such research is "descriptive rather than evaluative." Given the relatively large effect sizes reported in the literature, our view of these papers is more optimistic than that of Skogan and Frydl. However, caution is warranted in characterizing this literature as having detected unassailable evidence of deterrence.

2.3.1.4 Procedural Justice

There is literature on legal compliance that well summarized by Nagin and Telep (2017), who provide a comprehensive overview of the literature covering both the theory and what evidence exists to date on the question of whether interventions that increase citizens' perceptions of procedural justice and police legitimacy through officer trainings or community outreach can significantly raise legal compliance. These questions may be especially poignant now given the national spotlight on racial discrepancies in policing, specifically in use of force and officer-involved fatalities.

These racial disparities in policing, which cannot be explained by controlling for crime or geolocation, end up undermining community trust in the legal justice system. These discrepancies are reflected in the racial gaps of trust in police: A Gallup poll in 2016 illustrated that 67% of blacks believed that the police treat blacks less fairly than whites, while 16% report having personally unfair interactions with them (Gallup, Inc., 2016). In turn, the disconnect from the community results in worse policing as it lowers crime reporting and introduces reticence for those who may testify as witnesses (Desmond et al., 2016). The discrimination felt by black persons atrophies police legitimacy and induces legal cynicism, rendering many policing strategies ineffective, which Kirk and Papachristos (2011) argue leads to persistence of neighborhood violence despite decreases in poverty and in crime elsewhere (Sunshine and Tyler, 2003).

Nagin and Telep (2017) share studies that suggest perceptions of procedurally just treatment are closely correlated with perceptions on police legitimacy, which subsequently is strongly associated with legal compliance. As Nagin and Telep (2017) rightly caution, these results should be taken as preliminary evidence of a nascent empirical foundation. Many of these results are difficult to quantify given their subjective ratings on perception. Further, for many of these studies, it is not feasible to disentangle a cause-and-effect narrative given their empirical designs, particularly since this theory is riddled with the potential of reverse causality. Put differently, one could feasibly interpret police-citizen interactions in a game theory sense: police are administering differing levels of procedural justice based on citizen compliance, which partially shapes how citizens will respond and perceive police orders.

However, there have been a few studies analyzing natural experiments which have plausibly exogenous shifts in police legitimacy, allowing us to study the effects on legal compliance. An important caveat should be that researchers are not able to experimentally manipulate perception of procedural justice or police legitimacy, which means they are proxied by policies, trainings, or community outreach that is aimed at increasing both. Owens et al. (2016) is one of the few studies which examine how randomly-assigned trainings with their supervisor on topics pertaining to procedural justice affect officer behavior in Seattle. They found that officers that took part in the training were less likely to settle disputes through arrests or force. Importantly, these results were not driven by reductions in citizen interactions (i.e., there was no evidence to suggest de-policing).

Understanding that trainings are one way to elicit behavior change from officers, we can also shape perceptions of the community through outreach. A handful of these programs which work to increase police legitimacy programs have been shown to decrease aggregate crime on the neighborhood level (Papachristos et al., 2007). However, Wallace et al. (2016) seems to be the first quasi-experimental study to illustrate that programs aimed to influence police legitimacy impact legal compliance **at the individual level**.

Wallace et al. study the case of Project Safe Neighborhoods in Chicago, a program modeled after the POP "pulling levers" approach utilized in Operation Ceasefire in Boston. One of the prongs of this multifaceted program was an hour-long Offender Notification Forum, where violent felons who were recently released from prison were informed that any additional violent behavior would not be tolerated. The forum also focused on respect, fairness, and individual choice. The former intervention was a focused deterrence package and the latter underlined concepts of procedural justice and police legitimacy.

Wallace et al. use hazard models to determine the differential likelihood of recidivism and time until re-incarceration for this high-risk population using two comparison districts and two treatment districts. Possessing individual level data, they were able to distinguish the effect within a district (i.e., comparing criminal behavior between those who did and did not attend the forum in the treated district) and between districts (i.e., those in the treated district versus those in the non-treated district). Given that authors were able to discover significant differences in terms of recidivism rates and time to recidivate between districts and especially within districts provides compelling evidence for the effectiveness of the forum. However, it is quite difficult to attribute what impact came from the legitimacy portion of the intervention versus the focused deterrence and social service treatments. Only if we first rely on the untested theoretical premise that the forum bolstered notions of procedural justice and police legitimacy may we then conclude that these legitimacy-based interventions can be an effective means of increasing legal compliance for violent populations who are at a high risk of re-offending.

Combining the preferred elements of hot-spots and community policing, POP uses sophisticated resource-targeting strategies of hot-spots policing with intense community engagement to tackle the most persistent problems that engender crime and disorder.

2.3.1.5 Proactive Policing

The last discussed police tactic is proactive policing, particularly relevant for police departments who are resource constrained but have the capacity to increase the policing intensity. As there is no standardized way to assess the extent to which individual police departments engage in police work that is proactive, in practice, this literature seeks to understand if the increased intensity of arrests for minor infractions has an effect on the incidence of more serious crimes. Building on a proposition in Wilson and Kelling (1982), such an empirical operationalization was first proposed by Sampson and Cohen (1988) and has been replicated to various degrees by MacDonald (2002) and Kubrin et al. (2010). The general strategy for evaluating this high-focus, low diversity tactic is to run a regression with crime rates on a measure of policing intensity. In practice, policing intensity has been operationalized using the number of driving under the influence (DUI) and disorderly conduct arrests made per police officer. Using this approach has, in some cases, led to findings that are consistent with a deterrence effect of proactive policing. However, in the best controlled models, coefficients on the proactive policing metrics become small and insignificant. More importantly, these models are plagued by problems of simultaneity bias, omitted variables, and the inevitable difficulty involved in finding a credible proxy for the concept of proactive policing, as opposed to simply an environment that is rich in opportunities for police officers to make arrests. These issues coalesce, making it difficult to make any credible causal statements about the proactive policing tactic.

The idea of increasing police intensity can be traced back to the concept of "broken windows," or disorder policing, introduced by Wilson and Kelling (1982). It refers to the notion that, just like fixing a broken window sends a message to would-be vandals that the community cares about maintaining social order, arresting individuals for relatively minor infractions— such as vandalism and turnstile jumping—sends a message to potential offenders that the police are watchful. Broken-windows policing, in theory, operates primarily through perceptual deterrence—if offenders observe that police are especially watchful for minor infractions, they may update their perceived probability of apprehension for a more serious crime and accordingly will decrease their participation in crime. With less vandalism and visible low level offenses present, the criminogenic effect of these areas is heavily diminished, purveying a signal of order and police presence to other potential offenders or more serious criminals. In the popular media, broken-windows policing is an idea that is heavily associated with Mayor Rudolph Giuliani and New York Police Commissioner William J. Bratton, who has attributed the dramatic decline in crime in New York City after 1990 to its rollout (Kelling and Bratton, 2015).

Corresponding research has arisen to evaluate the effectiveness of broken-windows policing—in practice, this literature has focused disproportionately on New York City, which experienced the largest decline in crime among major US cities. This literature produces mixed findings. On the one hand, time series analyses by Kelling and Sousa (2001), and Corman and Mocan (2005) find that misdemeanor arrests are negatively associated with future arrests for more serious crimes such as robbery and motor vehicle theft. On the other hand, later research has pointed out that these studies omit a control group and that New York's aggregate crime trends, while steeper, are broadly similar to those of other cities that did not institute a policy of broken-windows policing (Eck and Maguire, 2000; Rosenfeld, Fornango, and Baumer, 2005; Harcourt and Ludwig, 2006). The two most credible analyses—Harcourt and Ludwig (2006) and Rosenfeld, Fornango, and Rengifo (2007)—use precinct-level data on misdemeanor arrests and violations and find either no effect or very small effects.

Chandrasekher (2016) uses evidence from a work slowdown in 1997 when the New York City Police Department was negotiating in a union contract with the city. The work slowdown drastically reduced the number of ticket citations the NYPD issued. She finds that work slowdowns, at least in the case of acquiring public favor for negotiation, result in minimal effects on public safety. However, despite the decrease in ticket citations in every category, arrest levels were either maintained or increased for all serious crimes. Thus, it suggests that the substitution of policing away from citation of low level offenses does not substantially alter the serious crime rate, potentially contradicting the broken-windows theory. The key caveat from her study illustrating the lack of effects of reduced policing on public safety was the focus on citations, not arrests, with external validity only being applied to other contract-related slowdowns.

More fundamentally, there are a number of alternative explanations for New York's dramatic reduction in crime including the receding of the crack epidemic (Blumstein, 1995), changes in demographics that were originally poorly measured using coarse and aggregated data, general strategies to address disorder such as boarding abandoned buildings, and the implementation of the data-driven CompStat system (Weisburd et al., 2003). Accordingly, even if identification problems can be set aside, it is unclear that this literature can isolate the impact of disorder policing from other changes that drove crime down in New York City. Given the dramatic rollback of the New York City Police Department's "stop-and-frisk" policy in 2014 and the continued decline in serious crime, as well as the failure of the most careful studies to find evidence of large effects, we are skeptical that disorder policing has played a large role in the decline in crime in New York City. Overall, our reading of this literature is that there is a substantial lack of evidence in favor of proactive policing having any substantial effect on crime.

Of course, New York City is not the only municipality to experiment with disorder policing and, in our view, some of the strongest evidence can be found in research from other cities. Three papers that employ especially strong research designs are worth mentioning. Braga et al. (1999) provides the first experimental evaluation of a strategy designed explicitly to address disorder. In Jersey City, NJ, twelve of twenty-four crime hot spots were randomly assigned to receive an intervention that involved disorder policing as well as other place-specific treatments that were intended to reduce crime. Such treatments include clearing vacant lots, requiring store owners to clean store fronts, and facilitating more frequent trash removal. Treated places experienced large declines in both crime and calls for service. In a follow-up study in Lowell, MA, Braga and Bond (2008) attempted to further isolate disorder policing from other types of disorder reduction, randomly assigning seventeen Lowell hot spots to receive a general disorder policing strategy. This study also showed strong reductions in crime in treated areas. However, the greatest gains were found in areas with an especially heavy focus on situational crime prevention, as opposed to arresting larger numbers of low-level offenders.

Evidence in favor of an effect of misdemeanor arrests is far more limited. Finally, a particularly careful paper by Caetano and Maheshri (2014) finds no evidence of an effect of "zero tolerance" law enforcement policies on crime using microdata from police precincts in Dallas. Taken as a whole, the evidence suggests that reducing disorder is a promising strategy for controlling crime. However, it is difficult to characterize these reductions as deterrence. In particular, disorder reduction may simply help people to feel better about their neighborhoods, thus representing a shift in preferences, rather than movement along the curve that is induced by an increase in the perceived probability of capture by police. We remain skeptical that disorder policing provides evidence of deterrence.

2.3.2 Natural Experiments of Policing Deployment

The ubiquity of the hot spots, problem-oriented, and proactive policing literatures in criminology has spawned a parallel literature in economics that seeks to learn from natural experiments in police deployments. This literature is conceptually similar to the hot-spots literature with two exceptions. First, the identifying variation is naturally occurring in contrast to experimental manipulation, which may be excessively contrived. Second, several of the natural experiments identify the impact of a diffuse growth in resources, rather than a concentration of resources at particular hot spots.

Three prominent studies are those of Di Tella and Schargrodsky (2004); Klick and Tabarrok (2005); and Draca, Machin, and Witt (2011). Each of these studies leverages a redeployment of police in response to a terrorist threat or act. The appeal of these studies is that terrorist threats are plausibly exogenous with respect to trends in city-level crime and therefore represent a unique opportunity to causally infer about the response of crime to changes in policing routines. An important caveat is the underlying assumption that offending patterns are not affected by the terrorist event. Di Tella and Schargrodsky study the response of police in Buenos Aires to the 1994 bombing of a Jewish community center. In the aftermath of the bombing, Argentine police engaged in a strategy of "target hardening" synagogues by deploying disproportionate numbers of officers to blocks with synagogues or other buildings housing Jewish organizations. Di Tella and Schargordsky report that the intervention led to a large decline in motor vehicle thefts on the blocks that received additional police patrols. Notably, this result has been called into question by Donohue, Ho, and Leahy (2013), who reanalyzed the original data and report evidence that is more consistent with spatial displacement of crime rather than crime reduction. However, with respect to identifying behavioral changes among offenders, both stories are equally consistent with deterrence.

In a similar study, Klick and Tabarrok (2005) utilize the fact that increases in the terror alert levels set by the US Department of Homeland Security result in large deployments of police presence. They find that property crime (but not violent crime) tends to fall in Washington DC, with especially significant declines in areas that receive the more substantial police protection. With respect to the United Kingdom, Draca, Machin, and Witt (2011) study the 2005 London tube bombings which resulted in sizable shifts in the deployments of police from the suburbs to central London and find that "street crimes" such as robbery and theft are reduced considerably in areas that received additional officers.

With respect to studying variation in the spatial concentration of police, two additional papers are worth noting. Cohen and Ludwig (2003) exploit short-term variation in the intensity of police patrols by day of the week in several different Pittsburgh patrol areas. They found that shootings were considerably lower in areas and on days that received more intensive police patrols. With respect to the long-term consequences of patterns of police deployments, MacDonald, Klick, and Grunwald (2016) use a spatial regression discontinuity (RD) design to study the impact of especially intensive policing around the University of Pennsylvania, a large urban university campus. In particular, areas directly adjacent to the university received police patrols from both the university and municipal police. Areas slightly further away from the campus received only municipal police patrols. The finding is that street crimes are substantially higher in the blocks just outside the area patrolled by the university police relative to the blocks just inside the university patrol area.

2.3.3 Deciphering Deterrence versus Incapacitation

The literature has reached a consensus that increases in policing reduce crime, at least for a population-weighted average of US cities. With respect to police deployments and tactics, the literature supports the idea that crime is responsive to a visible police presence in hot spots and pulling-levers strategies of POP that involve focused deterrence, while evidence in favor of an effect of proactive policing strategies such as broken-windows and disorder policing is more questionable. A remaining issue is to address the degree to which each of these literatures is informative with respect to disentangling deterrence from incapacitation.

With respect to the aggregate staffing literature, Levitt (1998) provides the first attempt to systematically unpack the relationship between deterrence and incapacitation by empirically examining the link between arrest rates and crime, a relationship that is negative. Levitt posits that this negative relationship can be explained either by deterrence, incapacitation, or measurement errors in crime. Ruling out measurement errors as a likely culprit, he differentiates between deterrence and incapacitation using the effect of changes in the arrest rate for one crime on the rate of other crimes. As Levitt notes, "in contrast to the effect of increased arrests for one crime on the commission of that crime, where deterrence and incapacitation are indistinguishable, it is demonstrated that these two forces act in opposite directions when looking across crimes. Incapacitation suggests that an increase in the arrest rate for one crime will reduce all crime rates; deterrence predicts that an increase in the arrest rate for one crime will lead to a rise in other crimes as criminals substitute away from the first crime." Naturally, his interpretation rests on the sizable assumption that police deterrence is crime-type specific while offenders are not. Levitt concludes that deterrence appears to be the more important factor, particularly for property crimes.

Owens (2013) reports a similar finding, examining whether variation in police staffing resulting from the COPS hiring program led to increased arrests. Despite the fact that the program, which provided funding to increase the number of patrol officers in US cities, appears to have led to a decline in crime, no significant effect is found on the level of arrests. As a result, Owens concludes that there is little evidence in favor of incapacitation, which necessarily must operate through arrests, thus implying a large role for deterrence.

While analyses by Levitt (1998) and Owens (2013) are suggestive of deterrence having a meaningful role in crime reduction, it is nevertheless difficult to disentangle deterrence from incapacitation in this way. In particular, a null relationship between police and arrests is also consistent with the idea that police productivity decreases when there are fewer crimes to investigate. Moreover, the imprecise parameter estimates on arrest in Owens (2013) render it difficult to make strong claims regarding the null effect of police on arrests. For this reason, while the aggregate-data literature is ideal for understanding the overall relationship between police and crime, it is only somewhat informative with respect to the magnitude of deterrence. This point is further compounded by the observation that research has yet to document the degree to which offenders perceive or are aware of increases in police staffing (Nagin, 1998).

We suspect that the literature on police tactics is considerably more informative with respect to identifying deterrence. Specifically, offenders are more likely to be aware of an enhanced police presence in small, local areas than relatively small changes in the number of police in a city spread out over a large geographic area. Likewise, while offenders tend to commit crimes locally, in order for incapacitation to explain the large declines in crime that occur in hot spots, it would have to be the case that offending is so local so as to be specific to a group of one or two blocks. The large drops in crime that occur in crime hot spots after they are more aggressively policed are more consistent with deterrence than with incapacitation. Focused deterrence strategies are also particularly informative in that declines in crime have been shown to be specific to the focus of the intervention. To the extent that at least some offenders are generalists, rather than specialists who commit only a certain type of offense, such a pattern is more consistent with deterrence than with incapacitation.

In summation, while it remains possible that an increased police presence lowers crime by situating police officers in locations where they are more likely to arrest and incapacitate potential offenders, on the whole, the high degree of visibility around police crackdowns or hot spots policing suggests a potentially greater effect from deterrence.

2.4 Theoretical Foundations of Optimal Policing

Leveraging two potentially independent measures of police staffing (one from the FBI's Uniform Crime Reports and another from the US Census's Annual Survey of Government Employment) for a sample of 242 US cities over a fifty-one-year time period, Chalfin and McCrary (2018) construct measurement error corrected IV models and estimate police elasticities with remarkable precision, reporting elasticities of -0.67 $\pm$ 0.48 for murder, -0.56 $\pm$ 0.24 for robbery, -0.34 $\pm$ 0.20 for motor vehicle theft, and -0.23 $\pm$ 0.18 for burglary.

Translating into monetary terms, for every dollar spent on police, the average U.S. city can expect about $1.60 in reduced crime costs. Their estimates suggest that if you increase police officers by 10 percent, you can get something like a 5 percent reduction in cost of crime. For cities with a high level of crime, where police are relatively inexpensive, that's a tradeoff a city should be willing to make.

The above cost-benefit analysis (CBA) conducted in Chalfin and McCrary (2018) only internalizes the direct cost of crime, eschewing away from integrating the additional social costs from policing. Particularly in confrontational methods of policing, such as the Stop, Question, Frisk (SQF) tactic utilized in New York City for over two decades, there are nontrivial costs imposed upon the public, often concentrated among young black males. Though researchers may be far from empirically estimating what these costs may be and integrating them into CBAs, it is important to address their existence when discussing optimal levels of policing.

A recent paper by Manski and Nagin (2017) presents a theoretical model to understand where the optimal level of policing lies when focusing on a policing tactic that is confrontational, leading to social costs to innocent and stopped citizens. Although relatively simple, it provides a good framework to analyze tradeoffs between social benefits of crime reduction, opportunity cost of police to enforce that tactic, and social cost of police officers on innocent bystanders. There is an additional layer since the model can be disaggregated to understand these optimal levels of policing by demographic group and neighborhood, by assessing racial differences in crime rates and social cost of stops. Using case studies and crime trends from Chicago, New York City, and the entire United States, Manski and Nagin use their model to make generalized policy implications and introduce measures of relative and attributive risk to comprehend the disparate impact of policing social costs by age-race cohorts.

For the model to become practically valuable, there needs to be two advances in the literature: (1) There needs to be a clear and defined way to quantify the social costs of policing. For example, the cost could potentially be extracted by assessing how police intensity affects perceptions of police legitimacy and thereby impacts legal compliance and crime reporting. If there are ways to isolate what proportion of the potential increase in crime is from lack of compliance or poor reporting, we should be able to obtain a proxy for the parameter estimate of social cost. (2) In the model, there is only one policing tactic that can be used, and it implicitly involves nontrivial amounts of social cost. Preferentially, the model should be expanded to include substitutable policing tactics, each with their own efficacy of crime reduction, time or labor intensive cost, and social costs from confrontation. Overall, Manski and Nagin (2017) do a good job of integrating overlooked social costs into an optimal policing model and introducing unintended consequences from more intensive policing practices. Currently, we still support the conclusion of Chalfin and McCrary (2018) that many cities can increase their police staffing to reduce crime and make net savings, while also hoping the literature produces better estimates of the social cost of policing on the community.

2.5 Conclusion

This chapter was meant to provide a comprehensive yet manageable glance at the literature of policing. The extensive reference list in the chapter provides a springboard if you are interested in getting more into the details of the theory and empirical methodology of these studies. In general, criminologists and economists view potential offenders as rational actors who make the determination to engage in crime when their expected benefit is greater than their expected cost. To influence their decision-making, there are three main concepts in deterrence theory: certainty (their risk of apprehension), severity (cost of punishment if they are arrested), and celerity (how swiftly they will be punished for their actions).

These concepts motivate much of the research in policing, where we estimate the effect of increases in police staffing or certain police strategies on crime. The general consensus from the literature is that raising the number of police officers, which leads to a respective rise in the apprehension risk for offenders, *does* result in substantial crime reduction. These estimates often struggle with imprecision and vary quite a bit by location, type of crime, and estimation technique. Chalfin and McCrary (2018) advance the literature by circumventing the issues of measurement error in the crime and police staffing data, providing much more precise estimates. They use their estimates to provide the policy implication that suggests that cities can net save money by investing more in law enforcement, thereby preventing costly crimes. Manski and Nagin (2017) extend the theoretical analysis by insisting that the social cost of confrontational policing must be taken into account, since it alters the optimal level of policing. We recommend additions to the Manski and Nagin model to ensure practical model applications, but we recognize its contribution of providing a solid foundation for the literature to progress.

As we discuss in subsection 2.3.1, there is a substantial heterogeneity in the efficacy of policing strategies. There is a consensus that the standard model of policing, which focuses on random preventive patrols and rapid response time, does not significantly reduce crime or even fear of crime. We caution against drawing definitive causal conclusions from many of these studies, as they lack the empirical rigor to meaningfully evaluate their estimates. With that caveat on mind, we illustrate that there is limited to no evidence to suggest that disorder or proactive policing is an effective means to deter crime, but there are evaluations which support the use of hot-spots and problem-oriented policing, specifically in programs such as Operation Ceasefire in Boston and Project Safe Neighborhoods in Chicago. We finally discussed a literature that needs more quantitative evidence, but provides a good theoretical frame for understanding how disparate treatment in policing can result in reduced legal compliance and even crime reporting (Nagin and Telep, 2017; Desmond et al., 2016).

For future research in policing, it is incredibly important—especially given the current national dialogue on policing and race—that we internalize the genuine discrepancies in overall social cost and direct cost by race. It is also imperative that we do not forget the effective and necessary role that police officers have in deterring crime and disorder in the community, while also providing many other additional emergency services. There must be a poignant and nuanced discussion going forward ensuring that any policies or regulations set on law enforcement are evidence-based and provide a better sense of procedural justice by limiting disparate impact.

Chapter 3

Rural Hospital Closings and Patient Welfare: The Effects of Closures on Quality and Availability of Hospitals[1]

3.1 Introduction

The population shift from rural to urban regions has decreased the population density around hospitals in small towns and rural areas. Meanwhile, improved road systems and improved ability to deliver online health services may make urban hospitals more attractive for rural patients. Rural hospitals find it increasingly difficult to maintain a large enough patient base to cover their costs on specialized care and threaten their ability to provide even the most common procedures. These financial exigencies threaten the future economic viability of rural hospitals (American Hospital Association, 2011). As shown in Figure 3.1, the number of rural hospitals declined steadily after 1990, continuing a decline that started in the 1970s (Capalbo and Heggem, 1999). Urban hospitals experienced their own consolidations, but their numbers have rebounded over the past 10 years.

Rural hospitals are particularly dependent on publicly subsidized health care provision with almost 60% of their revenues coming from Medicare and Medicaid. Medicare's Critical Access Hospital program started in the 1990s. It was devised to support hospitals in isolated areas where residents had few other health care options. Recently, Iowa had 82 out of 119 hospitals receiving funds from the program. Federal budgetary constraints may lead to cuts in the subsidy which would lead to further closure of rural hospitals in Iowa and elsewhere. This paper evaluates how rural residents respond to the loss of a local hospital and sheds light on how to reallocate federal funds to minimize those costs.

We use a conditional logit specification to model a rural patient's choice of local, urban, or specialty hospitals as a function of distance to and quality of each of the three hospital

[1]This chapter is coauthored with Dave Jones and Peter Orazem. We thank Mark Imerman for the help in locating the data set, Juan Francisco Rosas for helping to compile the information, and conference and meeting participants at the State Capitol of Iowa, Midwest Economics Association, and International Health Economics Association. Policy briefings about this project are found in Iowa State University's CARD Agricultural Policy Review and University of California Giannini Foundation's ARE Update.

options. We derive separate estimates of hospital choice parameters for inpatient visits, for outpatient visits, and for hospital care types—emergency, urgent, or elective. We use these estimates to simulate how potential hospital closings will alter the hospital choices made by rural Iowa patients, providing a closure cost in terms of changes in expected distance to and quality of hospital. We explore two potential hospital closing scenarios: 1) closing 15% of the least used rural hospitals and 2) closing 25% of the lowest quality rural hospitals. Differences between the two scenarios provide lessons on how to consolidate hospitals with the least harm to patients.

The empirical model shows that hospital choice is strongly affected by distance from home and hospital quality. Simulations using the model parameters show that closing 15% of the least used rural hospitals results in a marginal increase in expected distance of 1.8 miles and a small increase in expected quality. Closing 25% of the lowest quality rural hospitals results in a larger increase in expected distance of 2.9 miles with a significant increase in expected quality. The findings suggest that if hospital closures are necessary in the future, closing the lowest quality hospitals is a better policy prescription than closing the least used hospitals since it results in a substantial increase in quality with only a marginal increase in distance.

3.2 Literature Review

The impact of rural hospital closure depends on the extent to which rural residents bypass their existing rural options. Liu et al. (2007) reported that 20-50% of rural patients bypass their nearest rural hospital. The bypass decision varies by patient age and gender, and by method of payment (Basu, 2005; Basu and Cooper, 2000; Bronstein and Morrisey, 1991). The adverse consequences of a hospital closure will differ by the urgency of the diagnosis. For example, Kozhimannil et al. (2018) found that the closure of rural obstetrics services was correlated with more premature births. Buchmueller et al. (2006) found that the closure of urban hospitals resulted in a lower survival probability from heart attacks.

Radcliff et al. (2003) concluded that patients view rural hospitals as a viable option for minor ailments, but strongly prefer urban hospitals for more serious health issues. Regardless of illness, the elderly are less likely than younger patients to bypass their nearest hospital (Hogan, 1988). Chandra et al. (2016) discovered that patients who have more ability to choose among hospitals, such as transfer patients or non-emergency admissions, pick higher quality hospitals. Lin, Allan and Pennington (2002) found that patients seeking general medical or obstetrical care have lower bypass rates that those receiving complex medical, general surgery, or specialty surgery services.

Radcliff et al. (2003) found that Medicare and uninsured patients have lower bypass rates while patients in managed care or commercial insurance have higher bypass rates. Using hospital discharge data for patients in obstetric care in California, Ho and Pakes (2014) showed the price paid by the insurer affects the patient allocation among hospitals within a given network, particularly for patients whose plans pay higher capitation fees for physician groups. In many studies, rural residential choice of hospital is very sensitive to distance (Rieber et al., 1996). Studies consistently found a decrease in the likelihood of bypassing local hospitals as distance to alternative hospitals increases (Bronstein and Morrisey, 1991;

Buczko, 1994; Radcliff, et al., 2003), but residents of communities with greater per capita incomes are more likely to bypass the nearest hospital (Basu and Cooper, 2000; Dranove et al., 1993; Goldsteen et al., 1994).

Hospital quality is likely to be a key factor explaining the incentives to bypass rural hospitals. Liu et al. (2007) conducted a preliminary examination of the role of hospital quality in hospital choice. Their study surveyed 647 hospital inpatients for their assessments as to why patients would bypass a local hospital. Following the lack of local specialists, the second most common reason cited for bypassing a local hospital was poor reputation or quality of local care. Chandra et al. (2016) used three different measures of quality: patient survival and readmission, whether care provided met established guidelines, and patient satisfaction, to show that higher quality hospitals gain market share over time. To our knowledge, no previous study of rural health has examined the tradeoff between distance and quality in deciding whether to use or bypass rural hospitals or to examine how those decisions vary by inpatient, outpatient, or type of health condition.

3.3 Methodology

In order to assess how possible hospital closures affect rural hospital choice, we need to establish how rural patients weight hospital quality and distance. While a hospital closure may lower utility because of increased distance to the nearest hospital, the decrease in utility may be partially offset if the remaining hospital options are of higher quality. Consistent with that reasoning, we assume that rural residents choose hospitals to maximize expected utility from hospital services with a particular medical urgency. Utility is assumed to take the following form:

$$V_{ij}^k = V(D_j^k, Q_j^k, I_{ij}, M_{ij}, Z_{ij}); k \in \{S, U, R\} \tag{3.1}$$

V_{ij}^k is the indirect utility that individual i from residence j gets from receiving health service from hospital of type k. The three hospital types are Rural (R), Urban (U), and Specialty (S) hospitals. We assume that patients will evaluate these three options based on their closest alternative of each hospital type. As a result, the distance to (D_j^k) and quality of (Q_j^k) the kth hospital option will vary across rural areas, and so individual choice will depend in part on the convenience and quality of the available choices in the individual's location. We expect that distance to hospital k lowers utility from that hospital ($V_{D^k}^{'} < 0$). Utility increases in hospital quality ($V_{Q^k}^{'} > 0$). Hospital choice will also depend on whether the patient is insured and the type of insurance (I_{ij}), both of which affect the patient out-of-pocket cost of each hospital option. The anticipated cost and quality will also depend on the particular medical urgency (M_{ij}). Finally, individual attributes (Z_{ij}) including age, gender, education, tastes, and value of time will affect relative utility from the three hospital types.

Individual i in location j will choose hospital option k if

$$H_{ij}^k = \begin{cases} 1 \text{ } iff \text{ } V(D_j^k, Q_j^k, I_{ij}, M_{ij}, Z_{ij}) \geq V(D_j^l, Q_j^l, I_{ij}, M_{ij}, Z_{ij}) \forall k \neq l \\ 0 \text{ } otherwise \end{cases} \tag{3.2}$$

Inspection of Equation 3.2 shows that only distance and quality will affect the choice of hospital type, k. Factors that only vary across individuals but not across hospital choices such as tastes, education, wages, medical ailment, and insurance status will not affect choice of hospital type directly as they raise or lower utility equally across all hospital types. However, they can affect hospital choices indirectly to the extent that these factors are not separable from distance and quality. For example, if insurance status affects the marginal disutility of distance, then it can affect hospital choice through its interaction with distance.

To make Equation 3.2 operational, we need to approximate the form of the indirect utility functions, $V(\cdot)$. To simplify notation, define the vector $X_{ij} = (I_{ij}, M_{ij}, Z_{ij})$. Then

$$V_{ij}^k = \alpha_d D_j^k + \alpha_q Q_j^k + \alpha_{dq} D_j^k * Q_j^k + \beta_{dx} D_j^k * X_{ij} + \beta_{qx} Q_j^k * X_{ij} + \gamma_x X_{ij} + \epsilon_{ij}^k \tag{3.3}$$

The probability of choosing hospital type k over alternative l will be

$$\begin{aligned} Pr_{ij}^k = Pr(H_{ij}^k = 1) = Pr(\{\alpha_d\{D_j^k - D_j^l\} + \alpha_q\{Q_j^k - Q_j^l\} + \alpha_{dq}\{D_j^k * Q_j^k - D_j^l * Q_j^l\} \\ +\beta_{dx}\{D_j^k - D_j^l\} * X_{ij} + \beta_{qx}\{Q_j^k - Q_j^l\} * X_{ij}\} > \{\epsilon_{ij}^l - \epsilon_{ij}^k\}) \end{aligned} \tag{3.4}$$

If the error terms are independent draws from an extreme value distribution, the parameters can be estimated using a conditional logit specification. Because the distance and quality vary with each hospital type k, the parameters will reflect how quality and distance interact to make any given choice with a common set of coefficients: $\alpha_d, \alpha_q, \alpha_{dq}, \beta_{dx}, \beta_{qx}$. These coefficients represent the utility weights attached to each factor affecting hospital choice. Note that in comparing Equation 3.4 with Equation 3.3, it is only the interacted terms with X_{ij}, distance, and quality that will influence the hospital choice. All individual attributes and their higher moments are common fixed effects across hospital types, and so they are differenced away in the estimation.

In Equation 3.4, medical urgency only affects hospital choice through the channel of altering the marginal disutility of hospital distance or the marginal utility of hospital quality. Initially, we assume that the decision is unaffected by medical urgency except for the broad distinction of whether the condition can be handled on an outpatient basis or requires inpatient services. We accommodate that difference by estimating the conditional logit specification separately for inpatient and outpatient choices. We then relax this urgency restriction by estimating separate hospital choice equations for specific inpatient admission types. For each set of results, it is important to interpret the coefficients within the context of the admission type. For example, we will have one set of results that averages the hospital choice effects of factors across all inpatient hospitalizations, and then another set of results for each admission type, which separate the patients by urgency.

Because patient attributes are constant across the three hospital types, they are controlled as a fixed effect. However, some of these attributes may alter the marginal effects of distance and quality. Hence, we include them interacted with the hospital attributes. We control for age and gender. Insurance is an indicator of whether the individual is covered by a private or employer-provided insurance plan. Self-pay is whether the visit is paid for out-of-pocket. Medicare, Medicaid, or some other type of public insurance covers the rest of the visits.

The coefficients are difficult to interpret directly, so we transform the results for ease of interpretation. First, we estimate the marginal effect of each factor on the probability of choosing a hospital, evaluated at sample means. Afterward, we run a separate logit regression to obtain the uninteracted coefficients of distance and quality, allowing the resulting coefficients to correlate with missing interacted variables. We then compute the elasticities for distance and quality for each patient group.

3.4 Data

We require a data set that has sufficient variation to allow us to estimate the tradeoff between quality and distance in hospital choice. As urban areas have easily accessible hospital choices, a cross-section of rural patients is more likely to provide the needed variation in distance and quality. Additionally, rural communities are more typically afflicted with hospital closures, providing a more appropriate setting in which to estimate closure costs.

The Iowa Hospital Association records include every inpatient admission and outpatient visit to Iowa hospitals (Iowa Hospital Association, 2004). We have access to the recorded visits occurring between January 1, 2002 and December 31, 2002. The inpatient database includes 209,687 patients that were treated and discharged from an Iowa hospital during this period. Patients who were admitted prior to this period, but discharged during this period are not in the database. Similarly, patients who had not yet been discharged at the end of this period are not in the database. Non-Iowa residents in Iowa hospitals are not included, neither are Iowa residents who received their care out of state. The database also includes records on 138,685 outpatient records. Inclusion in the outpatient database does not require admission or release from the hospital, but only that the patient received treatment at an Iowa hospital.

The data set is old but it is unique to our purpose. First, it is the universe of every patient in Iowa in that year and so there is no selection issue save patients that seek treatment out of state. Second, we were able to access a consistent measure of hospital quality across the universe of hospitals which enables us to measure the tradeoffs between distance and quality. Finally, the data set includes both inpatient and outpatient visits and visits by emergency status and by diagnosis, and so we can assess how the quality-distance tradeoff changes with the type of health care required.

We divide hospitals into three groups: rural, urban, and specialty. Rural hospitals are those located in counties with rates 4-9 using the USDA Rural Urban Continuation Code. Those counties had an average population of 17,000. Urban hospitals are those located in counties containing a metropolitan statistical area which in Iowa would be the counties with populations of at least 80,000 at that time. The final group includes the specialty and research hospitals in Des Moines and Iowa City.

Our focus is on the determinants of hospital choice for rural residents—patients whose residence is in a zip code listed as rural by the U.S. Census in 2000. To calculate distance, this study measures the straight-line distance from the latitude and longitude of the patient's home zip code to the latitude and longitude of the nearest rural, the nearest urban, and the nearest specialty hospital even if none of those hospitals were chosen. Had we used distance to the chosen hospital, distance would be endogenous. While the choice set is the hospital type (rural, urban or specialty), the distance and quality measures are to the nearest rural, urban or specialty hospital.

As shown in Table 3.1, hospital choices did not differ much between inpatient and outpatient treatments. Almost 70% of rural residents chose a rural hospital for inpatient and outpatient service. Urban hospitals served 12% of rural residents and 18% were served by specialty hospitals. The average rural patient lives about five miles from a rural hospital, but lives 51 miles from the nearest urban hospital and 71 miles from the nearest specialty hospital.

Health Grades, Inc. reported information on hospital quality. Quality is reported by one-to-five-stars on various criteria. We use the two most common ailments, heart failure and pneumonia, to measure hospital quality.[2] Deaths from heart failure were recorded during the length of the patients' admissions at a hospital, and again on whether additional patients died within six months after discharge. Deaths from pneumonia are also recorded during the stay and within six months of discharge. A third measure recorded deaths within one month of discharge. Rather than using a single measure, we use the average of the reported measures across the five outcomes. When a hospital is evaluated only on a subset of the measures, we take the average of the reported measures. This strategy is consistent with the presumption that the most common reason for partial reporting would be the lack of sufficient cases to generate a reliable measure. Presuming random measurement error, an average of several indexes will have a lower measurement error than would any of the individual indexes. Table 3.1 shows that specialty and urban hospitals were of higher average quality than rural hospitals.

3.5 Conditional Logit Results

The results of our model of inpatient and outpatient hospital choice are presented in Table 3.2. The key variables of interest are distance to and quality of the nearest hospital of each type regardless of which hospital was actually chosen.[3] Distance is the single largest driving factor in the choice of hospital. At sample means, a 10% increase in distance lowers the probability of choosing that hospital type for both inpatient and outpatient services by just over 13%. There is an apparent tradeoff between distance and quality. A 10% increase in quality raises the probability of inpatient hospital choice by 4.2% and outpatient hospital choice by 0.9%. Because quality attracts increased patient use, some patients will opt for more distant but higher quality options. The interaction terms suggest that women, older patients, and patients who do not have private insurance or who self-pay are more distance sensitive.[4] These effects are similar for inpatient and outpatient care. Quality is more important for young patients, and more important for inpatients with public insurance,

[2]We were able to get information on hospital quality based on heart and pneumonia deaths for 117 of the 119 hospitals in Iowa. Quality measures based on other criteria were missing for at least 31% of the hospitals.

[3]Using the attributes of the chosen hospital rather than the nearest hospital would add a source of endogeneity to the estimation as the hospital factors would enter as a result of the choice.

[4]Most of the patients have access to public insurance through Medicare or Medicaid. Only 3.6% of inpatients and 2% of outpatients pay out-of-pocket.

although only marginally. As quality of a patient's choice set increases, distance becomes a larger disincentive, presumably because the value of traveling farther for an alternative is reduced.

We expect that those patients with more time sensitive needs might be more sensitive to distance and less sensitive to quality. For inpatient hospitalizations, the three admission codes we used in our analysis are emergency, urgent, and elective—ordered from most to least critical. Consistent with our expectations, emergency and urgent admissions are much more sensitive to distance than elective ones Table 3.3. A 10% increase in distance leads to a 18.1% and 16.4% reduction in the probability of choosing a hospital for emergency and urgent patients respectively, while it only leads to a 8.5% drop for elective procedures. Similar to distance, emergency and urgent care admissions are significantly more sensitive to quality than elective admissions. These results suggests that there are apparent tradeoffs between distance and quality even for the most time-sensitive admissions.

Under all admission types, the interaction between distance and quality is negative and statistically significant. As distance rises, quality becomes less of a draw. And as quality increases, there is less value to travel farther for an alternative. The fact that this outcome holds for all estimates suggests that the finding is a behavioral regularity in hospital choice.

In Table 3.4, we summarize the results from the estimation in Table 3.2 and Table 3.3, plus comparable estimates for the most common inpatient and outpatient diagnoses. The distance elasticities are all negative and statistically significant. All the quality elasticities are positive and all but one are statistically significant. Rural hospitals are chosen most commonly for the admission types with the largest distance elasticities. When hospital choice is relatively insensitive to distance, rural patients are more likely to bypass their nearest option for a more distant and higher quality hospital. Quality is particularly important for inpatient treatment including emergency and urgent admissions and for procedures related to pregnancy and newborns. The distance and quality elasticities are quite large in general, suggesting substantial utility tradeoffs between distance and quality. Thus, there is a mechanism to reduce the lost utility from the closure of nearby rural hospitals if the remaining hospital choices are of higher quality.

3.6 Hospital Closing Simulations

3.6.1 Simulation Model of Hospital Closings

The conditional logit specification performs well in modeling the rural patient decisions to select or bypass the nearest rural hospital, highlighting the importance of quality as well as distance in these choices. The model parameters also allow us to simulate how hospital closings will alter hospital choice, changing the expected distance to and quality of a patient's hospital choice—a proxy for the closure cost.

If the nearest rural hospital (R) for individual i has closed, the new indirect utility is computed using R', the next nearest open rural hospital with its associated quality and distance:

$$V_{ij}^{k'} = V(D_j^{k'}, Q_j^{k'}, I_{ij}, M_{ij}, Z_{ij}); k' \in \{S, U, R'\} \tag{3.5}$$

The difference between equations Equation 3.5 and Equation 3.1 is the lost utility to individual i associated with the closing of the nearest rural hospital. Individuals whose nearest rural hospital is unaffected by the closings experience no lost utility. We shall reference Equation 3.1 as the 'baseline' model and Equation 3.5 as the 'post-closing' model.

Given Equation 3.5, the implied hospital choice post-closing model is derived in the same manner as in Equation 3.2 except with new distance and quality for hospital R':

$$H_{ij}^{k'} = \begin{cases} 1 \ \textit{iff} \ V(D_j^{k'}, Q_j^{k'}, I_{ij}, M_{ij}, Z_{ij}) \geq V(D_j^{l}, Q_j^{l}, I_{ij}, M_{ij}, Z_{ij}) \forall k' \neq l \\ 0 \ \textit{otherwise} \end{cases} \tag{3.6}$$

Similarly, we can use the same probability model in Equation 3.4 using Equation 3.6 to determine the probability $Pr_{ij}^{k'}$ that individual i will choose hospital k' given that we replaced R for R'. Rural hospital closings affect hospital choice through changes in expected distance and quality of selected hospitals. The predicted probability of choosing each hospital type for both the baseline model and the post-closing model allows for the computation of expected quality (EQ) and expected distance (ED):

Baseline:

$$EQ_{ij} = \sum_k Q_j^k * Pr_{ij}^k; \qquad ED_{ij} = \sum_k D_j^k * Pr_{ij}^k, \quad k \in \{S, U, R\} \tag{3.7}$$

Post-closing:

$$EQ_{ij}^{'} = \sum_k Q_j^{k'} * Pr_{ij}^{k'}; \qquad ED_{ij}^{'} = \sum_k D_j^{k'} * Pr_{ij}^{k'}, \quad k' \in \{S, U, R'\} \tag{3.8}$$

We generate two simulations of possible rural hospital closings to examine the impacts on rural patients: 1) closing 15% of the least used rural hospitals, and 2) closing 25% of the lowest quality rural hospitals.[5] We estimate the expectations in Equation 3.7 and Equation 3.8 for both hospital closing scenarios, separating the analysis by inpatient admission type and whether the procedure was inpatient or outpatient. The costs of closing rural hospitals are estimated in terms of changes in expected distance and quality of hospital choice. In the simulation where we close the least used hospitals, we opt for a utilization threshold of 50%. Therefore, we shutter all of the hospitals that were chosen less than 50% of the time. These hospitals would be the most threatened if the Medicare Critical Access Hospital standards become more stringent. Rural patients living nearest to the closed hospitals under this simulation selected them between 11% and 41% of the time for inpatient hospitalizations, and between 6% and 41% of the time for outpatient services.

The second simulation assumes that the reduction in subsidy would be tied to hospital quality rather than current use, closing 25% of the lowest quality rural hospitals. The 24 hospitals closing under this scenario had quality rankings ranging from 1 to 1.83, well below the average hospital quality of 3.4.

[5]The least used hospitals differ slightly between inpatient and outpatient visits, leading to minor differences in the hospitals dropped in the simulations depending on diagnosis.

3.6.2 Simulation Results: Closing the 15% Least Used Rural Hospitals

Table 3.5 illustrates the simulated effect of closing the 15% least-used hospitals. To examine heterogeneity in existing tastes for hospitals, we present different estimates based on all rural patients ('Chose Any Hospital'), the subset of rural patients who originally chose a rural hospital ('Chose Rural Hospitals'), and the subset of rural patients who originally chose one of the hospitals closed in the simulation ('Chose Closed Hospitals'), with each respective group facing more adverse effects. Closing the least-used hospitals has a surprisingly small effect on expected distance traveled and on quality. In the outpatient data, distance increases by only 2.2-miles (15.3%) and quality rises by 0.09 (3.1%). Though the percentage term seems high for distance, the magnitude of these changes are rather small. To put in perspective, the baseline expected distance for outpatient procedures is 14.1 miles and for expected quality is 2.8.

The effects are larger when we confine the estimates to patients whose baseline choice was a rural hospital or a closed hospital. The patients who sought out outpatient procedures and selected closed hospitals have an expected distance increase of 7.0 miles (279%) and a decrease in expected quality of -0.36 (-28.0%). In the inpatient data, there remains a consistent increase in distance and decrease in quality for those who chose rural hospitals and those who chose a closed hospital—the more adversely affected groups. Patients who chose closed hospitals experience an increase in expected distance of 5.9 miles and a decrease of -0.43 in expected quality.

Closure costs may be greater depending on urgency of care. Examining subsets of the inpatient data by admission type (Emergency, Urgent, and Elective), all types see a small reduction in the probability of choosing their closest rural hospital (Table 3.5), likely because it is the least used hospitals that are closed. The drop in probability of choosing the closest rural hospital is largely shifted to the urban hospital for patients with emergency and urgent conditions, but is distributed between urban and specialty hospitals about evenly for elective admissions. The elective admission types had the lowest baseline probability of choosing a rural hospital to begin with since these patients had less time-sensitive conditions and were most able to bypass the closest rural option, illustrated by the low distance elasticity in Table 3.4. The hospital closings result in a small overall increase in expected distance ranging from 1.2 to 1.7 miles across admission types. With increases in expected quality being all below 1%, quality is not significantly affected when averaging across the entire admission type. The results do not change drastically when looking at the sample of patients who chose a rural hospital.

However, focusing on patients who chose a closed hospital under the simulation, the costs become more apparent. Patients with emergency and urgent admissions experience a loss in expected quality of -37.2% and -32.7% respectively, while quality for elective admissions falls by -13.4%. On the other hand, expected distance in absolute terms does not differ appreciably across admission type for patients who chose a least-used hospital. Expected travel distance rises by 6.5 miles for patients undergoing elective procedures, while patients with emergency and urgent conditions are expected to travel 5.4 and 5.0 miles farther, respectively. However, the percent changes between admission types are radically different since the baseline distance for elective patients who chose a closed hospital is quite small.

The conclusion from Table 3.5 is that there are only modest increases in distance and nearly negligible changes in quality from closing the least utilized rural hospitals. The magnitude of these changes vary by admission type, but even for the most concentrated negative impacts, we see at most a 7.0 mile increase in expected distance.

These small changes may seem surprising except that optimizing patients are already adapting to the available choices of hospitals. As the population ages and anticipates increased likelihood of necessary emergency or critical hospital care, rural residents move closer to hospitals that they trust. As a result, many patients have already voted with their feet regarding their preferred hospital and so closing less-favored hospitals has a smaller impact than would be true if rural populations could not move.

3.6.3 Simulation Results: Closing the 25% Lowest Quality Rural Hospitals

The findings from the simulated closure of 25% of the lowest quality rural hospitals are presented in Table 3.6. Compared to eliminating the least-used hospitals, closing the lowest quality hospitals has a larger effect on distance but results in improved quality of the remaining choices.[6] For outpatient data, there is an increase in expected distance of 2.8 miles (19.9%) and an increase in expected quality of 0.26 (9.2%). The patients who chose rural hospitals and closed hospitals experience more pronounced effects in terms of distance and quality. Patients with outpatient procedures who chose a closed hospital saw an increase of 9.5 miles in expected distance but with an increase of 0.7 in expected quality.

Similar magnitudes are found in the inpatient simulations, where we witness an increase in expected distance of 2.9 miles (20.5%) and an increase in expected quality of 0.22 (7.6%). Patients who chose a closed hospital in the baseline experience a rise of 9.6 miles in expected distance partially offset by an increase of 0.7 in expected quality. Closing 25% of the lowest quality hospitals also produces similar effects across admission types as it did in the inpatient and outpatient data sets (Table 3.6). The probability of choosing a rural hospital falls by 1.2 to 2.5 percentage points. This drop is smaller than in the previous scenario because we are not closing any higher quality rural hospitals that might attract more distant rural patients.

Across all admission types, there is a reduction in the probability of choosing a rural hospital resulting in a marked increase in expected distance and expected quality. Elective admission types experience similar effects to the inpatient and outpatient results, with the patients who chose rural hospitals and closed hospitals getting increased expected quality of 7.8% and 59.3% respectively. Patients with emergency and urgent admissions face an increase in expected distance of 2.6 and 3.8 miles respectively—which is 1-2 miles more than the expected distance from closing the least used hospitals. However, expected hospital

[6] It is important to note that this simulation does close more hospitals, and some of the differences between the scenarios are mechanical. The discrepancy between the number of hospitals closed in the least used and lowest quality scenarios is for aforementioned feasibility reasons. See the end of subsection 3.6.1 for more information.

quality rises by nearly 9%. For patients who chose the lowest quality hospitals originally, the effects are accentuated with expected distance increases of 9.9 and 10.7 miles for the emergency and urgent types, almost double the increase in expected distance instigated from closing the least used hospitals. However, because these patients had previously selected low quality hospitals, emergency and urgent admissions experience rising expected quality of 76.7% and 49.3% respectively.

These simulations suggest that closing the lowest quality hospitals is preferable to closing the least used hospitals. The rise in expected distance is partially offset by the rise expected service quality. Nevertheless, these increases in expected distance from either hospital closing scenario may be potentially too large for patients with critical time-sensitive conditions. It is plausible that the closures could be procedure specific so that the most distance sensitive services such as urgent/emergency procedures could be retained while less distance sensitive services are consolidated among the higher quality alternative hospitals.

3.7 Discussion and Conclusion

The outmigration of rural residents to urban areas requires a re-evaluation of rural health care delivery. Federal budgetary constraints and waning rural hospital demand threaten the sustainability of many rural hospitals. This paper highlights distance as the most significant factor in hospital choice, while illustrating the burden of plausible hospital closing scenarios on rural residents.

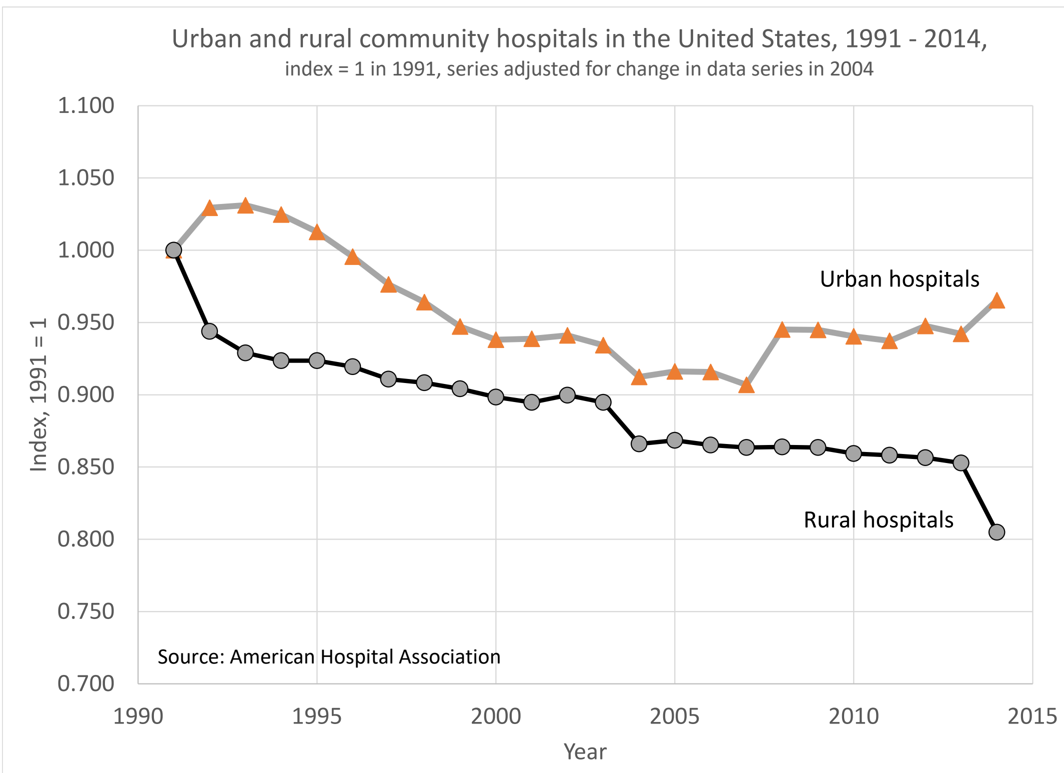
A notable finding of this paper is that rural patients value hospital quality and will travel greater distance to access better hospital services. The tradeoff is most salient for inpatient treatments and for emergency or urgent care. Proximity largely drives hospital choice for elective procedures and outpatient services. Our results are consistent with previous research that concludes patients with severe or complicated issues will seek out higher quality care, while people with time-sensitive conditions and the elderly are more distance sensitive.

Our simulations show that closing 15% of the least used rural Iowa hospitals results in a marginal increase in distance (around 1.8 miles) and a negligible change in quality, while closing 25% of the lowest quality hospitals results in a marginal increase in distance (around 2.9 miles) and a significant increase in quality. Closing the 15% least-used hospitals has a more pronounced effects on expected quality and distance for outpatient than inpatient admissions, while closing the 25% lowest-quality hospitals has similar effects on inpatient and outpatient hospital choice.

We conclude that if consolidation occurs, targeting low quality hospitals rather than the least used hospitals for closure is the better policy alternative. Closing the poorest quality hospitals results in a substantial increase in quality with only a marginally higher increase in distance.

Figures

Figure 3.1: Trends in Hospital Closings by Type



Tables

Table 3.1: Mean Values of Variables by Hospital Location and Inpatient/Outpatient Status

	Inpatient					**Outpatient**			
	Total	Rural	Urban	Specialty		Total	Rural	Urban	Specialty
Hospital Choice (Share)		0.69	0.12	0.19	Hospital Choice (Share)		0.70	0.12	0.18
Distance/100	0.43	0.05	0.51	0.71	Distance/100	0.42	0.05	0.51	0.70
Quality	3.43	2.45	3.72	4.13	Quality	3.39	2.46	3.69	4.01
Age/10	5.45				Age/10	5.57			
Male	0.41				Male	0.44			
Insurance	0.35				Insurance	0.45			
Self-pay	0.036				Self-pay	0.020			

Table 3.2: Conditional Logit Estimation of Rural Resident Choice of Hospital Quality and Distance by Inpatient and Outpatient Status

	Inpatient (N=209,687)		**Outpatient (N=138,685)**	
	Coefficient	Marginal Effect	Coefficient	Marginal Effect
Distance	-3.4564	-0. 6553	-0.6890	-0.1160
	(-53.01)	(-61.10)	(-8.48)	(-8.80)
Age x Distance	-0.0202	-0.0038	-0.2755	-0.0464
	(-3.76)	(-3.74)	(-38.19)	(-34.76)
Male x Distance	0. 8717	0.1653	0.3669	0.0617
	(30.92)	(30.92)	(10.95)	(10.95)
Insurance x Distance	0. 3123	0.0592	-0.9684	-0.1630
	(9.26)	(9.32)	(-25.19)	(-24.12)
Selfpay x Distance	-0. 6374	-0. 1209	-0.8261	-0.1390
	(-7.51)	(-7.48)	(-7.57)	(-7.54)
Quality x Distance	-0. 3013	-0. 0571	-0.4507	-0.0759
	(-22.20)	(-20.36)	(-26.22)	(-22.48)
Quality	0. 3268	0.0619	0.2023	0.0341
	(53.58)	(49.05)	(24.51)	(23.24)
Age x Quality	-0. 0189	-0.0036	-0.0138	-0.0023
	(-18.54)	(-19.35)	(-9.82)	(-10.16)
Male x Quality	-0.0046	-0.0009	-0.0081	-0.0014
	(-0.71)	(-0.71)	(-1.01)	(-1.01)
Insurance x Quality	-0.0337	-0.0063	0.0132	0.0022
	(-4.88)	(-4.92)	(1.62)	(1.63)
Selfpay x Quality	-0.0052	-0.0010	0.219	0.0369
	(-0.26)	(-0.26)	(7.09)	(7.03)
Pseudo R2	.39		.40	
Log likelihood	-139883.5		-91342.9	
Elasticities				
Distance	-1.31		-1.32	
	(-243.4)		(-201.2)	
Quality	0.42		0.09	
	(53.92)		(8.17)	

Dependent Variable is Choice of Rural, Urban, or Specialty Hospital.
z-statistics are in parentheses.

Table 3.3: Conditional Logit Estimation of Rural Inpatient Hospital Choice by Admission Type

	(1) Emergency **(N = 52,096)**	(2) Urgent **(N = 68,648)**	(3) Elective **(N = 71,012)**
Distance	-4.7664	-4.1796	-5.1869
	(-26.95)	(-31.20)	(-5.54)
Age x Distance	-0.1036	-0.1408	-0.1601
	(-6.99)	(-11.97)	(-18.87)
Male x Distance	0.8056	1.3875	1.0440
	(11.20)	(23.84)	(26.88)
Insurance x Distance	0.3662	0.0523	-0.1399
	(4.24)	(0.75)	(-3.07)
Selfpay x Distance	0.4232	-1.4890	-0.5691
	(2.49)	(-7.65)	(-4.72)
Quality x Distance	-0.1130	-0.1517	-0.5321
	(-3.35)	(-5.52)	(-28.27)
Quality	0.5021	0.2778	0.2856
	(36.62)	(22.57)	(29.38)
Age x Quality	-0.0339	-0.0066	-0.0170
	(-14.57)	(-3.06)	(-10.07)
Male x Quality	-0.0085	-0.0428	0.0197
	(-0.59)	(-3.26)	(1.90)
Insurance x Quality	-0.0355	-0.0238	-0.0156
	(-2.22)	(-1.76)	(-1.50)
Selfpay x Quality	-0.1447	0.0703	0.0883
	(-3.57)	(1.58)	(-2.67)
Pseudo R2	0.50	0.48	0.26
Log Likelihood	-28580.5	-39360.6	-57359.6
Elasticities			
Distance	-1.81	-1.64	-0.85
	(-119.7)	(-143.4)	(-129.3)
Quality	0.76	0.51	0.18
	(43.0)	(32.2)	(15.4)

Dependent Variable is Choice of Rural, Urban, or Specialty Hospital.
z-statistics are in parentheses.

Table 3.4: Cross Sample Comparisons of Elasticities and Hospital Selection Rates

Inpatient Elasticities				**Inpatient Hospital Selection Rates**		
	Distance	Quality	N	Rural	Urban	Research
All Inpatient	-1.32***	0.42***	209,976	0.69	0.12	0.19
Emergency	-1.82***	0.76***	52,219	0.74	0.08	0.18
Urgent	-1.64***	0.51***	68,648	0.74	0.14	0.12
Elective	-0.85***	0.18***	71,097	0.6	0.14	0.26
Diagnosis						
Respiratory	-2.13**	0.080**	24,362	0.81	0.08	0.11
Circulatory	-0.91**	0.33**	37,630	0.61	0.16	0.23
Digestive	-1.72**	0.047	21,277	0.76	0.1	0.14
Muscular	-0.87**	0.40**	20,745	0.58	0.19	0.23
Pregnancy	-1.73**	0.56**	20,033	0.7	0.13	0.18
Newborn	-1.65**	0.56**	19,238	0.69	0.13	0.18
Outpatient Elasticities				**Outpatient Hospital Selection Rates**		
	Distance	Quality	N	Rural	Urban	Research
All Outpatient	-1.32***	0.09***	138,685	0.7	0.12	0.18
Diagnosis						
Circulatory	-0.74**	0.15**	13,118	0.6	0.11	0.29
Digestive	-1.69**	0.11**	52,327	0.74	0.12	0.14

*Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level

Table 3.5: Simulated Impacts of Closing the 15% Least Used Rural Hospitals on the Distance to and Quality of Hospital Choice

Estimates are in changes relative to the baseline hospital choice							
Patient Set	Patients who:	Average	Probability	Expected Distance (miles)	Expected Quality	% Distance	% Quality
Outpatient	chose any hospital	Rural Average	-0.037	1.125	-0.072	32.3%	-3.9%
		Urban Average	0.017	0.327	0.053	5.5%	9.0%
		Specialty Average	0.021	0.702	0.105	15.1%	28.1%
		All	0.000	2.155	0.086	15.3%	3.1%
	chose rural hospitals	Rural Average	-0.017	0.617	-0.029	18.6%	-1.4%
	chose closed hospitals	Closed Average	-0.194	6.979	-0.358	278.5%	-28.0%
Inpatient	chose any hospital	Rural Average	-0.025	0.825	-0.070	25.5%	-3.9%
		Urban Average	0.016	0.382	0.057	6.4%	9.4%
		Specialty Average	0.009	0.319	0.031	6.2%	6.1%
		All	0.000	1.526	0.018	10.6%	0.6%
	chose rural hospitals	Rural Average	-0.011	0.429	-0.031	13.7%	-1.6%
	chose closed hospitals	Closed Average	-0.154	5.947	-0.429	224.0%	-28.4%
Emergency	chose any hospital	Rural Average	-0.025	0.542	-0.070	14.8%	-3.7%
		Urban Average	0.016	0.394	0.058	8.7%	11.1%
		Specialty Average	0.009	0.306	0.035	7.5%	7.0%
		Total Average	0.000	1.242	0.022	10.1%	0.8%
	chose rural hospitals	Rural Average	-0.011	0.308	-0.038	8.5%	-1.8%
	chose closed hospitals	Closed Average	-0.193	5.413	-0.660	177.9%	-37.2%
Urgent	chose any hospital	Rural Average	-0.027	0.634	-0.086	20.4%	-4.4%
		Urban Average	0.020	0.422	0.067	8.7%	12.8%
		Specialty Average	0.007	0.258	0.024	6.9%	6.5%
		Total Average	0.000	1.315	0.004	11.2%	0.2%
	chose rural hospitals	Rural Average	-0.008	0.228	-0.027	7.8%	-1.3%
	chose closed hospitals	Closed Average	-0.171	5.027	-0.593	124.4%	-32.7%
Elective	chose any hospital	Rural Average	-0.020	1.149	-0.045	38.2%	-2.8%
		Urban Average	0.011	0.289	0.040	3.5%	5.3%
		Specialty Average	0.009	0.306	0.028	3.8%	4.2%
		Total Average	0.000	1.744	0.023	9.1%	0.8%
	chose rural hospitals	Rural Average	-0.012	0.769	-0.019	25.8%	-1.1%
	chose closed hospitals	Closed Average	-0.099	6.519	-0.159	342.3%	-13.4%

Table 3.6: Simulated Impacts of Closing the 25% Lowest Quality Rural Hospitals on the Distance to and Quality of Hospital Choice

Estimates are in changes relative to the baseline hospital choice							
Patient Set	**Patients who:**	**Average**	**Probability**	**Expected Distance (miles)**	**Expected Quality**	**% Distance**	**% Quality**
Outpatient	**chose any hospital**	Rural Average	-0.020	2.006	0.152	57.7%	8.3%
		Urban Average	0.007	0.224	0.013	3.8%	2.2%
		Specialty Average	0.013	0.583	0.093	12.5%	24.9%
		All	0.000	2.813	0.258	19.9%	9.2%
	chose rural hospitals	Rural Average	-0.019	2.274	0.168	68.5%	8.4%
	chose closed hospitals	Closed Average	-0.083	9.536	0.683	294.8%	59.7%
Inpatient	**chose any hospital**	Rural Average	-0.017	2.164	0.156	67.0%	8.7%
		Urban Average	0.010	0.408	0.036	6.9%	5.9%
		Specialty Average	0.008	0.369	0.030	7.1%	5.8%
		All	0.000	2.941	0.222	20.5%	7.6%
	chose rural hospitals	Rural Average	-0.018	2.449	0.169	78.5%	8.5%
	chose closed hospitals	Closed Average	-0.072	9.677	0.668	316.3%	58.8%
Emergency	**chose any hospital**	Rural Average	-0.012	2.021	0.187	55.3%	9.9%
		Urban Average	0.007	0.299	0.025	6.6%	4.8%
		Specialty Average	0.005	0.267	0.023	6.5%	4.6%
		Total Average	0.000	2.587	0.235	21.0%	8.1%
	chose rural hospitals	Rural Average	-0.012	2.254	0.208	62.0%	9.9%
	chose closed hospitals	Closed Average	-0.051	9.868	0.912	270.4%	76.7%
Urgent	**chose any hospital**	Rural Average	-0.025	2.701	0.165	86.8%	8.5%
		Urban Average	0.013	0.550	0.045	11.3%	8.6%
		Specialty Average	0.011	0.537	0.042	14.2%	11.4%
		Total Average	0.000	3.788	0.252	32.3%	8.9%
	chose rural hospitals	Rural Average	-0.028	3.145	0.181	107.3%	8.6%
	chose closed hospitals	Closed Average	-0.094	10.746	0.618	368.9%	49.3%
Elective	**chose any hospital**	Rural Average	-0.011	1.932	0.136	64.3%	8.5%
		Urban Average	0.007	0.269	0.027	3.3%	3.6%
		Specialty Average	0.004	0.208	0.016	2.6%	2.5%
		Total Average	0.000	2.409	0.180	12.6%	6.0%
	chose rural hospitals	Rural Average	-0.012	2.108	0.139	70.6%	7.8%
	chose closed hospitals	Closed Average	-0.050	8.832	0.583	307.8%	59.3%

References

American Hospital Association. “The Opportunities and Challenges for Rural Hospitals in an Era of Health Reform.” *Trendwatch* (April 2011): 3.

Anderson, Michael L. 2008. “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” *Journal of the American Statistical Association* 103 (484): 1481–95. https://doi.org/10.1198/016214508000000841.

Andenaes, Johannes. 1974. *Punishment and Deterrence.* Ann Arbor: University of Michigan Press.

Anwar, Shamena, and Hanming Fang. 2006. “An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence.” *American Economic Review* 96 (1): 127–51. https://doi.org/10.1257/000282806776157579.

Auxier, Richard C. 2020. “What Police Spending Data Can (and Cannot) Explain amid Calls to Defund the Police.” Urban Institute. June 9, 2020. https://www.urban.org/urban-wire/what-police-spending-data-can-and-cannot-explain-amid-calls-defund-police.

Basu, Jayasree. “Severity of Illness, Race, and Choice of Local versus Distant Hospitals among the Elderly.” *Journal of Health Care for the Poor and Underserved* 16, no. 2 (May 2005): 391–405.

Basu, Jayasree, and James Cooper. “Out-of-Area Travel From Rural and Urban Counties: A Study of Ambulatory Care Sensitive Hospitalizations for New York State Residents.” *The Journal of Rural Health* 16, no. 2 (March 1, 2000): 129–38.

Battaglini, Marco, Rebecca B Morton, and Eleonora Patacchini. 2020. “Social Groups and the Effectiveness of Protests.” Working Paper 26757. National Bureau of Economic Research. https://doi.org/10.3386/w26757.

Beccaria, Cesare. 1764. *On Crimes and Punishments.* Oxford: Clarendon Press, 1957.

Bialik, Carl. 2015a. "An Ex-Cop Keeps The Country's Best Data Set On Police Misconduct." FiveThirtyEight (blog). April 22, 2015. https://fivethirtyeight.com/features/an-ex-cop-keeps-the-countrys-best-data-set-on-police-misconduct/.

———. 2015b. "Scare Headlines Exaggerated The U.S. Crime Wave." FiveThirtyEight, September 11, 2015. https://fivethirtyeight.com/features/scare-headlines-exaggerated-the-u-s-crime-wave/.

Bies, Katherine. 2017. "Let the Sunshine In: Illuminating the Powerful Role Police Unions Play in Shielding Officer Misconduct." *Stanford Law & Policy Review* 28 (109): 110–49.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169–217.

Bentham, Jeremy. 1789. *An Introduction to the Principles of Morals and Legislation*. Oxford: Clarendon Press.

Berk, Richard, and John MacDonald. 2010. "Policing the Homeless: An Evaluation of Efforts to Reduce Homeless-Related Crime." *Criminology and Public Policy* 9 (4): 813–40.

Blumstein, Alfred. 1995. "Youth Violence, Guns, and the Illicit-Drug Industry." *Journal of Criminal Law and Criminology* 86 (1): 10–36.

Braga, Anthony A. 2001. "The Effects of Hot Spots Policing on Crime." *Annals of the American Academy of Political and Social Science* 578 (1): 104–25.

———. 2005. "Hot Spots Policing and Crime Prevention: A Systematic Review of Randomized Controlled Trials." *Journal of Experimental Criminology* 1 (3): 317–42.

———. 2008. *Problem-Oriented Policing and Crime Prevention*, Second edition. Monsey, NY: Willow Tree Press.

Braga, Anthony A., and Brenda J. Bond. 2008. "Policing Crime and Disorder Hot Spots: A Randomized Controlled Trial." *Criminology* 46 (3): 577–607.

Braga, Anthony A., David M. Kennedy, Elin J. Waring, and Anne Morrison Piehl. 2001. "Problem-Oriented Policing, Deterrence, and Youth Violence: An Evaluation of Boston's Operation Ceasefire." *Journal of Research in Crime and Delinquency* 38 (3): 195–225.

Braga, A. A., Pierce, G. L., McDevitt, J., Bond, B. J., and Cronin, S. 2008. "The strategic prevention of gun violence among gang-involved offenders." *Justice Quarterly*, 25 (1), 132–162.

Braga, Anthony A., and David l. Weisburd. 2012. "The Effects of 'Pulling Levers' Focused Deterrence Strategies on Crime." *Campbell Systematic Reviews* 6.

Braga, Anthony A., David L. Weisburd, Elin J. Waring, Lorraine Green Mazerolle, William Spelman, and Francis Gajewski. 1999. "Problem-Oriented Policing in Violent Crime Places: A Randomized Controlled Experiment." *Criminology* 37 (3): 541–80.

Bronstein, J. M., and M. A. Morrisey. "Bypassing Rural Hospitals for Obstetrics Care." *Journal of Health Politics, Policy and Law* 16, no. 1 (1991): 87–118.

Buchmueller, Thomas C., Mireille Jacobson, and Cheryl Wold. 2006. "How far to the hospital?: The effect of hospital closures on access to care." *Journal of Health Economics* 25(4): 740-761.

Buczko, W. "Bypassing of Local Hospitals by Rural Medicare Beneficiaries." *The Journal of Rural Health* 10, no. 4 (1994): 237–46.

Caetano, Gregorio, and Vikram Maheshri. 2014. "Identifying Dynamic Spillovers in Criminal Behavior." Unpublished

Caeti, Tory J. 1999. "Houston's Targeted Beat Program: A Quasi-experimental Test of Police Patrol Strategies." Ph.D. Dissertation, Sam Houston State University.

Cameron, Samuel. 1988. "The Economics of Crime Deterrence: A Survey of Theory and Evidence." *Kyklos* 41 (2): 301–23.

Capalbo, Susan M., and Christine N. Heggem. "Innovations in the Delivery of Health Care Services to Rural Communities: Telemedicine and Limited-Service Hospitals." *Rural Development Perspectives* 14, no. 3 (1999): 8–13.

Chalfin, Aaron. 2015. "Economic Costs of Crime." In *The Encyclopedia of Crime & Punishment*, 1–12. Wiley. https://doi.org/10.1002/9781118519639.wbecpx193.

Chalfin, Aaron, and Justin McCrary. 2017. "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature* 55 (1):5–48. https://doi.org/10.1257/jel.20141147.

———. 2018. "Are U.S. Cities Underpoliced? Theory and Evidence." *Review of Economics and Statistics*, March.

Chandra, Amitabh, Amy Finkelstein, Adam Sacarny, and Chad Syverson. "Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector." *American Economic Review* 106, no. 8 (August 2016): 2110–44.

Chandrasekher, Andrea Cann. 2016. “The Effect of Police Slowdowns on Crime.” *American Law and Economics Review*, September. https://doi.org/10.1093/aler/ahw008.

Cheng, Cheng, and Wei Long. 2018. “The Spillover Effects of Highly Publicized Police-Related Deaths on Policing and Crime: Evidence from Large US Cities.” Working Paper, July.

Cloninger, Dale O. 1991. “Lethal Police Response as a Crime Deterrent: 57-City Study Suggests a Decrease in Certain Crimes." *American Journal of Economics and Sociology* 50 (1): 59–69.

Cohen, Jacqueline, and Jens Ludwig. 2003. “Policing Crime Guns.” In *Evaluating Gun Policy: Effects on Crime and Violence*, edited by Jens Ludwig and Philip J. Cook, 217–50. Washington, DC: Brookings Institution Press.

Corman, Hope, and Naci Mocan. 2005. “Carrots, Sticks, and Broken Windows.” *Journal of Law and Economics* 48 (1): 235–66.

Corsaro, Nicholas, Rod K. Brunson, and Edmund F. McGarrell. 2010. “Evaluating a Policing Strategy Intended to Disrupt an Illicit Street-Level Drug Market.” *Evaluation Review* 34 (6): 513–48.

Corsaro, Nicholas, and Edmund McGarrell. 2010. “Reducing Homicide Risk in Indianapolis between 1997 and 2000.” *Journal of Urban Health* 87 (5): 851–64.

Corsaro, Nicholas, Eleazer D. Hunt, Natalie Kroovand Hipple, and Edmund McGarrell. 2012. “The Impact of Drug Market Pulling Levers Policing on Neighborhood Violence: The Impact of Drug Market Pulling Levers Policing on Neighborhood Violence.” *Criminology and Public Policy* 11 (2): 167–99.

Davey, Monica, and Mitch Smith. 2015. “Murder Rates Rising Sharply in Many U.S. Cities.” The New York Times, August 31, 2015, sec. U.S. https://www.nytimes.com/2015/09/01/us/murder-rates-rising-sharply-in-many-us-cities.html.

DeAngelo, Gregory, and Benjamin Hansen. 2014. “Life and Death in the Fast Lane: Police Enforcement and Traffic Fatalities.” *American Economic Journal: Economic Policy* 6 (2): 231–57. https://doi.org/10.1257/pol.6.2.231.

Desmond, Matthew, Andrew Papachristos, and David Kirk. 2016. “Police Violence and Citizen Crime Reporting in the Black Community.” *American Sociological Review* 81 (5):857–76.

Desmond, Matthew, Andrew Papachristos, and David Kirk. 2020. “Evidence of the Effect of Police Violence on Citizen Crime Reporting.” *American Sociological Review* 85 (1):

184–90. https://doi.org/10.1177/0003122419895979.

Devi, Tanaya, and Roland Fryer. 2020. "Policing the Police: The Impact of 'Pattern-or-Practice' Investigations on Crime." Working Paper 27324. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w27324.

Di Tella, Rafael, and Ernesto Schargrodsky. 2004. "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack." *American Economic Review* 94 (1):115–33. https://doi.org/10.1257/000282804322970733.

Donohue, John J., Daniel E. Ho, and Patrick Leahy. 2013. "Do Police Reduce Crime? A Reexamination of a Natural Experiment." In *Empirical Legal Analysis: Assessing the Performance of Legal Institutions*, edited by Yun-chien Chang, 125–44. London and New York: Taylor and Francis, Routledge.

Donohue, John, and Steven Levitt. 2001. "The Impact of Race on Policing and Arrests." *Journal of Law and Economics* 44 (2): 367–94. https://doi.org/10.1086/322810.

Draca, Mirko, Stephen Machin, and Robert Witt. 2011. "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks." *American Economic Review* 101 (5):2157–81. https://doi.org/10.1257/aer.101.5.2157.

Dranove, D., W. D. White, and L. Wu. "Segmentation in Local Hospital Markets." *Medical Care* 31, no. 1 (January 1993): 52–64.

Durlauf, Steven N., and Daniel S. Nagin. 2011. "Imprisonment and Crime: Can Both Be Reduced?" *Criminology and Public Policy* 10 (1): 13–54.

Eck, John E., and Edward R. Maguire. 2000. "Have Changes in Policing Reduced Violent Crime? An Assessment of the Evidence." In *The Crime Drop in America*, edited by Alfred Blumstein and Joel Wallman, 207–65. Cambridge and New York: Cambridge University Press.

Eck, John E., and William Spelman. 1987. "Problem solving: Problem oriented policing in Newport news." Washington, DC: Police Executive Research Forum.

Elephrame. 2018. "2017 Report on the Black Lives Matter Movement." Elephrame. January 2, 2018. Link.

Engel, Robin S., Marie Skubak Tillyer, and Nicholas Corsaro. 2013. "Reducing Gang Violence Using Focused Deterrence: Evaluating the Cincinnati Initiative to Reduce Violence (CIRV)." *Justice Quarterly* 30 (3): 403–39.

Evans, William N., and Emily G. Owens. 2007. "COPS and Crime." *Journal of Public Economics* 91 (1): 181–201. https://doi.org/10.1016/j.jpubeco.2006.05.014.

"Fatal Encounters." 2005–2016. Fatal Encounters. Accessed November 6, 2014. Link.

Federal Bureau of Investigation. 2004. "Uniform Crime Reporting Handbook." 2004483104. Department of Justice. https://lccn.loc.gov/2004483104.

Federal Bureau of Investigation. 2016. Latest Crime Statistics Released (2015). Story.

Federal Bureau of Investigation. 2016. Uniform Crime Reporting Program Data: Offenses Known and Clearances by Arrest, 2005–2015. DOI: 10.3886/ICPSR36122.v1.

Federal Bureau of Investigation. 2017. FBI Releases Preliminary Semiannual Crime Statistics for 2016. Press Release.

Friedersdorf, Conor. 2015. "Black Lives Matter Takes Aim at Police-Union Contracts." The Atlantic. December 7, 2015. https://www.theatlantic.com/politics/archive/2015/12/black-lives-matter-takes-aim-at-police-union-contracts/418530/.

Friedman, Barry, and Maria Ponomarenko. 2015. "Democratic Policing." *NYU Law Review* 90 (6). https://www.nyulawreview.org/issues/volume-90-number-6/democratic-policing/.

Fryer, Roland G. 2019. "An Empirical Analysis of Racial Differences in Police Use of Force." *Journal of Political Economy* 127 (3): 1210–61. https://doi.org/10.1086/701423.

Gallup, Inc. 2016. "Before July Shootings, Blacks Divided on Police Behavior." Gallup.com. July 1, 2016. http://www.gallup.com/poll/193652/july-shootings-blacks-divided-police-behavior.aspx.

Gardiner, Dustin. 2019. "California Senate Approves Strict Police Use-of-Force Bill, Citing Police Killings of Black People - SFChronicle.com." July 9, 2019. Link.

Goldsteen, R. L., D. J. Falcone, R. W. Broyles, K. Goldsteen, and B. J. Reilly. "Local Factors Affecting the Tendency to Bypass Local Hospitals for Inpatient Mental Health Care: An Exploratory Analysis." *The Journal of Rural Health* 10, no. 2 (1994): 89–97.

Goldstein, Herman. 1979. "Improving policing: A problem oriented approach." *Crime and Delinquency* 24:236–58.

Goncalves, Felipe, and Steven Mello. 2017. "A Few Bad Apples? Racial Bias in Policing." Working Paper. Link.

Grunwald, Ben, and John Rappaport. 2020. "The Wandering Officer." *Yale Law Journal* 129 (6): 1600–1945.

Harcourt, Bernard E., and Jens Ludwig. 2006. "Broken Windows: New Evidence from New York City and a Five-City Social Experiment." *University of Chicago Law Review* 73: 271–320.

Hausman, David, and Dorothy Kronick. 2019. "Policing Police." SSRN Scholarly Paper ID 3192908. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3192908.

Hawkins, Derek. 2016. "'Ferguson Effect'? Savagely Beaten Cop Didn't Draw Gun for Fear of Media Uproar, Says Chicago Police Chief." Washington Post, October 7, 2016. https://www.washingtonpost.com/news/morning-mix/wp/2016/10/07/ferguson-effect-savagely-beaten-cop-didnt-draw-gun-for-fear-of-media-uproar-says-chicago-police-chief/.

Heaton, Paul. 2010. "Understanding the Effects of Antiprofiling Policies." *Journal of Law and Economics* 53 (1): 29–64. https://doi.org/10.1086/649645.

Ho, Kate, and Ariel Pakes. "Hospital Choices, Hospital Prices, and Financial Incentives to Physicians." *American Economic Review* 104, no. 12 (December 2014): 3841–84.

Hogan, C. "Patterns of Travel for Rural Individuals Hospitalized in New York State: Relationships between Distance, Destination, and Case Mix." *The Journal of Rural Health* 4, no. 2 (July 1988): 29–41.

Hope, Timothy J. 1994. "Problem-Oriented Policing and Drug-Market Locations: Three Case Studies." *Crime Prevention Studies* 2 (1): 5–32.

Hughes, David. 2020. "Opinion | I'm a Black Police Officer. Here's How to Change the System." *The New York Times*, July 16, 2020, sec. Opinion. https://www.nytimes.com/2020/07/16/opinion/police-funding-defund.html.

Iowa Hospital Association. "Hospital and Health System Characteristics." (2004).

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers." *American Economic Review* 83 (4): 685–709.

Kalven, Jamie. 2015. "Chicago Police Say They Killed a Black Teen in Self-Defense. Then Why Did They Shoot Him 16 Times?" Slate Magazine. February 11, 2015. Link.

Kaplan, Jacob. 2018. Uniform Crime Reporting (UCR) Program Data: Arrests by Age, Sex, and Race, 1974-2016. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2018-12-29. Link.

Kaplan, Jacob. 2019. Uniform Crime Reporting Program Data: Offenses Known and Clearances by Arrest, 1960-2017. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-02-10. Link.

Kelling, George L., and William J. Bratton. 2015. "Why We Need Broken Windows Policing." *City Journal* Winter.

Kelling, George, Tony Pate, Duane Dieckman, and Charles Brown. 1974. "The Kansas City preventive patrol experiment: Technical report." Washington, DC: Police Foundation.

Kelling, George L., and William H. Sousa, Jr. 2001. "Do Police Matter? An Analysis of the Impact of New York City's Policy Reforms." Manhattan Institute Center for Civic Innovation Civic Report 22.

Kelly, Kimbriell, Wesley Lowery, and Steven Rich. 2017. "Police Chiefs Are Often Forced to Put Officers Fired for Misconduct Back on the Streets." Washington Post. August 3, 2017. https://www.washingtonpost.com/graphics/2017/investigations/police-fired-rehired/.

Kennedy, David M., Anthony A. Braga, Anne M. Piehl, and Elin J. Waring. 2001. "Reducing Gun Violence: The Boston Gun Project's Operation Ceasefire." US Department of Justice National Institute of Justice Research Report 188741.

Kirk, David, and Andrew Papachristos. 2011. "Cultural Mechanisms and the Persistence of Neighborhood Violence." *American Journal of Sociology* 116 (4): 1190–1233.

Klick, Jonathan, and Alexander Tabarrok. 2005. "Using Terror Alert Levels to Estimate the Effect of Police on Crime." *Journal of Law and Economics* 48 (1):267–79.

Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (1): 203–29. https://doi.org/10.1086/318603.

Kozhimannil, Katy B., Peiyin Hung, Carrie Henning-Smith, Michelle M. Casey, and Shailendra Prasad. 2018. "Association between loss of hospital-based obstetric services and birth outcomes in rural counties in the United States." *Journal of the American Medical Association* 319(12): 1239-1247.

Krishnan, Kaushik, Justin McCrary, and Deepak Premkumar. 2017. "Understanding Civilian Deaths at Police Hands: Evidence from Crowdsourced Data." Working Paper.

Kubrin, Charis E., Steven F. Messner, Glenn Deane, Kelly McGeever, and Thomas D. Stucky. 2010. "Proactive Policing and Robbery Rates across U.S. Cities." *Criminology* 48 (1): 57–97.

Lacoe, Johanna, and Jillian Stein. 2018. "Exploring the Policy Implications of High-Profile Police Violence." Criminology & Public Policy 17 (4): 859–63. Link.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. 2018. "School Finance Reform and the Distribution of Student Achievement." *American Economic Journal: Applied Economics.* 10 (2): 1–26. https://doi.org/10.1257/app.20160567.

Lee, David S., and Justin McCrary. 2017. "The Deterrence Effect of Prison: Dynamic Theory and Evidence." In *Regression Discontinuity Designs*, 38:73–146. Advances in Econometrics 38. Emerald Publishing Limited. Link.

Legewie, Joscha. 2016. "Racial Profiling and Use of Force in Police Stops: How Local Events Trigger Periods of Increased Discrimination." *American Journal of Sociology* 122 (2):379–424.

Legewie, Joscha, and Jeffrey Fagan. 2016. "Group Threat, Police Officer Diversity and the Deadly Use of Police Force." SSRN Scholarly Paper ID 2778692. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=2778692.

Levitt, Steven D. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime." *American Economic Review* 87 (3): 270–90.

Levitt, Steven D. 1998. "Why Do Increased Arrest Rates Appear to Reduce Crime: Deterrence, Incapacitation, or Measurement Error?" *Economic Inquiry* 36 (3):353–72. https://doi.org/10.1111/j.1465-7295.1998.tb01720.x.

Levitt, Steven D. 2002. "Using Electoral Cycles in Police Hiring to Estimate the Effects of Police on Crime: Reply." *American Economic Review* 92 (4): 1244–50.

Levitt, Steven D., and Thomas J. Miles. 2006. "Economic Contributions to the Understanding of Crime." *Annual Review of Law and Social Science* 2: 147–64.

Lin, Ge, Diane E. Allan, and Margaret J. Penning. "Examining Distance Effects on Hospitalizations Using GIS: A Study of Three Health Regions in British Columbia, Canada." *Environment and Planning A* 34, no. 11 (November 1, 2002): 2037–53.

Liu, Jiexin Jason, Gail R. Bellamy, and Melissa McCormick. "Patient Bypass Behavior and Critical Access Hospitals: Implications for Patient Retention." *The Journal of Rural Health* 23, no. 1 (2007): 17–24.

Lin, Ming-Jen. 2009. "More Police, Less Crime: Evidence from US State Data." *International Review of Law and Economics* 29 (2): 73–80.

Long, Wei. 2019. “How Does Oversight Affect Police? Evidence from the Police Misconduct Reform." *Journal of Economic Behavior & Organization* 168 (December): 94–118. https://doi.org/10.1016/j.jebo.2019.10.003.

Lum, Cynthia, Christopher S. Koper, and Cody W. Telep. 2011. “The Evidence-Based Policing Matrix.” *Journal of Experimental Criminology* 7 (1):3–26. https://doi.org/10.1007/s11292-010-9108-2.

Mac Donald, Heather. 2015. “The New Nationwide Crime Wave.” *Wall Street Journal*, May 29, 2015, sec. Opinion. https://www.wsj.com/articles/the-new-nationwide-crime-wave-1432938425.

MacDonald, John M. 2002. “The Effectiveness of Community Policing in Reducing Urban Violence.” *Crime and Delinquency* 48 (4): 592–618.

MacDonald, John M., Jonathan Klick, and Ben Grunwald. 2016. “The Effect of Private Police on Crime: Evidence from a Geographic Regression Discontinuity Design.” *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 179 (3):831–46. https://doi.org/10.1111/rssa.12142.

Maltz, Michael D. 2006. “Analysis of Missingness in UCR Crime Data.” 215343. Criminal Justice Research Ctr, Ohio State University. Link.

Manski, Charles F., and Daniel S. Nagin. 2017. “Assessing Benefits, Costs, and Disparate Racial Impacts of Confrontational Proactive Policing." *Proceedings of the National Academy of Sciences* 114 (35): 9308–13. https://doi.org/10.1073/pnas.1707215114.

Mas, Alexandre. 2006. “Pay, Reference Points, and Police Performance." *Quarterly Journal of Economics* 121 (3): 783–821. https://doi.org/10.1162/qjec.121.3.783.

Mastrorocco, Nicola, and Arianna Ornaghi. 2020. “Who Watches the Watchmen? Local News and Police Behavior in the United States.” Working Paper, June. Link.

McCrary, Justin. 2007. “The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police.” *American Economic Review* 97 (1):318–53.

McCrary, Justin, and Deepak Premkumar. 2019. “Why We Need Police." in *Cambridge Handbook on Policing in the United States.* Cambridge University Press.

McGarrell, Edmund F., Steven Chermak, Jeremy Wilson, and Nicholas Corsaro. 2006. “Reducing Homicide through a ‘Lever-Pulling’ Strategy.” *Justice Quarterly* 23 (2): 214–31.

McGarrell, Edmund F., Nicholas Corsaro, Natalie Kroovand Hipple, and Timothy S. Bynum. 2010. "Project Safe Neighborhoods and Violent Crime Trends in US Cities: Assessing Violent Crime Impact." *Journal of Quantitative Criminology* 26 (2): 165–90.

Mello, Steven. 2019. "More COPS, Less Crime." *Journal of Public Economics* 172 (April): 174–200. https://doi.org/10.1016/j.jpubeco.2018.12.003.

Morgan, Stephen, and Joel Pally. 2016. "Ferguson, Gray, and Davis: An Analysis of Recorded Crime Incidents and Arrests in Baltimore City, March 2010 through December 2015." Working Paper. Link.

Nagin, Daniel S. 1998. "Criminal Deterrence Research at the Outset of the Twenty-First Century." *Crime and Justice: A Review of Research* 23: 1–42.

Nagin, Daniel S. 2013. "Deterrence: A Review of the Evidence By a Criminologist for Economists." *Annual Review of Economics* 5: 83–105.

Nagin, Daniel S., and Cody W. Telep. 2017. "Procedural Justice and Legal Compliance." *Annual Review of Law and Social Science* 13 (1):5–28. https://doi.org/10.1146/annurev-lawsocsci-110316-113310.

National Institute of Justice. n.d. "Practice: Street-Level Drug Law Enforcement - CrimeSolutions.gov." Accessed December 14, 2017. https://www.crimesolutions.gov/PracticeDetails.aspx?ID=49.

Noblet, Ronald, and Urban Peace Institute. 2015. "Nationwide Crime Spike Has Law Enforcement Retooling Its Approach." All Things Considered, National Public Radio. http://www.npr.org/2015/07/01/418555852/nationwide-crime-spike-has-law-enforcement-retooling-their-approach.

Ouss, Aurélie, and John Rappaport. 2020. "Is Police Behavior Getting Worse? Data Selection and the Measurement of Policing Harms." *The Journal of Legal Studies* 49 (1): 153–98. https://doi.org/10.1086/708705.

Owens, Emily Greene. 2013. "COPS and Cuffs." In *Lessons from the Economics of Crime: What Reduces Offending?*, edited by Philip J. Cook, Stephen Machin, Olivier Marie, and Giovanni Mastrobuoni, 17–44. Cambridge, MA and London: MIT Press.

Owens EG, Weisburd D, Alpert G, Amendola KL. 2016. "Promoting Police Integrity Through Early Engagements and Procedural Justice in the Seattle Police Department." Washington, DC: Natl. Inst. Justice, US Dep. Justice.

Papachristos, Andrew V., Tracey L. Meares, and Jeffrey Fagan. 2007. "Attention Felons: Evaluating Project Safe Neighborhoods in Chicago." *Journal of Empirical Legal Studies*

4 (2):223–72. https://doi.org/10.1111/j.1740-1461.2007.00096.x.

Persico, Nicola. 2002. "Racial Profiling, Fairness, and Effectiveness of Policing." *American Economic Review* 92 (5): 1472–97. https://doi.org/10.1257/000282802762024593.

Polinsky, A. Mitchell, and Steven Shavell. 1999. "On the Disutility and Discounting of Imprisonment and the Theory of Deterrence." *Journal of Legal Studies* 28 (1): 1–16.

Prendergast, Canice. 2001. "Selection and Oversight in the Public Sector, With the Los Angeles Police Department as an Example." Working Paper 8664. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w8664.

Radcliff, Tiffany A., Michelle Brasure, Ira S. Moscovice, and Jeffrey T. Stensland. "Understanding Rural Hospital Bypass Behavior." *The Journal of Rural Health* 19, no. 3 (2003): 252–59.

Raphael, Steven, and Jens Ludwig. 2003. "Prison Sentence Enhancements: The Case of Project Exile." In *Evaluating Gun Policy: Effects on Crime and Violence*, edited by Jens Ludwig and Philip J. Cook, 251–76. Washington, DC: Brookings Institution Press.

Rieber, G. M., D. Benzie, and S. McMahon. "Why Patients Bypass Rural Health Care Centers." *Minnesota Medicine* 79, no. 6 (June 1996): 46–50.

Rivera, Roman, and Bocar Ba. 2019. "The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago." Working Paper.

Rosenfeld, Richard, Robert Fornango, and Eric Baumer. 2005. "Did Ceasefire, Compstat, and Exile Reduce Homicide?" *Criminology and Public Policy* 4 (3): 419–49.

Rosenfeld, Richard, Robert Fornango, and Andres F. Rengifo. 2007. "The Impact of Order-Maintenance Policing on New York City Homicide and Robbery Rates: 1988–2001." *Criminology* 45 (2): 355–84.

Rosenfeld, Richard, and Joel Wallman. 2019. "Did De-Policing Cause the Increase in Homicide Rates?" *Criminology & Public Policy* 18 (1): 51–75. Link.

Rozema, Kyle, and Max Schanzenbach. 2019. "Good Cop, Bad Cop: Using Civilian Allegations to Predict Police Misconduct." *American Economic Journal: Economic Policy* 11 (2): 225–68. https://doi.org/10.1257/pol.20160573.

Rushin, Stephen. 2017. "Police Union Contracts." *Duke Law Journal* 66 (6): 1191–1266.

Rushin, Stephen. 2019. "Police Disciplinary Appeals." *University of Pennsylvania Law Review* 167 (3): 545.

Samaha, A. 2017. "How Video Finally Proved That Cops Lie." Buzzfeed News, January 17, 2017.

Sampson, Robert J., and Jacqueline Cohen. 1988. "Deterrent Effects of the Police on Crime: A Replication and Theoretical Extension." *Law and Society Review* 22 (1): 163–90.

Sandler, Danielle, and Ryan Sandler. 2014."Multiple Event Studies in Public Finance and Labor Economics: A Simulation Study with Applications." *Journal of Economic and Social Measurement* 39 (1,2):31–57.

Schanzenbach, Max. 2015. "Union Contracts Key to Reducing Police Misconduct." Chicago Tribune. November 23, 2015. Link.

Schmidt, Michael S., and Matt Apuzzo. 2015. "F.B.I. Chief Links Scrutiny of Police With Rise in Violent Crime Index." The New York Times, October 2015. https://www.nytimes.com/2015/10/24/us/politics/fbi-chief-links-scrutiny-of-police-with-rise-in-violent-crime.html.

Schwartz, Joanna C. 2014. "Police Indemnification." *NYU Law Review*, June. https://papers.ssrn.com/abstract=2297534.

Sherman, Lawrence W. 1990. "Police Crackdowns: Initial and Residual Deterrence." *Crime and Justice* 12: 1–48.

Sherman, Lawrence W. 1995. "Hot Spots of Crime and Criminal Careers of Places." In *Crime and Place: Crime Prevention Studies*, Vol. 1, edited by John E. Eck and David Weisburd, 35–52. Monsey, NY: Willow Tree Press.

Sherman, Lawrence W., et al. 1995. "Deterrent Effects of Police Raids on Crack Houses: A Randomized, Controlled Experiment." *Justice Quarterly* 12 (4): 755–81.

Sherman, Lawrence W., Patrick R. Gartin, and Michael E. Buerger. 1989. "Hot Spots of Predatory Crime: Routine Activities and the Criminology of Place." *Criminology* 27 (1): 27–56.

Sherman, Lawrence W., and Dennis P. Rogan. 1995. "Effects of Gun Seizures on Gun Violence: 'Hot Spots' Patrol in Kansas City." *Justice Quarterly* 12 (4): 673–93.

Sherman, Lawrence W., and David Weisburd. 1995. "General Deterrent Effects of Police Patrol in Crime 'Hot Spots': A Randomized, Controlled Trial." *Justice Quarterly* 12 (4):

625–48.

Shi, Lan. 2009. “The Limit of Oversight in Policing: Evidence from the 2001 Cincinnati Riot.” *Journal of Public Economics* 93 (1): 99–113. Link.

Shjarback, John A., David C. Pyrooz, Scott E. Wolfe, and Scott H. Decker. 2017. “De-Policing and Crime in the Wake of Ferguson: Racialized Changes in the Quantity and Quality of Policing among Missouri Police Departments.” *Journal of Criminal Justice* 50 (Supplement C): 42–52. https://doi.org/10.1016/j.jcrimjus.2017.04.003.

Skogan, Wesley, and Kathleen Frydl, eds. 2004. *Fairness and Effectiveness in Policing: The Evidence.* Washington, DC: National Academies Press.

Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations.* New York: Modern Library, 1937.

Spelman, William, and Dale K. Brown. 1981. “Calling the Police: Citizen Reporting of Serious Crime.” Washington, DC: Police Executive Research Forum.

Sunshine, Jason, and Tom R. Tyler. 2003. “The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing." *Law & Society Review* 37 (3): 513–48. https://doi.org/10.1111/1540-5893.3703002.

Swaine, Jon. 2014. “Michael Brown Shooting: ‘They Killed Another Young Black Man in America.’ ” The Guardian, August 12, 2014, sec. Global. https://www.theguardian.com/world/2014/aug/12/ferguson-missouri-shooting-michael-brown-civil-rights-police-brutality.

Trump, Kris-Stella, Vanessa Williamson, and Katherine Levine Einstein. 2018. “Vol 16(2): Replication Data for: Black Lives Matter: Evidence That Police-Caused Deaths Predict Protest Activity,” May. https://doi.org/10.7910/DVN/L2GSK6.

Wallace, Danielle, Andrew V. Papachristos, Tracey Meares, and Jeffrey Fagan. 2016. “Desistance and Legitimacy: The Impact of Offender Notification Meetings on Recidivism among High Risk Offenders.” *Justice Quarterly* 33 (7):1237–64. https://doi.org/10.1080/07418825.2015.1081262.

Weisburd, David. 2005. “Hot Spots Policing Experiments and Criminal Justice Research: Lessons from the Field.” *ANNALS of the American Academy of Political and Social Science* 599 (1): 220–45.

Weisburd, David, Shawn Bushway, Cynthia Lum, and Sue-Ming Yang. 2004. “Trajectories of Crime at Places: A Longitudinal Study of Street Segments in the City of Seattle.”

Criminology 42 (2): 283–322.

Weisburd, David, and John E. Eck. 2004. "What Can Police Do to Reduce Crime, Disorder, and Fear?" *The Annals of the American Academy of Political and Social Science* 593:42–65.

Weisburd, David, and Lorraine Green. 1995. "Policing Drug Hot Spots: The Jersey City Drug Market Analysis Experiment." *Justice Quarterly* 12 (4): 711–35.

Weisburd, David, Lisa Maher, and Lawrence Sherman. 1992. "Contrasting Crime General and Crime Specific Theory: The Case of Hot Spots of Crime." In *New Directions in Criminological Theory: Advances in Criminological Theory: Volume 4*, edited by Freda Adler and William S. Laufer, 45–70. New Brunswick, NJ and London: Transaction Publishers.

Weisburd, David, Stephen D. Mastrofski, Ann Marie McNally, Rosann Greenspan, and James J. Willis. 2003. "Reforming to Preserve: Compstat and Strategic Problem Solving in American Policing." *Criminology and Public Policy* 2 (3): 421–56.

Weisburd, David, Nancy A. Morris, and Elizabeth R. Groff. 2009. "Hot Spots of Juvenile Crime: A Longitudinal Study of Arrest Incidents at Street Segments in Seattle, Washington." *Journal of Quantitative Criminology* 25: 443–67.

Weisburd, David, Cody W. Telep, Joshua C. Hinkle, and John E. Eck. 2010. "Is Problem-Oriented Policing Effective in Reducing Crime and Disorder? Findings from a Campbell Systematic Review." *Criminology and Public Policy* 9 (1): 139–72.

Weisburd, David, Laura A. Wyckoff, Justin Ready, John E. Eck, Joshua C. Hinkle, and Frank Gajewski. 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.

Wilson, James Q., and George L. Kelling. 1982. "Broken Windows: The Police and Neighborhood Safety." *Atlantic Monthly* 249 (3): 29–38.

Zimring, Franklin E., and Gordon Hawkins. 1973. *Deterrence: The Legal Threat in Crime Control*. Chicago and London: University of Chicago Press.

Zoorob, Michael. 2020. "Do Police Brutality Stories Reduce 911 Calls? Reassessing an Important Criminological Finding." *American Sociological Review* 85 (1): 176–83. https://doi.org/10.1177/0003122419895254.

Appendix A

Appendix: Intensified Scrutiny and Bureaucratic Effort: Evidence from Policing After High-Profile, Officer-Involved Fatalities

A.1 Appendix: Data

A.1.1 Inputting Rules for the Uniform Crime Report

The counts in the FBI data are based on the Hierarchy Rule, which states that for multiple-offense incidents the police department should only record the most serious offense/arrest, providing a ranking of Part I offenses and their severity (FBI, 2004). Naturally, this reporting suggests that the data is an undercount, but the National Incident-Based Reporting System—an FBI dataset that contains every reported criminal incident—suggests that for most offenses, around 90%, only one crime occurs. Thus, this undercount is not a grave concern, at least for the sample of jurisdictions that report to NIBRS. Additionally, offenses are distinct through the Separate of Time and Place rule. Even if a criminal commits numerous offenses in a short period of time, but they are perpetrated in different locations, then they must be recorded as separate incidents (FBI, 2004), providing assurance that the data is a close reflection of reality.

A.1.2 Cleaning the Uniform Crime Report

I extensively clean both the UCR and ASR datasets, and narrow the sample to departments that report at least six years to both datasets between 2005–2016. Relying on departments that report to both the UCR and the ASR provides some confidence in the nature of reporting quality. However, the annual FBI data has been panned for the lack of consistent reporting from police departments. In fact, the Department of Justice has previously provided grants for reports analyzing the patterns of "missingness," such as Maltz (2006). Additionally, I only keep departments report each month consistently across years, unlike certain departments who, for example, file only twice a year (i.e., input all of their crime and arrest statistics under July and December). Following Evans and Owens (2007) and Mello (2019), I fit a local polynomial function for both crime and arrests for each department and set all of values outside the 99.9% confidence interval as an outliers. The threshold was determined by visually inspecting a random sample of departments. These outliers are then set to missing. Additionally, observations are flagged as outliers if all the values in both violent and property crime (or arrests) are zero for the entire quarter and the population of the jurisdiction is above 5,000.

As pointed out in Chalfin and McCrary (2018), the population variable in the UCR and ASR jump discretely in census years for many jurisdictions. In order to alleviate that concern, I fit a local polynomial to smooth out the population variable at the census threshold and use the smoothed population variable as a control in all future analyses. Finally, for ASR, I use a "clean" subset of the data, whose sample inclusion is city police departments, who have a population greater than 10,000, and fewer than 9 outliers across the 12-year sample frame. In Evans and Owens (2007), they similarly use the city departments with a population threshold of 10,000 as determinants for sample inclusion, but they drop agencies if they have four or more outliers. I chose a higher threshold because I had additional methods detecting of outliers.

For the main results, I use the clean subset of the ASR data to ensure quality in the underlying data. Since this misreporting is less of an issue in the UCR data, I go through the same cleaning techniques to drop inconsistent reporting agencies and errant observations, but do not make the same "clean" sample cutoffs. Thus, the UCR analysis has more municipalities for both treatment and control, where the exact number can be found at the bottom of each of the event study graphs.

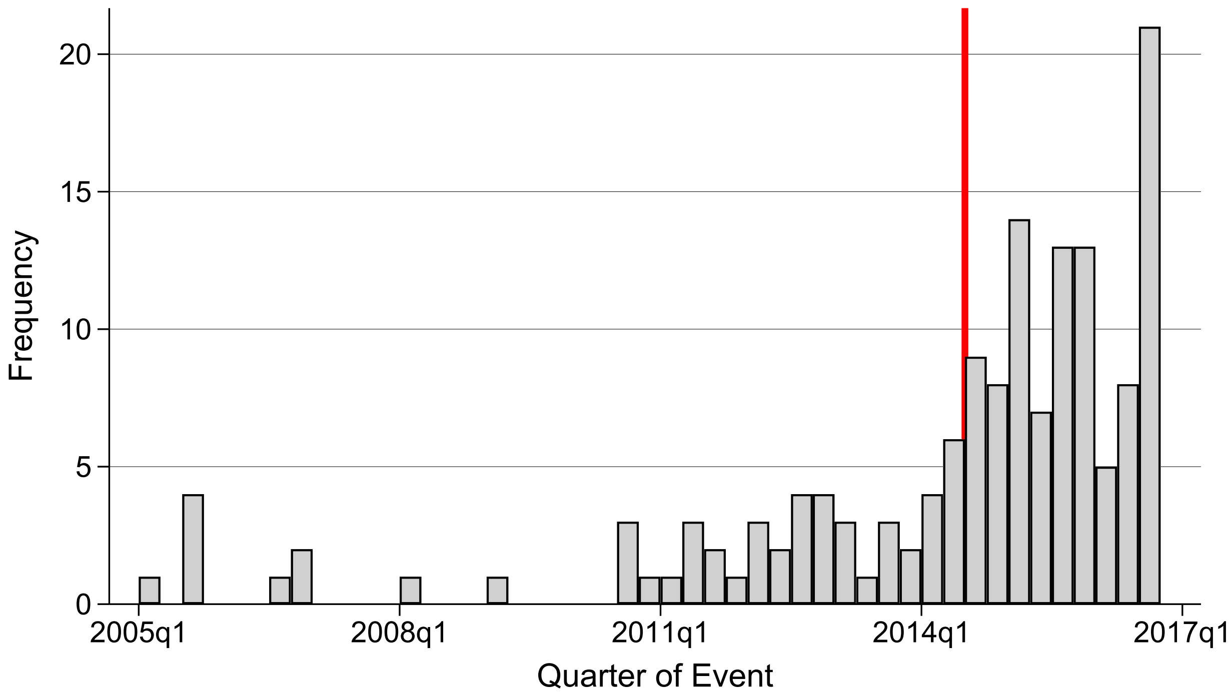
A.1.3 Scraping News Articles for Officer-Involved Fatalities

To measure how high-profile an officer-involved is, I scrape the number of news articles written about the incident, as determined by a search engine's news classification. Each search about the OF requires each article to contain the victim's full name, the city where the incident took place, and the word "police" or the word "killing." I read each case report on the incidents that have more than 1,000 articles to remove "falsely positive" high-profile occurrences. There are over 100 cases where I demote a fatality's news article number to zero if they are erroneously high-profile. Some examples of erroneous cases with inflated news hits are (1) articles about a different, more famous person with same name in the same city, (2) people participating in multiple crimes, often multiple homicides, which gave them notoriety, making the coverage not about the OF itself (e.g, one of the Boston Marathon bombers, Tamerlan Tsarnaev), or (3) victims engaging in shootouts with police. This is not an exhaustive list of reasons, but it provides the general overview for the type of OF that is be excluded from the event sample to study exogenous shocks of scrutiny on policing effort. Lowering the threshold, to 500 articles for example, results in many false positives and not as many protests in the affected jurisdictions, bolstering the argument for a higher threshold.

A.2 Appendix: Figures and Tables

A.2.1 Relevant Findings

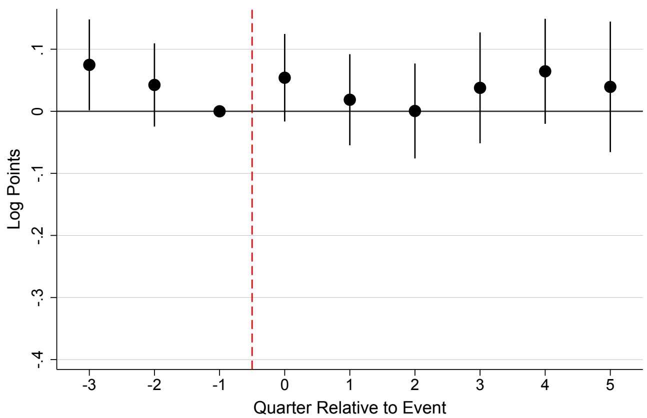
Figure A.1: Timing of High-Profile, Officer-Involved Fatalities (2005–2016)



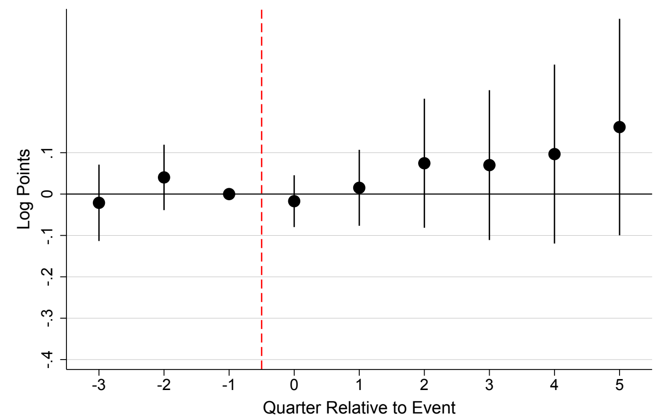
151 fatalities in total from 2005-2016. High profile is defined as ≥ 1,000 news articles. Red line depicts the death of Michael Brown in Ferguson, MO (August 2014). The map includes fatalities that will not be used in the main analysis sample, likely because the involved department poorly reported data to the FBI.

Figure A.2: Effect of Officer-Involved Fatality on Non-Low Marginal Benefit Arrests

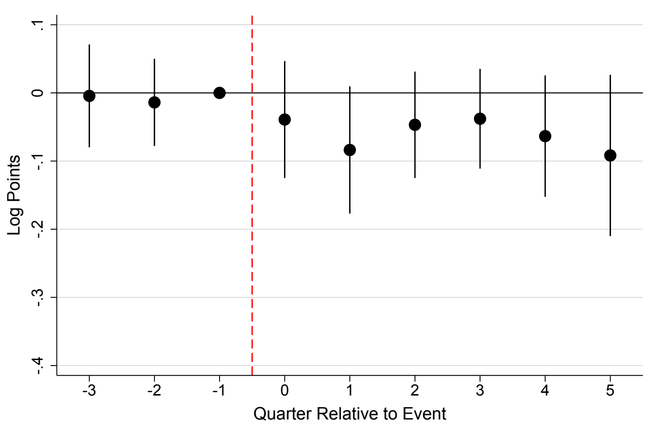
(a) Log(Robbery Arrests)



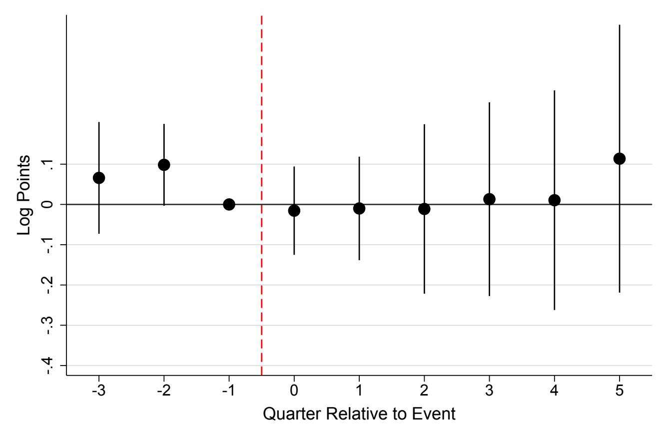
(b) Log(Motor Vehicle Theft Arrests)



(c) Log(Weapon Arrests)



(d) Log(Heroin/Cocaine Sale Arrests)



Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, violent and property crime, and their specific offense type if it's Part I. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Table A.1: Effect of the Highest Profile, Officer-Involved Fatality on Log(Violent Crime Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Violent Crime Arrests				
Treat*Post	0.002	0.008	-0.005	0.000
	(0.024)	(0.027)	(0.030)	(0.034)
Pretrend Test		0.012		0.010
		(0.025)		(0.037)
Black*Treat*Post			0.013	0.015
			(0.032)	(0.037)
Black Pretrend Test				0.003
				(0.038)
Panel B: Murder Arrests				
Treat*Post	-0.014	-0.020	-0.009	-0.014
	(0.017)	(0.021)	(0.024)	(0.031)
Pretrend Test		-0.013		-0.009
		(0.025)		(0.032)
Black*Treat*Post			-0.009	-0.013
			(0.038)	(0.048)
Black Pretrend Test				-0.007
				(0.053)
Panel C: Aggravated Assault Arrests				
Treat*Post	-0.010	0.003	-0.017	0.003
	(0.026)	(0.030)	(0.033)	(0.038)
Pretrend Test		0.026		0.040
		(0.024)		(0.034)
Black*Treat*Post			0.014	-0.000
			(0.034)	(0.037)
Black Pretrend Test				-0.028
				(0.039)
Observations	740,980	740,980	740,980	740,980
Number of Agencies	2,739	2,739	2,739	2,739

Coefficients from a double (DD) and triple difference (DDD) regression. Standard errors, clustered at the agency level, in parentheses. Uses only city police departments with less than 9 outliers and a population greater 10,000. Each regression controls for population, and violent and property crime, while Part I arrests control for their specific offense type. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the black dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Effect of the Highest Profile, Officer-Involved Fatality on Log(Property Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Property Crime Arrests				
Treat*Post	-0.071***	-0.055**	-0.050**	-0.024
	(0.021)	(0.022)	(0.025)	(0.029)
Pretrend Test		0.030*		0.051*
		(0.018)		(0.027)
Black*Treat*Post			-0.041*	-0.063*
			(0.025)	(0.033)
Black Pretrend Test				-0.042
				(0.035)
Panel B: Burglary Arrests				
Treat*Post	-0.052*	-0.068*	-0.070*	-0.073*
	(0.030)	(0.036)	(0.036)	(0.044)
Pretrend Test		-0.032		-0.006
		(0.030)		(0.045)
Black*Treat*Post			0.035	0.009
			(0.039)	(0.053)
Black Pretrend Test				-0.052
				(0.073)
Panel C: Theft Arrests				
Treat*Post	-0.092***	-0.078**	-0.062*	-0.043
	(0.031)	(0.033)	(0.034)	(0.039)
Pretrend Test		0.026		0.037
		(0.021)		(0.027)
Black*Treat*Post			-0.058*	-0.070**
			(0.031)	(0.035)
Black Pretrend Test				-0.022
				(0.034)
Observations	740,980	740,980	740,980	740,980
Number of Agencies	2,739	2,739	2,739	2,739

Coefficients from a double (DD) and triple difference (DDD) regression. Standard errors, clustered at the agency level, in parentheses. Uses only city police departments with less than 9 outliers and a population greater 10,000. Each regression controls for population, and violent and property crime, while Part I arrests control for their specific offense type. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the black dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Effect of the Highest Profile, Officer-Involved Fatality on Log(Low MB Arrests)

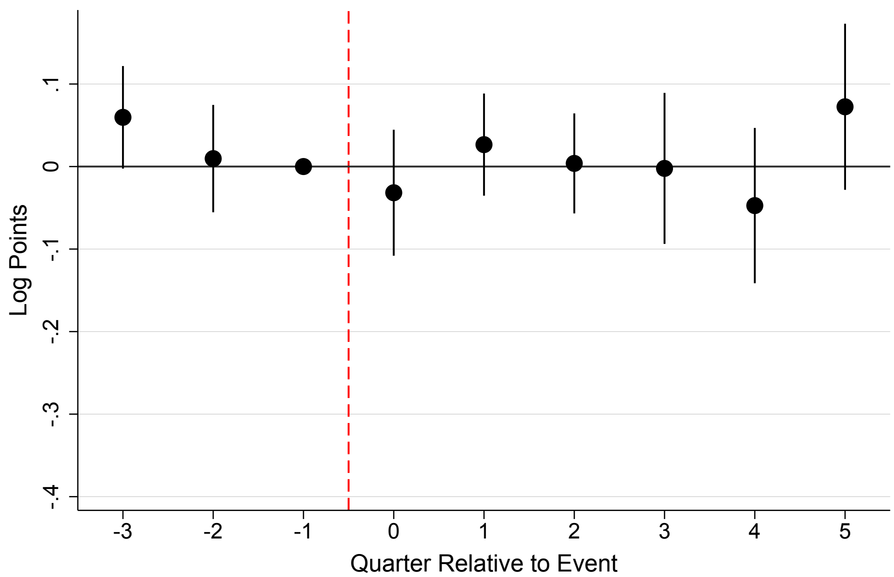
	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Low Marginal Benefit Arrests				
Treat*Post	-0.098**	-0.119**	-0.084	-0.106*
	(0.046)	(0.049)	(0.052)	(0.056)
Pretrend Test		-0.044*		-0.044
		(0.026)		(0.031)
Black*Treat*Post			-0.028	-0.027
			(0.031)	(0.034)
Black Pretrend Test				0.000
				(0.028)
Panel B: Disorderly Conduct Arrests				
Treat*Post	-0.147**	-0.164**	-0.157**	-0.171**
	(0.064)	(0.068)	(0.065)	(0.072)
Pretrend Test		-0.035		-0.029
		(0.037)		(0.044)
Black*Treat*Post			0.021	0.015
			(0.039)	(0.045)
Black Pretrend Test				-0.011
				(0.055)
Panel C: Marijuana Possession Arrests				
Treat*Post	-0.142***	-0.166**	-0.145**	-0.177**
	(0.054)	(0.065)	(0.057)	(0.070)
Pretrend Test		-0.047		-0.063
		(0.042)		(0.051)
Black*Treat*Post			0.006	0.022
			(0.053)	(0.056)
Black Pretrend Test				0.032
				(0.047)
Observations	741,156	741,156	741,156	741,156
Number of Agencies	2,739	2,739	2,739	2,739

Coefficients from a double (DD) and triple difference (DDD) regression. Standard errors, clustered at the agency level, in parentheses. Uses only city police departments with less than 9 outliers and a population greater 10,000. Each regression controls for population, and violent and property crime, while Part I arrests control for their specific offense type. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the black dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

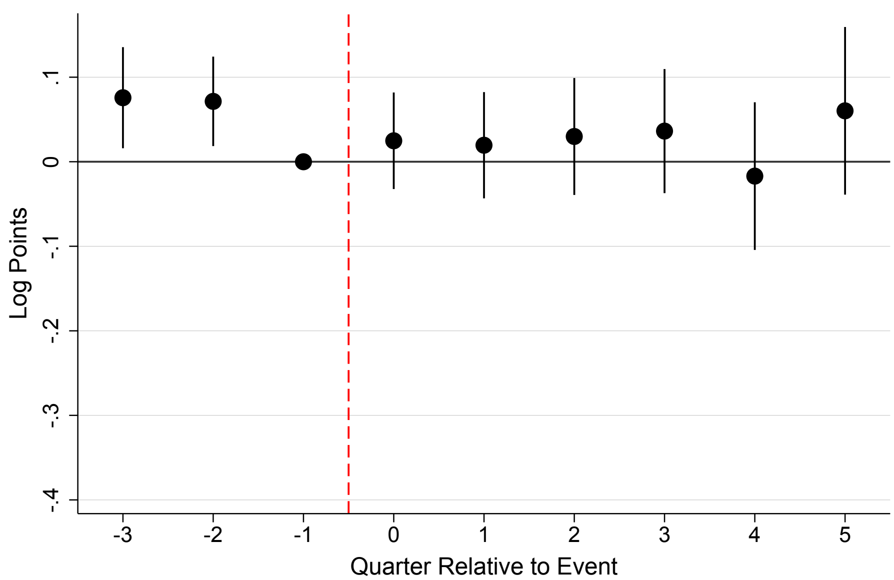
A.2.2 UCR Sample: Part I Offenses Known and Clearances by Arrests

Figure A.3: UCR Sample: Effect of Officer-Involved Fatality on Part I Arrests

(a) Log(Murder Arrests)



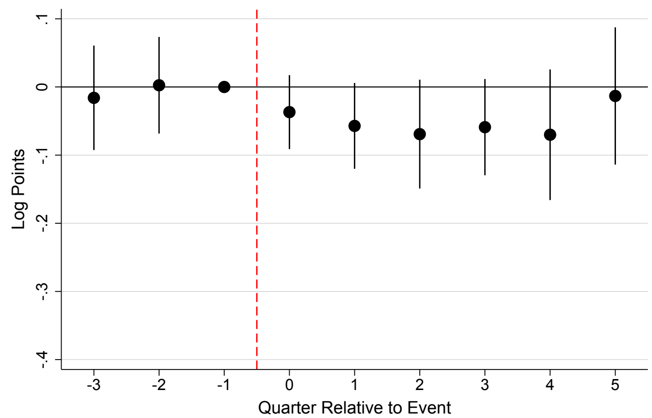
(b) Log(Aggravated Assault Arrests)



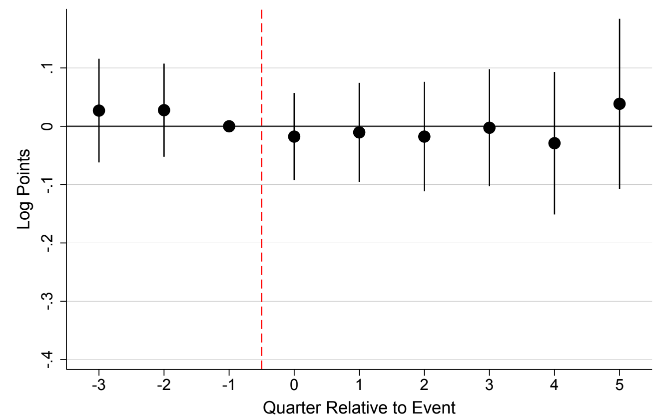
The event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and their specific offense type. All regressions have month-of-sample and department fixed effects. [1,017,449 observations; 77 treated; 9,176 control agencies]

Figure A.4: UCR Sample: Effect of Officer-Involved Fatality on Property Crime Arrests

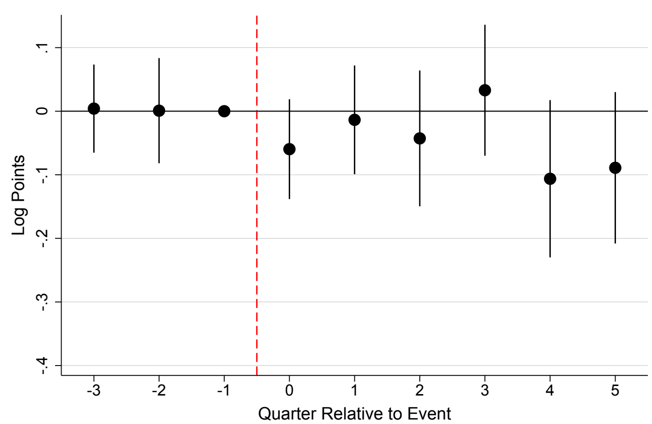
(a) Log(Property Crime Arrests)



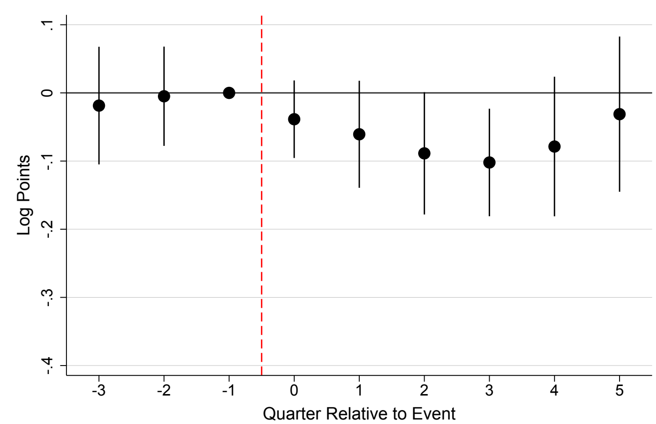
(b) Log(Motor Vehicle Theft Arrests)



(c) Log(Burglary Arrests)



(d) Log(Theft Arrests)

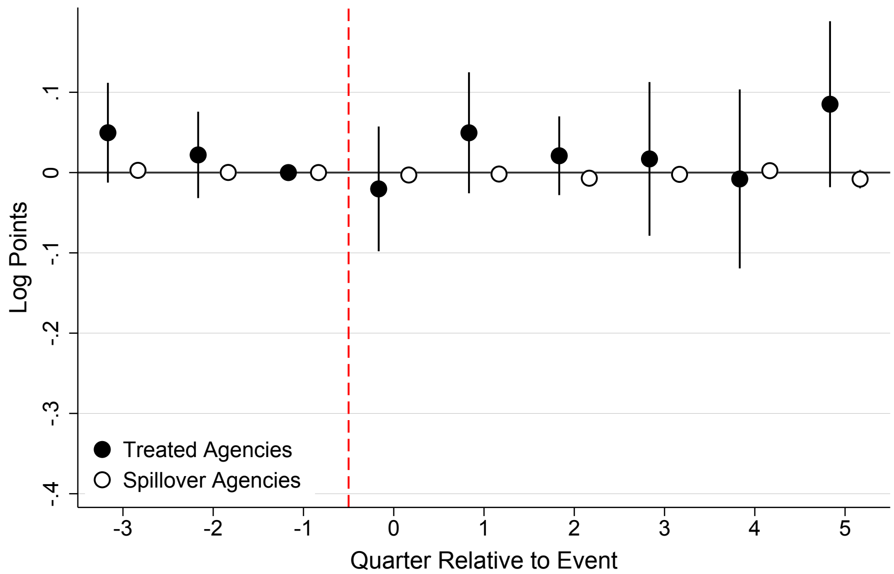


The event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and their specific offense type. All regressions have month-of-sample and department fixed effects. [1,017,449 observations; 77 treated; 9,176 control agencies]

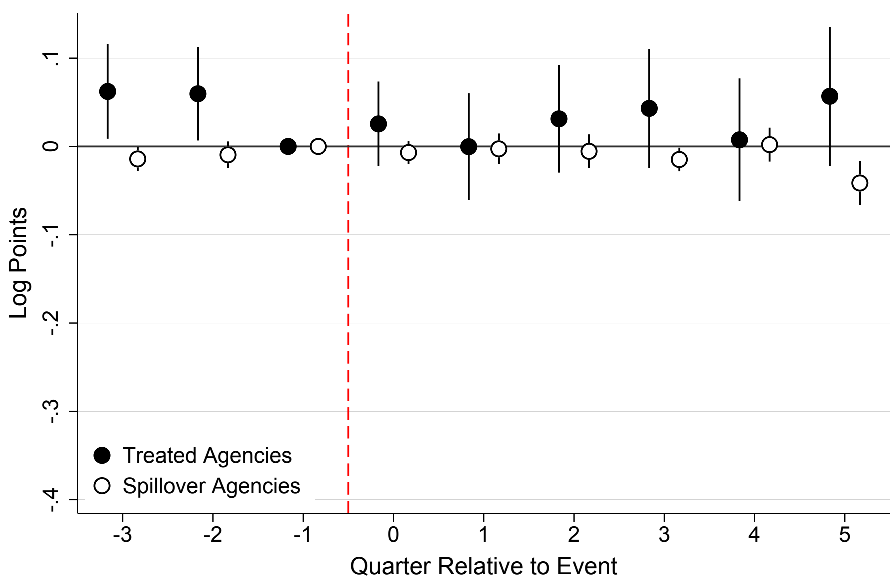
A.2.3 UCR Spillover Analysis of Adjacent Municipalities in the Same County

Figure A.5: UCR Spillover Analysis: Effect of Officer-Involved Fatality on Part I Arrests

(a) Log(Murder Arrests)



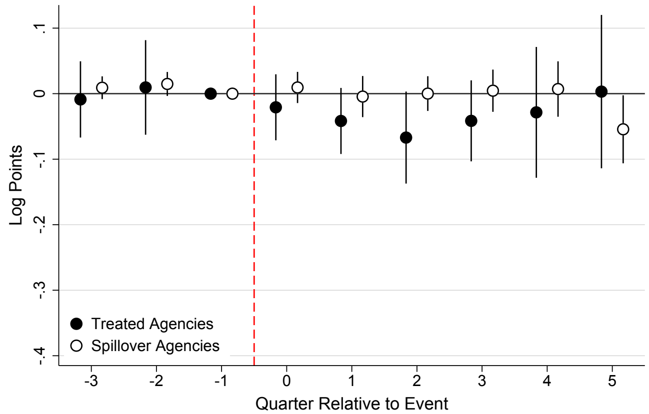
(b) Log(Aggravated Assault Arrests)



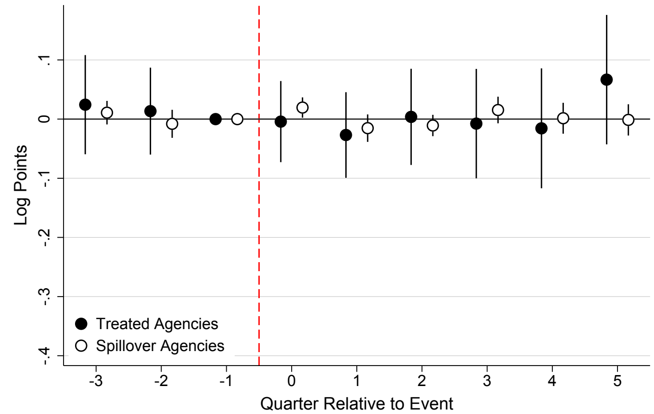
The event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and their specific offense type. All regressions have month-of-sample and department fixed effects. [1,017,335 observations; 74 treated; 725 spillover; 8,453 control agencies]

Figure A.6: UCR Spillover Analysis: Effect of Officer-Involved Fatality on Property Crime Arrests

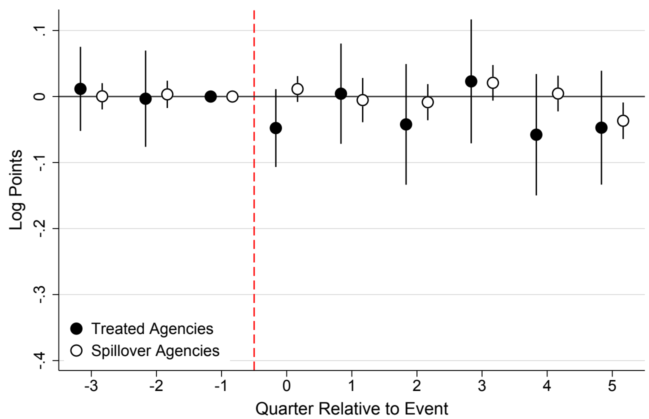
(a) Log(Property Crime Arrests)



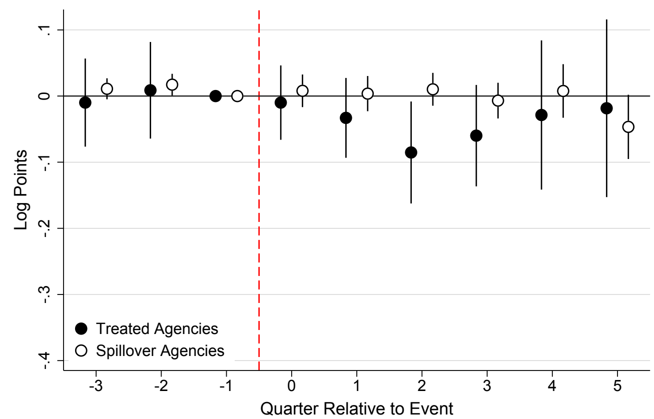
(b) Log(Motor Vehicle Theft Arrests)



(c) Log(Burglary Arrests)



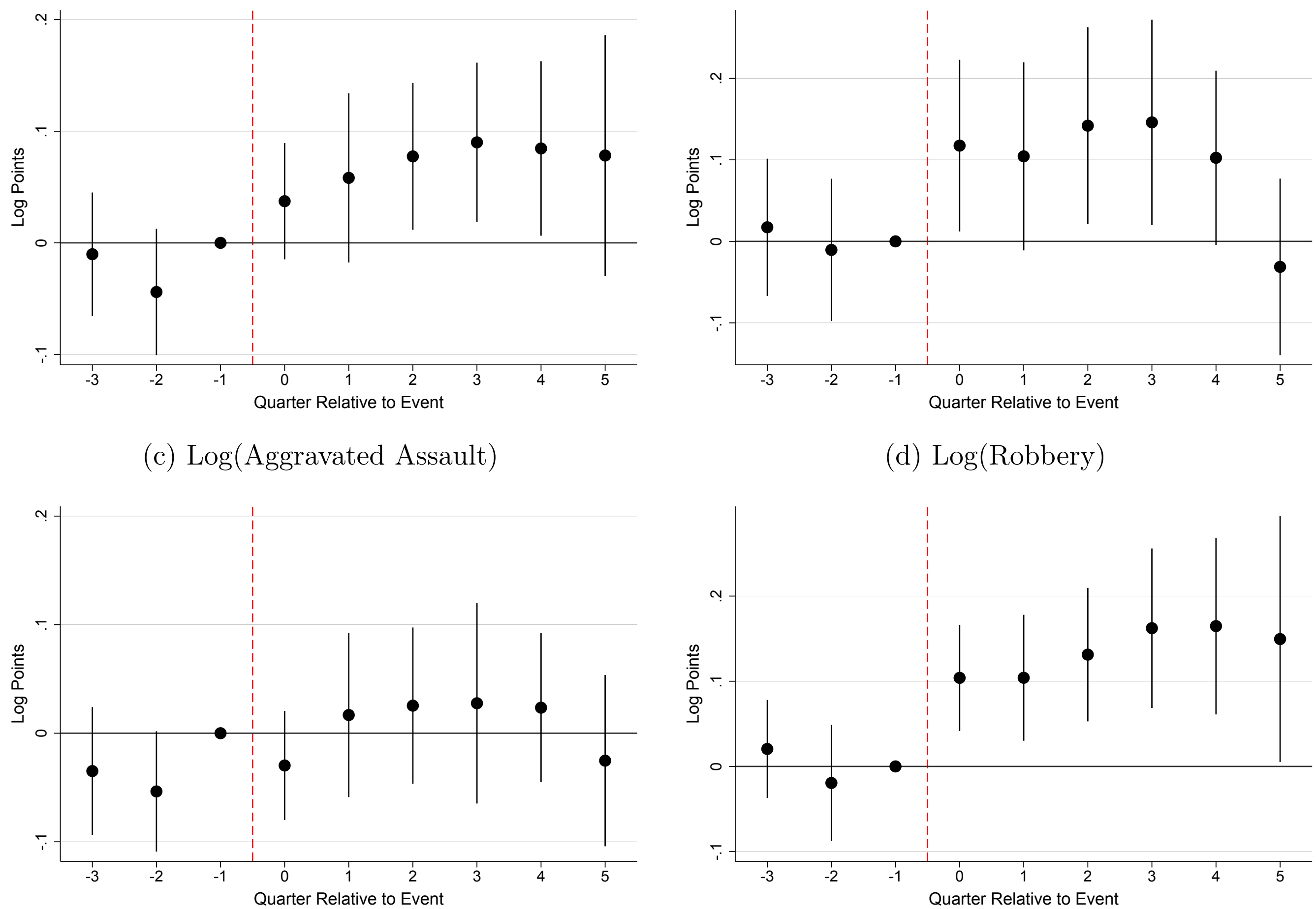
(d) Log(Theft Arrests)



The event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and their specific offense type. All regressions have month-of-sample and department fixed effects. [1,017,335 observations; 74 treated; 725 spillover; 8,453 control agencies]

A.2.4 ASR Sample: Effect of High-Profile, Officer-Involved Fatality on Part I Crimes

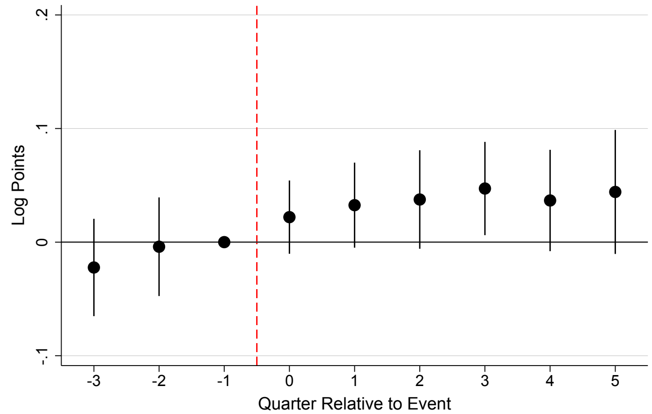
Figure A.7: Crime Analysis: Effect of Officer-Involved Fatality on Violent Crime



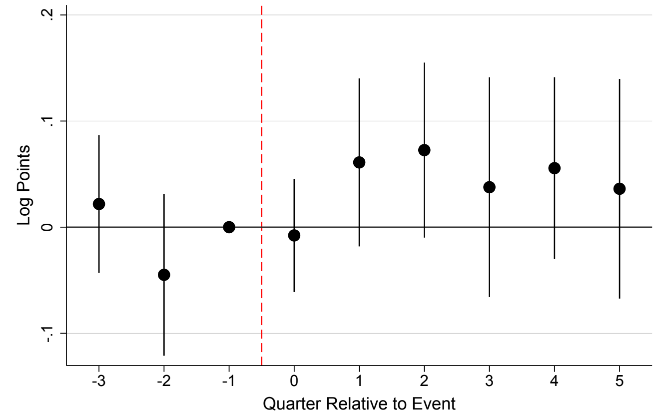
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in crime after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Figure A.8: Crime Analysis: Effect of Officer-Involved Fatality on Property Crime

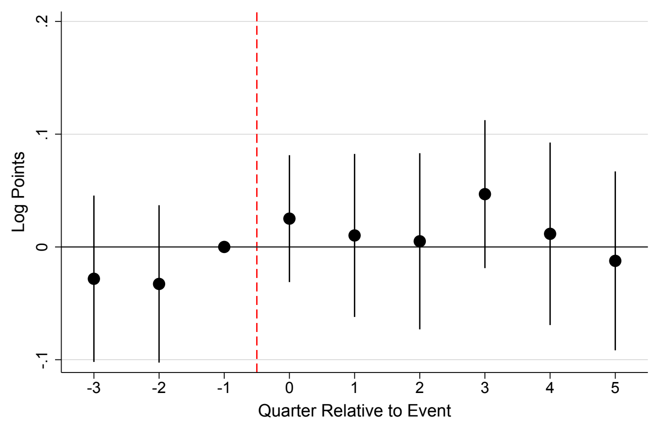
(a) Log(Property Crime)



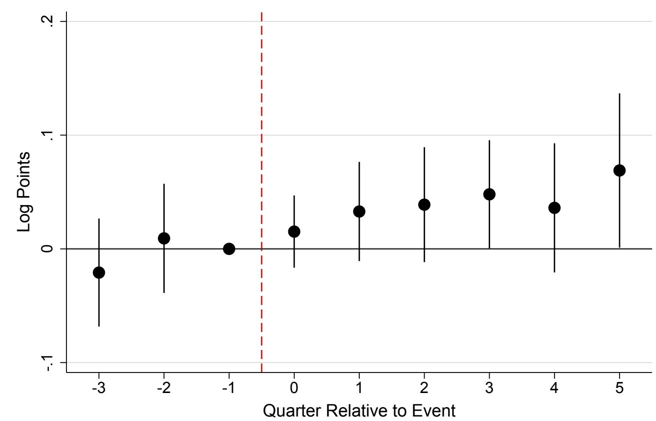
(b) Log(Motor Vehicle Theft)



(c) Log(Burglary)



(d) Log(Theft)

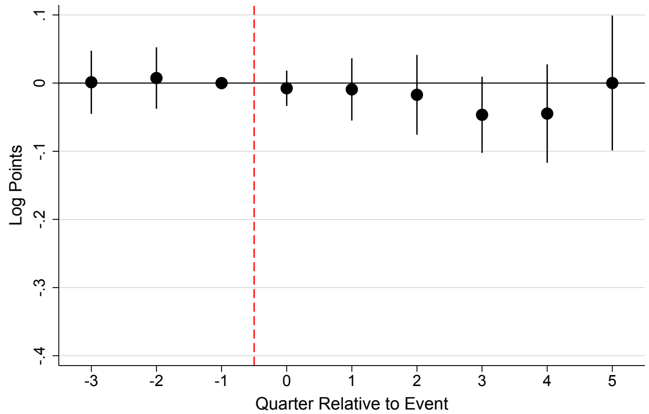


Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in crime after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

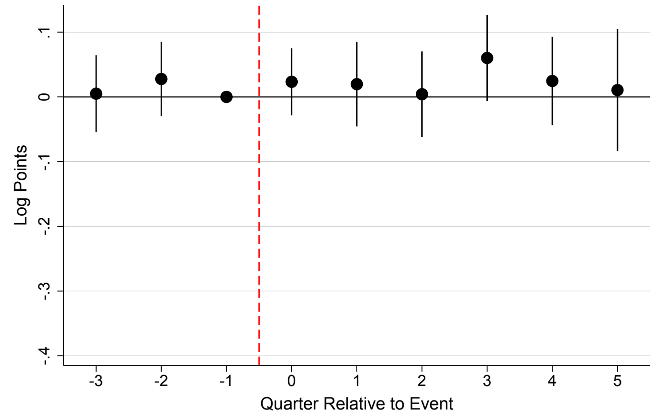
A.2.5 ASR Analysis without Crime Controls: Effect of High-Profile, Officer-Involved Fatality on Arrests

Figure A.9: Effect of High-Profile, Officer-Involved Fatality on Arrests without Crime Controls

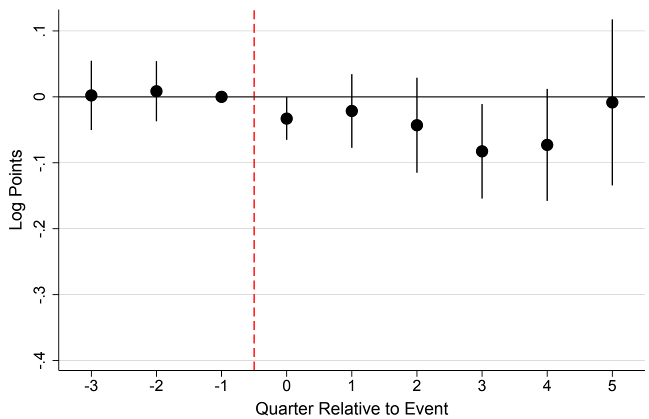
(a) Log(All Part I Arrests)



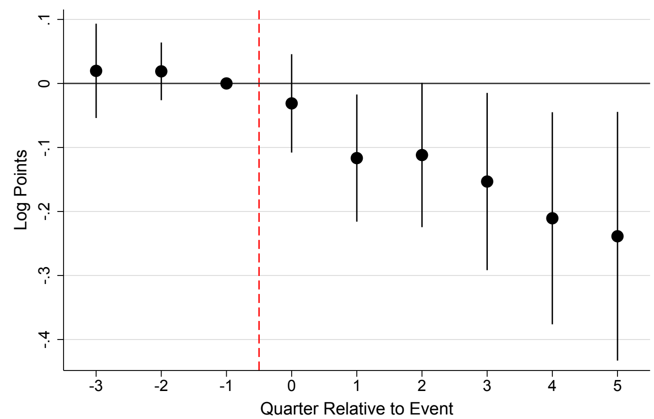
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)



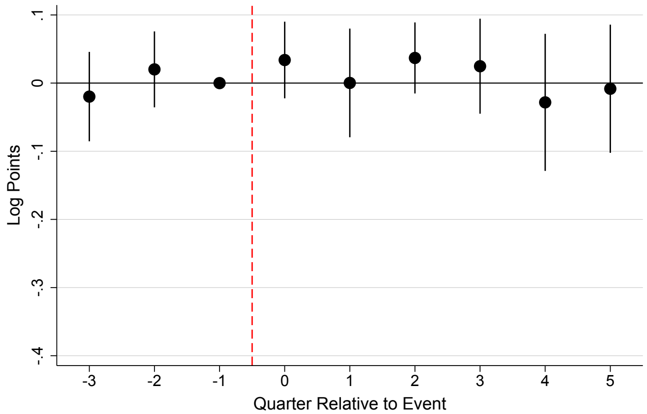
(d) Log(Part II Low MB Arrests)



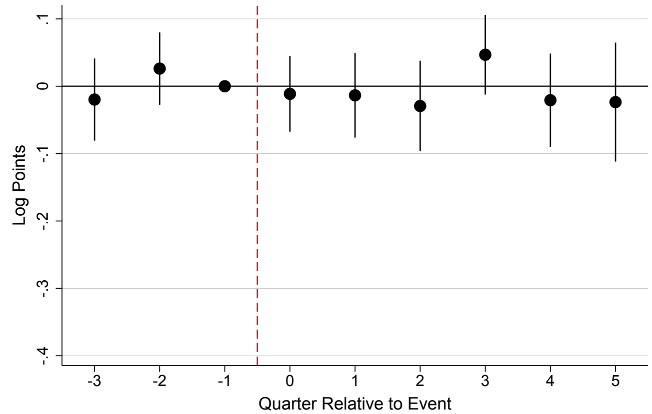
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Figure A.10: Effect of High-Profile, Officer-Involved Fatality on Violent and Property Crime Arrests without Crime Controls

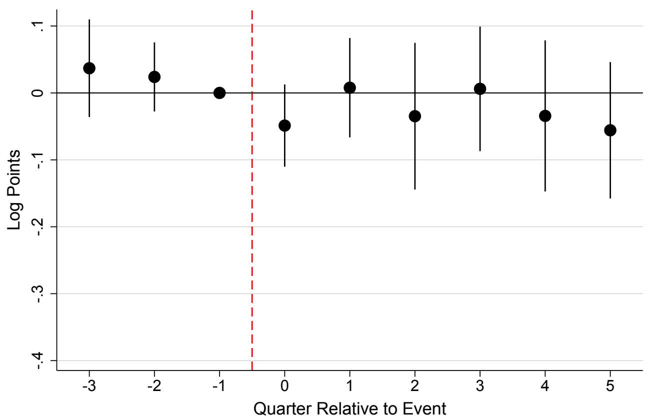
(a) Log(Murder Arrests)



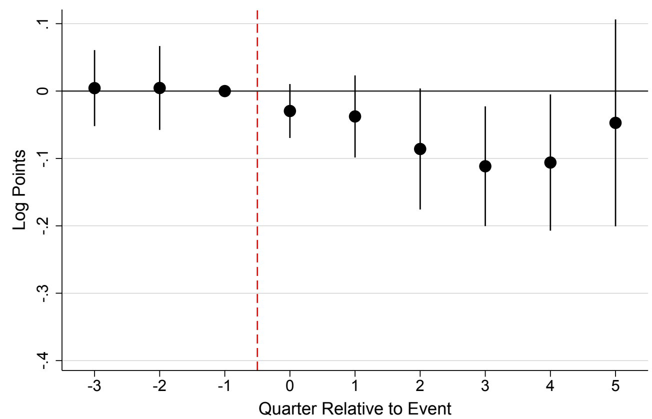
(b) Log(Aggravated Assault Arrests)



(c) Log(Burglary Arrests)



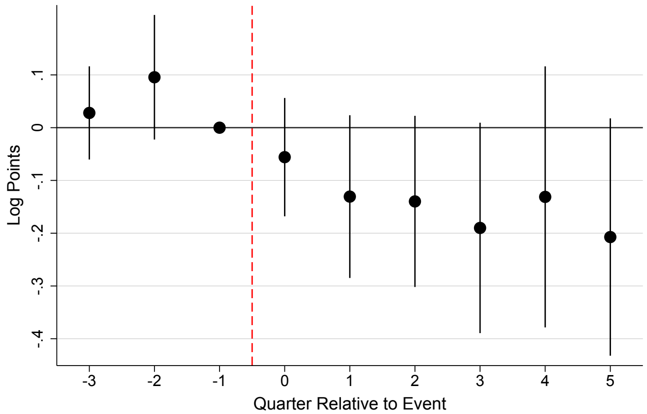
(d) Log(Theft Arrests)



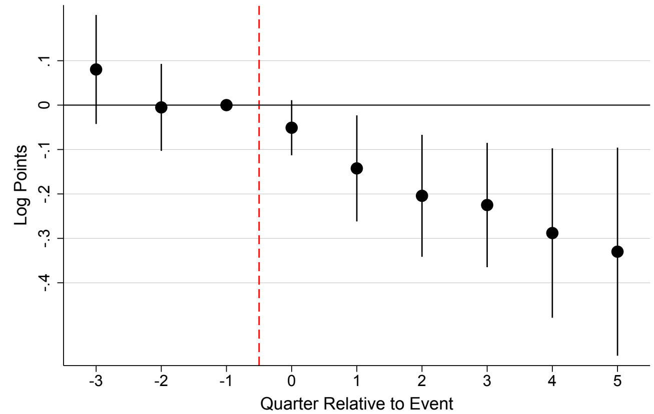
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,910 observations; 53 treated; 2,687 control agencies]

Figure A.11: Effect of High-Profile, Officer-Involved Fatality on Low MB Arrests without Crime Controls

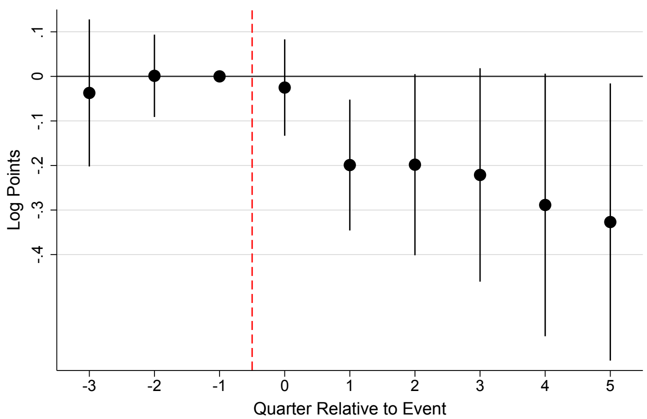
(a) Log(Disorderly Conduct Arrests)



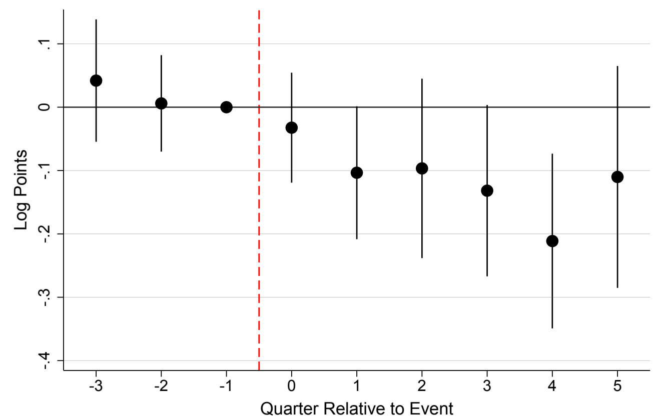
(b) Log(Marijuana Possession Arrests)



(c) Log(Liquor Law Violation Arrests)



(d) Log(Marijuana Sale Arrests)



Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [599,088 observations; 53 treated; 2,687 control agencies]