Public Scrutiny, Police Behavior, and Crime Consequences: Evidence from High-Profile Police Killings*

Deepak Premkumar†

March 17, 2022

Abstract

After a spate of protests touched off by high-profile incidents of police use of force, there has been a renewed focus on whether public scrutiny shapes policing behavior, otherwise known as the *Ferguson Effect*. This question has gained additional urgency as the country grapples with increases in murders. This paper provides the first national analysis showing that after police killings that generate significant public attention and scrutiny, officers reduce effort and crime increases. The effects differ by offense type: Reduced police effort yields persistently fewer arrests for low-level offenses (e.g., marijuana possession) but limited changes in arrests for violent or more serious property crimes. I show that decreased interaction with civilians through police stops may be driving the results. However, the increase in *offending* is driven by murders and robberies, imposing significant crime costs on affected municipalities. The effects only occur after there is broad community awareness of the incident. These findings are robust to numerous changes in empirical specification, transformations of the dependent variable, and varying levels of fixed effects that control for changes in state law and treatment spillovers. I also present evidence that suggests these effects are not driven by a pattern-or-practice investigation or a court-mandated monitoring agreement. To distinguish between the potential effects that may simultaneously impact arrest levels following a high-profile police killing, I develop a theoretical model that provides empirically testable predictions for each mechanism. I find that the reduction in low-level arrests corroborates public scrutiny as the causal channel. Finally, I provide evidence that the increase in offending is driven by *both* a response to the reduction in policing effort and a reaction to the police killing itself, suggesting that measures to reduce use of force should be prioritized.

JEL Codes: K42, H76, J15, D73, H41

Keywords: Ferguson Effect, police killings, public scrutiny, police effort, crime, arrests, stops

*This paper was previously circulated with the title "Intensified Scrutiny and Bureaucratic Effort: Evidence from Policing and Crime After High-Profile, Officer-Involved Fatalities." 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, Steve Raphael, and Yotam Shem-Tov for their incisive comments and careful read-throughs. I thank Abhay Aneja, Bocar Ba, Aaron Chalfin, Andrea Headley, Julien Lafortune, Magnus Lofstrom, Justin McCrary, Steve Mello, Matt Pecenco, Jeff Perloff, Evan Rose and Joel Wallman for their comments, as well as the seminar participants at UC Berkeley Empirical Legal Studies (BELS), UC Berkeley Ag. and Resource Economics, Cornerstone Research, US Treasury, Government Accountability Office, Federal Trade Commission, Yale's Public Health Modeling Unit, Public Policy Institute of California, Professor Jennifer Doleac's Virtual Crime Economics seminar, APPAM, and the 2021 ASSA meeting. 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 BELS Fellowship.

†Public Policy Institute of California. premkumar@ppic.org.

1 Introduction

How can law enforcement agencies provide safety without compromising trust or imposing undue burden on communities? In aftermath of the murder of George Floyd, calls to defund the police have popularized one answer to this question—that they cannot. As US experienced the largest recorded increase in murders in 2020 (Asher, 2021), the answer to this question becomes even more pressing. A consistent finding in previous economics literature has found that US cities are underpoliced, in the sense that the additional hiring of police officers would result in welfare gains from reduced crime costs (Evans and Owens, 2007; Chalfin and McCrary, 2018; Mello, 2019). Typically, these economic studies do not account for the social costs of policing (McCrary and Premkumar, 2019), broader costs imposed upon on a community from an increase in police presence or the confrontational nature of a policing strategy. Recent data from national newspapers and crowdsourced datasets such as Fatal Encounters reveal that officer-involved fatalities may be a significant source of social cost, finding that 7–9% of homicides (i.e., all cases where someone causes the death of another) in the US involve the police.[1]

Police departments may not internalize these social costs: First, these social costs may not be widely understood across the community if they are concentrated in a subset of the population (Edwards, Lee, and Esposito, 2019; Ang, 2021), 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 social costs 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 public scrutiny through protests, media coverage, or additional monitoring (e.g., cop-watch groups)—to raise awareness and advocate for their preferred policing allocation (Madestam et al., 2013; Battaglini, Morton, and Patacchini, 2020; Ouss and Rappaport, 2020). In the wake of highly-publicized police killings, some commentators have suggested that police effort has been reduced due to additional public scrutiny from the community and media, widely known as the *Ferguson Effect* (Comey, 2015; Mac Donald, 2015). That discussion has recently heightened as the murder of George Floyd unleashed one of the largest waves of civil protest and unrest seen in the US (Buchanan, Bui, and Patel, 2020).

This paper provides the first national analysis of a key question of policy interest: how does a high-profile police killing in a jurisdiction affect local policing behavior and crime? Exploiting the plausibly exogenous timing of these fatalities to credibly identify causal effects, I explore how police officer effort is affected by public scrutiny, as measured by community

[1]This statistic, of course, does not speak to whether or how many of these cases were "justified," but regardless, they are still relevant when calculating social costs.

awareness, media coverage, and protests. To determine what constitutes *high profile*, I create a novel dataset by merging crowdsourced information on officer-involved fatalities (OFs) with web-scraped data on the amount of media coverage and local search engine patterns of each incident.[2] I combine this with the FBI Uniform Crime Report, which collects arrest and crime data from most police departments. I use monthly arrest counts, disaggregated by race and crime type, to analyze policing effort in 2,740 police departments, 52 of which experience at least one high-profile, officer-involved fatality between 2005–2016. I then 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—adjusting for changes in population, temporal variation in arrests nationally, time invariant department- and jurisdiction-level characteristics, and county-specific linear trends. Finally, leveraging a similar empirical design, I test to see if these OFs induce any changes in offending behavior.

To include the variation arising from departments that have multiple high-profile OFs, I utilize an event study regression design that allows for more than one event per observational unit. I subset the events to a sample of officer-involved fatalities that generate 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 not plausible for the vast majority of OFs, given how common and low profile they are. 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 complement the national analysis with a detailed analysis of how killing of Laquan McDonald and subsequent cover-up by the Chicago Police Department affected the arrests, stops, and crimes in Chicago. McDonald's death is one of the highest-profile police killings in the sample, and the incident is notable since there is a year gap between his death and broad community awareness.

I measure changes in policing effort primarily through the standard metric of arrests, also known as clearances (Mas, 2006; Shi, 2009; Heaton, 2010).[3] After an OF, there are numerous potential mechanisms that could affect arrests: First, public scrutiny could increase psychic costs for officers (i.e., loss of morale when interacting with a more aggrieved community) and/or their perceived costs of mistakes, making it more costly for officers to exert effort. If the costs become too high, officers may consider leaving the department, reducing aggregate

[2]There are differences between the more general term officer-involved fatality and police killing, such as the former may include fatal vehicle accidents or pursuit deaths while the latter typically only includes direct use-of-force incidents. However, the primary subset of OFs I consider in this paper—high-profile ones—are all police killings, since they all involve the direct use of force. When describing this high-profile set of deaths, I use these terms interchangeably throughout the paper. But I still use "OFs" as the referring acronym since it is more general.

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

police effort at the municipal level. However, OFs may also affect arrests through other channels besides policing effort, including (2) reduced community cooperation in identifying and locating suspects, (3) reduced civilian reporting of crimes, and (4) changes in criminal behavior, in response to the OF itself and to subsequent signals of policing effort. This paper allays that concern by constructing a novel theoretical model of officer behavior that provides empirically testable predictions by arrest type to identify the causal mechanisms. Discerning which channels are occurring is integral to test the existence of the Ferguson Effect—that is, whether high-profile police killings cause reductions in policing effort due to increased public scrutiny of the police. As increased scrutiny drives up the marginal cost of officer effort, the model predicts that declines in equilibrium effort should disproportionately occur in arrest types that have low marginal benefit to the officer, primarily corresponding to offenses that have relatively low social cost of crime, such as drug possession.

Following a high-profile police killing, officers do indeed curtail effort, but not evenly across crime types: Theft arrests experience temporary reductions of 3–10%. Arrests fall more sharply among crime types that have lower social costs. I create an index of these low-level arrests (e.g., disorderly conduct, liquor law violation, and marijuana possession), and find that these arrests exhibit a sustained and marked decrease of up to 21%, exemplified most prominently by arrests for marijuana possession, which decline by up to 28%. Notably, there are no reductions in arrests for higher social-cost offenses, such as violent crimes or more serious property crimes. These findings are robust to numerous changes in empirical specification, transformations of the dependent variable, varying levels of fixed effects that control for changes in state law and treatment spillovers, and modifying the treatment sample by increasing the high-profile threshold. Moreover, I present evidence that suggests these effects are not driven by a policy mechanism such as a pattern-or-practice investigation or a court-mandated monitoring agreement.

While the decline in effort for clearing theft offenses is temporary, it persists for low-level arrests for at least 1.5 years. The sustained transition to a lower equilibrium effort level likely negates the possibility that these effects are driven by officers being re-assigned to crowd control duties for protests. These low-level arrests are the most sensitive to officer effort because they usually result from officer-initiated stops, where they have the most discretion in determining whether to investigate and/or intervene. The resulting effect pattern, where reductions are concentrated in the low-level arrests for the least serious offenses, align with public scrutiny being the causal channel. This notion is substantiated by evidence that higher-profile police killings generate larger reductions in low-level arrests. Notably, these results are not driven by any single incident, as the findings are robust to excluding a handful of the highest-profile cases. Analyzing interaction-level data from Chicago after the killing of Laquan McDonald, low-level arrests discontinuously fall by a quantitatively similar amount to the national analysis but *only after* the public becomes broadly aware of the incident—

over a year after McDonald's death. Similarly, there are significant reductions in pedestrian and vehicular stops of up to 44% and 12% respectively right after community awareness, providing of some of the first evidence of the impact on stops. This delayed effect negates other potential explanations, such as a police response to the incident itself absent any public scrutiny.

Generally, in the national analysis, the decline in arrests occurs for both Black and White suspects. These results provide evidence that public scrutiny drives a broader pullback of policing activity, rather than one solely focused on reducing interactions with Black suspects or engagement in majority-Black neighborhoods. Notably, for theft and marijuana possession arrests, the effects for Black suspects are suggestively larger than White suspects. Additionally, after the killing of Laquan McDonald, there are sharper changes in arrests and stops of Black civilians. The motivating anecdote of the Ferguson Effect may explain the race results: Officers may not want to make public, street-level arrests of Black individuals, potentially substituting their time to patrolling in their vehicle.

Despite reductions primarily occurring in less serious arrests, the national analysis reveals substantial increases in *serious offending* after a high-profile police killing. Most notably, there is a significant rise of 11–18% in murders and robberies. There are also smaller (and occasionally significant) increases in property crime, driven by theft. Similar to results for arrests, the effects are markedly larger for higher-profile police killings (e.g., an increase in murder of 31%), and does not seem to be driven by any individual incident.

The additional crimes, especially the rise in murder, impose significant crime costs on the affected municipalities. However, because community awareness of the incident and the subsequent public scrutiny drive the reductions of officer effort, it is difficult to identify the determinant of the offending response: is it a response to the reduction in policing effort or is it a reaction to police killing itself driving outrage and legal cynicism? Using heterogeneity tests from both the national analysis and the McDonald case study, I find suggestive evidence for both theories.

Building on past studies, this paper most closely relates to a literature that analyzes the determinants of police officer behavior and the associated crime consequences.[4,5] 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. 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

[4] 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); Heaton (2010); DeAngelo and Hansen (2014); Chandrasekher (2016); Morgan and Pally (2016); Pyrooz et al. (2016); Shjarback et al. (2017); Long (2019); Mello (2019); Rosenfeld and Wallman (2019); Devi and Fryer (2020)

[5] 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); Fryer (2020); Heckman and Durlauf (2020); Ang (2021)

violations. More recently, Cheng and Long (2018) estimate the spillover effects of the death of Michael Brown on eight, predominantly-black US cities relative to 39 other large cities, finding larger decreases in arrests for misdemeanors than felonies, as well as an increase in murders. Rivera and Ba (2019) contrast the effects of low- and high-profile events in shaping police behavior, discovering that two low-profile police events in Chicago result in officers receiving less civilian complaints without changes in crime and arrests; conversely, they find a high-profile scandal increases crime and civilian complaints without commensurate increases in arrests.

Previous studies have provided evidence from case studies or singular events, but this project provides the first national analysis of police effort and offending, estimating the effects from 72 high-profile police killings and 2,740 police departments. The paper extends the analysis through the unique case study of the killing of Laquan McDonald and its subsequent cover-up, allowing for detailed heterogeneity tests on micro-level arrests, crime, and stop data. Further, this paper's findings differentiates itself from recent studies by finding that treatment spillovers (Cheng and Long, 2018), and outside interventions in police departments—such as ACLU monitoring (Cassell and Fowles, 2018) or pattern-or-practice investigations (Devi and Fryer, 2020)—have limited explanatory power for the sharp decline in low-level arrests or rise in violent crime. Finally, the particular context of policing and crime is especially timely and has sizable welfare consequences (Chalfin, 2015; Asher, 2021).

The remainder of the paper is organized as follows: Section 2 provides the background on the Ferguson Effect. Section 3 presents a stylized model of the police officer's objective function and predictions for empirical analysis. Section 4 discusses the numerous datasets in the analysis, how they are cleaned, and why each one is chosen. Section 5 details the empirical specification used, and also highlights the results, showing how policing and offending behavior changes after a high-profile OF. Section 6 discusses the results in the broader context of the theory and literature, and identifies the mechanisms that may be driving the findings through additional heterogeneity tests. Finally, Section 7 recaps the highlights of the paper and concludes.

2 Background on the Ferguson Effect

Over the past eight years, there has been a renewed focus on policing issues following dogged 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"* (Comey, 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. More recently, the killing of George Floyd in Minneapolis caused arguably the largest protest movement in US history (Buchanan, Bui, and Patel, 2020). These events carry a common thread of large and frequent protests as well as continuous news coverage. The impetus for the unrest and attention is 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. Moreover, the consequences of that change in terms of crime costs is equally crucial to understand. The phenomenon has been studied before with various empirical strategies (Morgan and Pally, 2016; Pyrooz et al., 2016; Shjarback et al., 2017; Cheng and Long, 2018; Rivera and Ba, 2019; Rosenfeld and Wallman, 2019), but the broader implications are harder to extrapolate, because of limitations from study design.

If the Ferguson Effect exists, it should locally manifest in jurisdictions that experience high-profile police killings. Estimating the local effects of high-profile police killings necessitates a national analysis with comprehensive fatality data and an empirical specification that internalizes the time-varying aspect of public scrutiny. This paper combines data from 2,740 police departments, 52 of which are treated, between 2005–2016 to study the impact of high-profile OFs on police behavior and crime.

3 Theory and Stylized Model

After a high-profile police killing, it is plausible that there are (1) increases in public 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 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 a set of observed changes in arresting patterns.

The officer's objective is to maximize utility by determining the amount of effort (e) to exert. Effort in this model is conceptualized as various policing activity, such as patrolling, investigating suspicious activity, stopping civilians, and even making arrests. 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), community cooperation (F), and reported crime (RC_c).[7] The objective of the officer is to maximize utility:

[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 are still possible.

[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 monetary, personal, and/or professional utility returns based on the arrest type they clear. Officers take reported crime, not the true level of crime, as given because they are unable to police beyond the offenses known.

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

To reward effort in policing, police departments may partially determine promotions based on arrests, typically at the interview stage (Evans and Ba, 2021), valuing either "a high volume of them or a string of high-quality ones" (Samaha, 2017).[8] More arrests (A) increase the likelihood of promotion and/or ancillary payments, since officers are compensated by the number of hours worked and overtime pay for further policing work and court appearances related to their arrests (Samaha, 2017; Hughes, 2020; Mastrorocco and Ornaghi, 2020). My conversations with officers also reveal a personal utility return, where they highlight a commitment to service and valuing "worthwhile arrests." The officer's private 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 B.1 detailing the specific costs by offense.[9] 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 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

[8] Promotions are primarily driven by tenure and typically involve testing, but at the interview stage, arrests and clearance rates may play a role in promotion (Ornaghi, 2019; Evans and Ba, 2021).

[9] Table B.1 uses the cost of crime estimates from Chalfin and McCrary (2018). Reflecting the usage in 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.

$$\implies e^{\star} = \frac{1}{\phi}\left[\frac{df_c}{dA_c} \cdot \frac{\partial A_c}{\partial e_c}\right] \tag{1}$$

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.[10]

The Ferguson Effect—or reduced officer effort as a result of increased scrutiny from a high-profile police killing—manifests itself through exogenous increases in the cost of effort multiplier, ϕ. Increases in public scrutiny from media coverage, community attention, and protests may result in additional psychic costs, as officers experience disutility and a loss of morale from protests and interactions with a community that may be apprehensive and critical of them. Additionally, officers—as well as police departments more broadly—may perceive higher costs of mistakes and misconduct potentially shaping their behavior (or policing strategies) to serve a more vigilant community where civilians are recording officers as they make stops and more journalists are attentive to policing issues.[11] 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). If the overall cost of effort becomes too large for the officer in the department involved in the police killing, that may cause officers to retire early or voluntarily resign and seek employment elsewhere, as seen in a few jurisdictions (Westervelt, 2021). In this section, however, the focus is on the intensive margin decisions

[10] 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. Conversely, officer retirements and resignations represent lower aggregate effort exertion on the extensive margin if the calculation is at the jurisdiction level.

[11] 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 this model (Schwartz, 2014; Rushin, 2017; Rushin, 2019; Grunwald and Rappaport, 2020). 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.

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; Rushin, 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.

that officers make.

Thus, the comparative static of interest is $\frac{\partial e^\star}{\partial \phi}$: how does the rise in the marginal cost of effort—from additional public scrutiny—affect equilibrium effort? Let $e^\star$ be the choice of effort 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.[12]

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,

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

Additionally, $\frac{\partial A_L^\star}{\partial e_L} > \frac{\partial A_H^\star}{\partial e_H}$ and $\frac{df_H}{dA_H} \gg \frac{df_L}{dA_L}$, where differences between the highest and lowest social-cost crimes are several orders of magnitude (see Table B.1 for comparisons

[12]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.

between Part I crimes). Low marginal-benefit arrests are more sensitive to officer effort because they usually result from officer-initiated stops, where they have the most discretion in determining whether to investigate and/or intervene.

Thus,

$$\left|\frac{\partial A_L^\star}{\partial \phi}\right| > \left|\frac{\partial A_H^\star}{\partial \phi}\right| \tag{3}$$

Therefore, an exogenous increase in public scrutiny—which is modeled as an increase the marginal cost of effort—causes equilibrium officer effort and arrests to decline. Further, model predicts that this decrease in arrests will primarily occur for offenses that produce low marginal benefit to the officer (i.e., generally low social-cost crimes). I explore this mapping empirically in Section 5 when testing for reductions in arrests after controlling for population, and time-invariant department/municipality characteristics.

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

$$\frac{\partial A_c^\star}{\partial F} = \frac{\partial A_c^\star}{\partial e} \cdot \frac{-\dfrac{\partial^2 U}{\partial e_c \partial F}}{\dfrac{\partial^2 U}{\partial e_c^2}} = \frac{\partial A_c}{\partial e_c} \cdot \frac{-\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial F}}{\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} = \frac{-\dfrac{\partial A_c}{\partial e_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial F}}{\dfrac{\partial^2 A_c}{\partial e_c^2} - \phi \Big/ \dfrac{df_c}{dA_c}} > 0 \tag{4}$$

suggesting that decreases in community cooperation from a high-profile OF result in reductions in equilibrium arrests. Increased scrutiny and reduced 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., murder)[13]. Because $\frac{\partial^2 A_H}{\partial e_H \partial F} \gg \frac{\partial^2 A_L}{\partial e_L \partial F} \approx 0$, then

$$\frac{\partial A_H^\star}{\partial F} > \frac{\partial A_L^\star}{\partial F} \approx 0 \tag{5}$$

[13] 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.

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-benefit arrests (L), then the resulting mechanism is a decline in officer effort from increased 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), if the marginal offender finds the death of a civilian as a salient signal of interaction risk with law enforcement. Alternatively, it could increase as the high-profile police killing itself drives diminished perceptions of police legitimacy, leading to increases in legal cynicism and resulting in people *taking the law into their own hands* (Persico, 2002; Kirk and Papachristos, 2011; Leovy, 2015; Rosenfeld and Wallman, 2019). It could also increase if the marginal offender see visible reductions in police effort and internalizes that signal as diminished apprehension risk (Becker, 1968).

Similarly, civilian reporting of crime, the other component of RC_c, may change after a fatality (Desmond, Papachristos, and Kirk, 2016). In order to properly assess the existence of the Ferguson Effect, I am interested in changes in ϕ holding all else constant. Thus, in Section 6, I explore the relevant mechanisms by directly controlling for changes in reported crime, closing an intermediary channel where arrests may fluctuate from changes in offending or reporting. Thus, I focus on the two remaining channels: public scrutiny (Equation 3) and community cooperation (Equation 5).

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

$$\frac{\partial A_c^\star}{\partial RC_c} = \frac{\partial A_c^\star}{\partial e} \cdot \frac{-\dfrac{\partial^2 U}{\partial e_c \partial RC_c}}{\dfrac{\partial^2 U}{\partial e_c^2}} = \frac{\partial A_c}{\partial e_c} \cdot \frac{-\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial RC_c}}{\dfrac{df_c}{dA_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c^2} - \phi} = \frac{-\dfrac{\partial A_c}{\partial e_c} \cdot \dfrac{\partial^2 A_c}{\partial e_c \partial RC_c}}{\dfrac{\partial^2 A_c}{\partial e_c^2} - \phi \Big/ \dfrac{df_c}{dA_c}} > 0 \qquad (6)$$

Expectantly, as officers are aware of more crimes, the equilibrium level of arrests increases—or vice versa. But, that does not strictly hold for the least serious offenses (L), where $\dfrac{\partial^2 A_H}{\partial e_H \partial RC_H} \gg \dfrac{\partial^2 A_L}{\partial e_L \partial RC_L} \approx 0$. 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.[14] As a result,

$$\frac{\partial A_H^\star}{\partial RC_H} > \frac{\partial A_L^\star}{\partial RC_L} \approx 0 \tag{7}$$

4 Data Description

4.1 Arrest and Crime Data

The study combines government data on arrests, reported crime, stops, 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 contain data from nearly every police department in the US. In addition the national analysis, I also combine data on arrests, crimes, and stops in Chicago for an analysis of Laquan McDonald's death. I discuss the data in more detail in Section 5.6. Finally, I also integrate datasets such as the FBI's National Incident-Based Report System and the PBS Frontline's Pattern-or-practice Investigations to provide further insights on the analysis, which I discuss in Section 6.

The primary data contains monthly arrests for Part I offenses (murder, aggravated assault, robbery, motor vehicle theft, burglary, theft) and known offenses (i.e., crimes reported by civilians and/or observed by the police).[15,16] Separately, it details Part II arrests, a broader offense classification set than Part I—notably containing the least costly crimes, such as disorderly conduct or marijuana possession.[17]

I create an index of low-level arrests—Part II clearances that will likely have larger effects because officers have more discretion in determining whether to intervene: curfew/loitering, disorderly conduct, public drunkenness, Driving Under the Influence (DUI), liquor violations, marijuana possession, marijuana sale, prostitution, suspicion, vagrancy, and vandalism.[18] These arrests largely correspond to the low marginal-benefit arrests discussed in Section 3, since they typically produce little monetary, personal, and/or professional return to the officer, and generally have the lowest social costs of crime victimization. Lack of DUI enforcement can lead to substantially higher social costs (DeAngelo and Hansen, 2014), but the arrests

[14] In fact, Section 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.

[15] The exact rules for how they record crime and arrest data are further explained in Section A.1. It is important to highlight that theft and motor vehicle theft are mutually exclusive categories. Murder encompasses both murder and non-negligent manslaughter.

[16] 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.

[17] Many Part II offenses have no distinct victims—leading to the FBI to describe them as crimes against society—and subsequently, they are far less likely to be reported by the community. Thus, the FBI only reports them if they result in arrest.

[18] 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.

are included in the low-level index because officers have anecdotally indicated that there is less priority placed on their enforcement relative to other offenses (Murgado, 2012) and in certain municipalities, a far majority of them are misdemeanors. Importantly, the none of the results hinge on whether DUI is included or not. By compiling certain Part II offenses into a low-level arrest index, I reduce the false discovery rate—the likelihood of finding falsely significant, spurious effects (Anderson, 2008).

The arrest data is reported separately by race of suspect. I focus on the arrests of Black and White suspects, and the data is structured at the agency-month-race level. In order to ensure that the data maps accurately to the jurisdictions, I clean it using rules described in Section A.2, following Evans and Owens (2007) and Mello (2019).

For this study, I use monthly arrest and crime counts to analyze policing effort and offending behavior in the aftermath of 72 high-profile OFs in 52 of 2,740 police departments in the analysis sample between 2005–2016. Table 1 reports that treated jurisdictions have on average lower arrest rates for White suspects and higher rates for Black suspects than control jurisdictions across all types of crime: violent, property, and the low-level arrests. Moreover, across race, the arrest rates for Black suspects are substantially higher than White ones 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 Black people and fewer White people, while having similar educational attainment and slightly higher poverty rates, driven by Black poverty.[19]

4.2 Officer-Involved Fatality Data

The official government estimates for officer-involved fatalities, in the FBI Supplementary Homicide Report within the category of "Justifiable Homicides," are extremely undercounted, resulting in non-governmental estimates having up to three times the number of fatalities. The most comprehensive dataset is from Fatal Encounters, who has managed to create a sophisticated collection system: collating data from Freedom of Information Act requests, news sources, and using paid researchers and volunteers to run additional searches and data checks from public sources.[20] 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.[21] I create this high-profile subset by scraping the number

[19] 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 analysis.

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

[21] Most fatalities do not result in significant media coverage and subsequently produce limited community

of news articles written about each incident. In Section A.3, I describe the scraping procedure in more detail. Figure B.1 provides a histogram of the distribution of news articles in bins of 500 articles.[22] The two obvious break points in the distribution are when fatalities generate more than 1,000 and 2,500 news articles. 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 lead to larger treatment effects, there is reduced statistical power due to the small number of OFs.[23] Therefore, I define "high-profile" as having at least 1,000 news articles written about the incident.[24]

To see how the high-profile police killings vary across space, Figure B.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. As expected, there are a few clusters in population centers like Chicago, Los Angeles, and New York. 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 is captured on video or seen by witnesses, rather than the number of news outlets—as demonstrated by the death 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 B.3). 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.[25] 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). There are

awareness and protests. Additionally, any empirical design that includes large US cities and leverages the timing of all officer-involved fatalities would be hampered by the frequency in which they occur (e.g., Los Angeles experiences more than 12 OFs every year in the sample frame). The effects shown in this paper will be identified off of departments and time periods that experience no OFs or low-profile ones.

[22]The bin from 0-500 articles is excluded since it contained over 99.9% of officer-involved fatalities, rendering the rest of the histogram indecipherable. Similarly, given the heavy right skew of the distribution, the histogram does not depict some of the most high-profile OFs. There are nine OFs that are covered by more than 20,000 news articles each, which are inputted into the farthest right bin.

[23]In Section B.8, I show that the results are qualitatively robust to increasing the threshold levels of 2,500 and 5,000 articles.

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

[25]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.

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 further explore this issue in Section 5.4, 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 results, 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, when public scrutiny can plausibly begin. I describe the protocols used to determine whether to change the event time in Section A.4. 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 5.6. Overall, there are 10 events in the analysis that have timing changed.

The treatment sample is 72 high-profile OFs spread across 52 jurisdictions. The maximum number of OFs in a single jurisdiction is six (in Los Angeles). Table 2 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, and thus plausibly consistent with a public perception of injustice. This perception likely drives increases in scrutiny, which is evidenced by the additional media coverage of an incident.

5 Empirical Strategy and Results

In this section, I describe the empirical specifications and the corresponding results related to whether high-profile, officer-involved fatalities affect arresting and offending behavior. The identification strategy relies on the quasi-random timing of the OFs to induce a plausibly exogenous shock of increased scrutiny to the local police department. 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.[26] Thus, control jurisdictions include departments who experience no OFs or less publicized ones. Section 5.1 discusses difference-in-differences (DD) and triple difference (DDD) estimation using only the highest-profile, officer-involved fatality per jurisdiction. Then, Section 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. Finally, Section 5.3 explores the impact of a high-profile, officer-involved fatality on crime outcomes, using the DD and event study strategies.

[26] 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 B.3). The empirical analysis addresses the concern by including department and month-of-sample fixed effects, while controlling for population, and county-specific linear trends.

5.1 Difference-in-Differences and Triple Difference

5.1.1 Difference-in-Differences 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

$$Y_{it}^{cr} = \mu_i + \lambda_t + \rho(t)_i + \beta Pop_{iy} + \theta(Treat * Post_{it}) + \nu NegBin_{it} + \alpha PosBin_{it} + e_{it}^{cr} \quad (8)$$

where Y_{it}^{cr} represents log arrests of crime type c and race of suspect r in department i during month t.[27] Crime type c consists of a variety of violent, property, and low-level offenses, as well as those respective crime categories. For this analysis, I only include arrests where the race of the suspect (r) is Black or White.[28] The coefficient of interest is θ on $Treat * Post_{it}$, which represents the approximate percent change (divided by 100) in the involved department's arrests in the 8 months before a fatality to 16 months after relative to control departments.[29] To have θ measure a window of time 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).[30]

I use department and month-of-sample fixed effects (μ_i, λ_t), and county-specific linear trends ($\rho(t)_i$).[31] Pop_{iy} controls for yearly log population. Standard errors are clustered at the police department level. 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, in Section 3.

[27]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.

[28]The analysis is run at the agency-month-race level to have consistency between the DD and DDD designs. The results are robust to running the DD (or event study) at the agency-month level instead, rendering any distinction between the approaches trivial.

[29]The coefficient literally represents the difference in log points. Log points are a rough approximation for percent, where estimates are more precise for smaller changes in log points. The exact formula for conversion is $\% = 100 \cdot [exp(\hat{\beta}) - 1]$. All percent changes described in this paper are exact difference using the conversion formula. Thus, for larger log-point differences, there may be visual discrepancies between stated percent change and what is seen in log points.

[30]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.

[31]I employ county trends to control for the steady decline in arrests 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 quite similar, but computationally more intensive. As pointed out in Wolfers (2006), there is a concern that models with linear trends may be overly restrictive, resulting in the trend control capturing part of the treatment effect. Since the coefficient estimates are also quite similar—though slightly noisier—after excluding the linear trends from the model, that concern seems minor. The county-specific linear trends remain in the preferred specification because of the increased precision.

Table 2 provides descriptive evidence that police killings of Black civilians are higher profile, with an average of over 20,500 news articles written about each incident compared to White OFs which average around 3,000 mentions. It is possible that public scrutiny, operating along this racial dimension, increases the marginal cost of effort for arresting Black suspects—if, for example, police officers perceive higher error costs for these interactions—without significant changes for White suspects. To test for racial heterogeneity in arrests, the DDD specification is quite similar to the DD model

$$\begin{aligned} Y_{it}^{cr} = & Black * \mu_i + Black * \lambda_t + Black * \rho(t)_i + \beta Pop_{iy} + \omega(Black * Treat * Post_{it}) \\ & + \delta(Treat * Post_{it}) + \nu(Black * NegBin_{it}) + \alpha(Black * PosBin_{it}) + e_{it}^{cr} \end{aligned} \tag{9}$$

with the Treat*Post variable and the binned ends of the event window being interacted with an indicator variable (*Black*) denoting whether the arrested suspect was Black. Further, I interact the Black arrest indicator variable 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. For the DDD, the coefficient of interest is ω on the Black*Treat*Post variable, identifying racial differences in the Treat*Post coefficient.

5.1.2 DD and DDD Results

The impact on officer effort is measured by statistically significant drops in log arrests in the 16 months after the highest-profile OF per jurisdiction. After accounting for other potential channels, reductions in low-level arrests (e.g., generally low social-cost offenses) represent evidence of the effect of scrutiny on policing effort. Figure 1 presents the DD coefficients from separate regressions of individual arrest types in descending order of their respective social cost of crime. With the exception of motor vehicle theft arrests, none of these regressions have significant differences between the average pre-treatment trends of control and treated departments, providing a partial test of the 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).[32] However, property crime arrests (motor vehicle theft, burglary, and theft) decrease by 5.0% (Table B.3), driven by a significant decline in theft arrests (7.1%)—the least socially costly property offense

[32] The actual DD tables, such as Table B.2, are in the appendix. Column (1) provides the DD estimate, while Column (2) runs the test of parallel pre-trends. The "Pre-trend 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 pre-trend test is significant, then the identifying assumption of the DD is not met.

(Table B.1).[33]

Shifting to the least serious offenses, I see much sharper reductions in arrests for marijuana sale (12%), disorderly conduct (13%), and marijuana possession (13%). These decreases help drive the overall decline in low-level arrests of 9.1%, relative to the eight months before an OF (Table B.4).[34] This effect pattern is broadly reflective of the reductions in policing effort occurring along less socially costly crimes, evinced by Figure 1.

In Table B.2, I present the DDD race results for violent crime arrests in Column (3), while Column (4) runs the test of racially parallel pre-trends. The table highlights the lack of racial differences in the (lack of) change in violent crime arrests. On the other hand, Table B.3 highlights suggestive evidence that Black theft arrests fall by 5.9% relative to White theft arrests. The DDD empirical design does not uncover any significant racial differences for low-level clearances (Table B.4). Overall, the changes in arrests following an OF do not appear to vary significantly by race.

5.2 Event Study Design

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 indicators, where there is an indicator for each unit of time away from the event within a specified event window. For this study, the preferred event window is -3 to 5 quarters (-9 to 17 months) around the event with a time indicator for each quarter. The indicator 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 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 72 to 52 events—taking away the variation in timing of later treatments. Moreover, drops in arrests from future events would be incorrectly attributed to later event indicators.

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

[33] An important caveat is that property crime arrests experience statistically significant differences in pre-treatment trends between treated and control departments, driven from pre-treatment differential trends in motor vehicle theft arrests. I present the property crime arrest coefficient only for comparison with the respective event study coefficients, which are similar in size and do not suffer from differential pre-trends.

[34] For more explanation on the low-level index, refer to Section 4.

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 indicators 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 indicators should equal one). The preferred event window is -3 to 5 quarters (-9 to 17 months) around the event with a time indicator for each quarter. Because certain departments experience multiple OFs, the binned endpoints of the event window need to be the sum of pooled time indicators outside of the event window, taking on values up to six in the dataset for Los Angeles. Consequently, I define a flexible "indicators" setup of D_{it}^{j}[35]:

$$D_{it}^{j} = \begin{cases} \sum_{j=-J}^{a-1} \sum_{m=1}^{n_i} \mathbb{1}(t = \tau_{im} + j) \text{ for } j < a \\ \sum_{m=1}^{n_i} \mathbb{1}(t = \tau_{im} + j) \text{ for } a \leq j \leq b \\ \sum_{j=b+1}^{J} \sum_{m=1}^{n_i} \mathbb{1}(t = \tau_{im} + j) \text{ for } j > b \end{cases} \tag{10}$$

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 B.4 shows a visual example of the indicator structure in months, using simulated data of a department with two officer-involved fatalities: one in January 2006 and the other in May 2006—the highlighted months. Thus, the event time indicator for time 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 indicator 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

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

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 B.4, there are missing values for a few event time indicators, 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 indicators 1 to 3) need to be set to missing. Thus, the remaining data that 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 indicators into quarter indicators (-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, police-involved death by month, crime type, and race. Given parallel pre-treatment trends between treated and control departments, any change after a police killing is reflective of the causal impact of the event on arrests or crime. Following the DD model, the preferred estimating equation is

$$Y_{it}^{cr} = \mu_i + \lambda_t + \rho(t)_i + \beta Pop_{iy} + \sum_{j=a-1, j\neq -1}^{b+1} \theta_j D_{iq}^j + e_{it}^{cr} \tag{11}$$

where the indicator structure (Equation 10) replaces the Treat*Post variable. The θ_j's are the coefficients of interest on event time indicators, D_{iq}^j, where I pool the month indicators into quarter indicators (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, time-invariant city characteristics, county-specific linear trends, and temporal variation in arrests nationally. With all of the specifications, I cluster the standard errors at the department level.

To estimate changes in arrests by race, I create separate indicator variable sets for both Black and White arrests by interacting a Black and a White indicator variable with the original event time indicator set, using the same integrated approach detailed above.

$$\begin{aligned} Y_{it}^{cr} &= Black * \mu_i + Black * \lambda_t + Black * \rho(t)_i + \beta Pop_{iy} \\ &+ \sum_{j=a-1, j\neq -1}^{b+1} \chi_j * Black * D_{iq}^j + \sum_{j=a-1, j\neq -1}^{b+1} \gamma_j * White * D_{iq}^j + e_{it}^{cr} \end{aligned} \tag{12}$$

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.

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 (72) and richer variation. Event study graphs have the coefficient on log arrests on the vertical axis and event time indicators as the horizontal axis, where period 0 is when a police department is involved in a high-profile OF. Figure 2 shows a very slight suggestive reduction in all Part I arrests (violent and property crime plus simple assault and negligent manslaughter) that expands over the first few quarters in a downward arc shape, reaching the low point in Q3 and then returning to the initial equilibrium in Q5. Property crime arrests drive this pattern, experiencing reductions of 2–8%. On the contrary, violent crime arrests exhibit minor *increases*, culminating in a significant increase of 7% in Q3. There is a persistent and continual decline in low-level arrests, where the sharpest drop of 21% is about three times that of property crime arrests.

Consistent with the DD results, murder, aggravated assault, and burglary exhibit minimal change in arrests, aside from a suggestive uptick in assault arrests in Q3 and a suggestive decline in burglary arrests in Q1 (Figure 3). Theft arrests, again, show a significant decline of 3–10%, instigating the effects observed in property crime arrests. These changes are dwarfed by those of the low-level arrest categories (Figure 4).[36] Marijuana possession arrests experiences the most prominent change, falling by 14–28% in Q1–Q5.[37] Disorderly conduct arrests follow a similar pattern with reductions between 12–18% for Q1–Q5. These patterns are emblematic of significant reductions in other low-level arrests after Q0, such as liquor violations (19–28%) or marijuana sale arrests (8–18%). None of the aforementioned arrest categories demonstrate any statistically significant differences in the pre-treatment trends between control and treated departments.

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 the effects on arrests (Figure B.34, Figure B.35, Figure B.36). However, Figure 5 shows that theft arrests suggestively fall more for Black suspects (4-15%) compared to White ones (1-7%). Similarly, Black arrests for marijuana possession suggestively drop more than White arrests (7-34% versus 3-22% respectively).[38] Although

[36]The vertical axis of the log point scale may be shifted to accommodate larger confidence intervals from larger effects.

[37]Concern about reductions in marijuana possession arrests being driven by a consistent downward trajectory during 2005–2016 should be assuaged by (1) the inclusion of county-specific linear trends, (2) the effect pattern being robust to changes in state laws through the introduction of state-by-year fixed effects, and (3) the trend not continuing if the event window is expanded (i.e., for that regression, Q-3 is higher than Q-4).

[38]However, White marijuana possession arrests are significantly different in control and treated departments three quarters before an OF, violating the identification assumption.

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.

5.3 Crime Analysis: Difference-in-Differences (DD) and Event Study

5.3.1 DD Results

I employ the same DD regression as Equation 8 with Y_{it}^{c} now representing log offenses of crime type c.[39] Figure 6 presents the DD coefficients from separate regressions of individual offense types in descending order of their respective social cost of crime. None of these regressions have significant differences between the average pre-treatment trends of control and treated departments. After a high-profile, officer-involved fatality, the involved police department experiences a substantial increase in violent crime of 9.1% (Table B.5), driven by sharp increases in murders (15%) and robberies (13%). There are more moderate increases in property crime of 5.5% (Table B.6), largely from a significant rise in motor vehicle thefts (8.2%) and thefts (5.2%).

5.3.2 Event Study Results

Using the previous event study design, I employ Equation 11 with Y_{it}^{c} as log offenses, increasing the number of fatalities in the treatment sample from 52 to 72. Consistent with the DD findings, Figure 7 documents the significant and sizable increases of 11–18% in murders and robberies for Q0–Q4. The increases in murder taper off in Q5 after an OF, whereas robberies remain at elevated levels. There are also smaller, suggestive increases in motor vehicle theft and theft of up to 7% and 5% respectively (Figure 8). Unlike in the DD, the increases for motor vehicle theft and theft are too noisy to be statistically significant, but for theft and property crime in general, there seems to be consistent pattern of steady increases after an OF.

5.4 Spillover Analysis

One remaining concern is that the treatment exposure may not be specific to the involved department, potentially leading to spillover effects in other municipalities (Morgan and Pally, 2016; Cheng and Long, 2019). If these spillovers occur nationally, they are already being controlled for by the month-of-sample fixed effects. Additionally, the results are robust to spillovers that occur across jurisdictions with similar populations (Section B.2.1) or within a state (Section B.2.2), since the size and significance of effects are largely unchanged after interacting the population group or state with the month-of-sample fixed effects.

[39] The FBI crime data does not contain information on the race of the suspect when an arrest is not made, hence the exclusion of the race notation.

The most likely scenario is that these high-profile OFs affect policing and crime in geographically-adjacent agencies. I directly test this in Section B.2.3, where I use the DD empirical model (Equation 8) but expand the geographic definition of treatment to the county, rather than just the involved police department. Overall, the average effects on arrests and crime in the treated county are heavily attenuated (Treat*Post in Column 1). Separating the effects between the involved department and the spillover departments that reside in the same county in Column 3, I find that the spillover departments (Spillover*Treat*Post + Treat*Post) experience reductions in arrests that are much more minor than the involved agency (Treat*Post). The involved department sees reductions in theft (8%), disorderly conduct (12%), and marijuana possession arrests (15%) that are similar to the main results, while the spillover agencies have reductions of around 2%—none of which are significantly different than zero. This pattern is especially true for crime, where the spillover jurisdictions experience no change in offending behavior.

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, consistent with the geographic spillover analysis. 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 instigate much greater scrutiny, with nearly two-thirds of them facing protests (Table 1). Consequently, there are sharp drops in policing effort for less serious offenses, evinced by reductions in low-level arrests of up to 21%.

5.5 Addressing Potential Negative Weights with Two-way Fixed Effects Designs

Recent econometric papers have highlighted that difference-in-differences designs with two-way fixed effects (TWFE) and staggered treatment timing, such as Equation 8 in this paper, may have coefficients derived from implicit "negative weights" when there are heterogeneous treatments effects over time or across units (Goodman-Bacon, 2018; de Chaisemartin and D'Haultfœuille, 2020). One potential solution is to estimate the coefficients in an event study design, which does not suffer the same issue, as long as there is limited heterogeneity in treatment effects across cohorts (defined by initial treatment time). Since the preferred specification is the event study one (Equation 11), the lack of pre-treatment trends—evidenced by the flat pre-period—and clear post-treatment effects in most figures is reassuring. However, if the pattern of effects differ between earlier treated units and later treated ones, the event time indicators still could be contaminated with negative weights in specifications using two-way fixed effects (Sun and Abraham, 2020).

By estimating the specification with only observational unit fixed effects—no month-year fixed effects or county-specific linear trends—I circumvent concerns that the coefficients are a weighted average composed with negative weights. I find that the effects on arrests and violent crime are quite similar in the one-way fixed effects model (respectively, Figure B.10 and Figure B.11), while the effects on property crime are a bit attenuated. The stability of coefficients between the one-way and two-way fixed effect models suggests that negative weights are not a concern for the preferred specification, and that the effects are likely being driven from differences in timing of treated units.[40] The latter theory can be assessed by estimating the preferred specification with exclusively treated units (i.e., police departments that experience at least one high-profile police killing between 2005–2016). The results using only treated departments closely reflect the ones from the one-way fixed effects model, where the main results hold but the property crime coefficients are attenuated (Figure B.12, Figure B.13). This similarity is intuitive since the one-way model is only using the never-treated units to help estimate the population control.

To understand how the coefficients under the specifications with the two-way fixed effects change under TWFE estimators that are robust to heterogeneity across treatment group or time, I implement the cohort interaction-weighted (IW) estimator from Sun and Abraham (2020). Because Sun and Abraham (2020) does not address empirical settings with multiple treatments per observational unit, I limit the treatment sample to the highest-profile police killing for each jurisdiction, and define cohort as the quarter in which that occurs, focusing on never-treated units as the control group. I then explore the robustness of the effects from the DD design, as well as an event study that only uses the highest-profile event per jurisdiction, using the IW estimator. Since the DD coefficients accurately summarize what is seen in the event study figures, I present the cohort IW DD coefficient plots for arrests in Figure B.14 and crime in Figure B.15. The DD coefficient plots have a clear comparison, since I have previously shown the DD plots of the effect of the highest-profile police killing per jurisdiction, but I have not shown the event study ones.[41] The pattern of significant effects for arrests is similar, where the reductions are concentrated in those low-level arrests, but now theft arrests has a statistically significant pre-trend (despite having a similar coefficient). The effects on crime are also robust to using the IW estimator with the coefficients now being slightly larger and all statistically significant.

[40]Concerns about the negative weights were already partially diminished because the analysis sample is largely comprised of never-treated units—2,687 of the 2,739 agencies. These units provide much of the variation to estimate the month-year fixed effects. Since the coefficients of interest do not significantly change from the exclusion of these time fixed effects, it suggests that the never-treated units are not heavily influencing the coefficients.

[41]Generally, the arrest and crime results for the event study specification (Figure B.16, Figure B.17) using only the highest-profile police killing are roughly similar to the main event study findings, but are noticeable smaller in the actual coefficient estimates. These differences are expected given the observed discrepancies between the DD coefficients and the event study ones in the main results. Furthermore, since the event study results using the traditional estimator and including only the highest-profile police killing per agency have not been shown, I stick to discussing only the DD results for brevity.

5.6 Case Study of Laquan McDonald and the Chicago Police Department

To further explore the dynamics of public scrutiny on police behavior and crime, I consider the case study of the killing of Laquan McDonald by the Chicago Police Department, where there is a the year gap between the incident and broad community awareness. I employ Freedom-of-Information-Act requests to the Chicago Police Department (CPD) to acquire detailed micro-level data on murders, shootings, and pedestrian and vehicular stops, which I combine with publicly available data on arrests and crimes from CPD, and search trend information from Google.

In October 2014, Laquan McDonald was fatally shot in Chicago, where a police incident report incorrectly described McDonald lunging at an officer with a knife. Given the misinformation, the use of force was considered justified and there was a lack of awareness (and scrutiny) of the event, as evidenced by Google Trends (Figure B.18). In February 2015, a journalist uncovered the official autopsy report, which raised questions since McDonald was shot 16 times—with some of bullets entering from his back (Kalven, 2015). The journalist also revealed that there was footage of the incident that CPD was refusing to release. During this time, there is no broad awareness of the incident, according to locals searching for it on Google. However, in late November 2015, CPD was forced by a court order to release the video. The dash cam footage directly contradicted the initial incident report. Consequently, there was a surge in public scrutiny—measured by local Google searches of the incident (Figure B.18) and repeated protests—eventually culminating in tens of thousands of news articles on the incident.[42]

To assess the impact of these separate events on arrests, reported crime, and stops, I run an event study regression of log outcome (Y_{it}^{cr}) on monthly event time indicators (D_{it}^{j}), using beat-level fixed effects (μ_i) and linear trends ($\rho(t)_i$). Following the event study design (Equation 11) without control units, the preferred estimating equation is

$$Y_{it}^{cr} = \mu_i + \rho(t)_i + \sum_{j=a-1, j\neq -1}^{b+1} \theta_j D_{it}^{j} + e_{it}^{cr} \tag{13}$$

where θ_j's are the coefficients of interest on the monthly event time indicators (D_{it}^{j}), denoting time from the *video release in November 2015*. The month before the video release is omitted, and the coefficient estimates are normalized such that $\theta_{-1} = 0$. The other θ_j's represent the approximate percent change in the outcome in the month before the event to j months after, after controlling for time-invariant beat characteristics and beat-specific linear trends. I cluster the standard errors at the beat level.[43] For all regressions where race (r) data is

[42] The officer who shot McDonald was convicted of second-degree murder in October 2018, which may have increased the overall news results on this incident. Overall, that is a rare judicial result, even in these high-profile cases.

[43] Clustering at the district level (the larger geographic unit) produces qualitatively similar results, rendering a few event time indicators for the pre-treatment period statistically insignificant.

available, I focus on Black and White civilians to have comparable results to the national analysis, but the results including all races is quite similar.

Figure 9a provides the dynamic impacts of Laquan McDonald's death, autopsy release, and dash cam release on low-level arrests conducted by CPD. At the time of his death, there is minimal change in low-level arrests because of the erroneous incident reports resulting in a lack of awareness of the incident. Two months afterward there is a significant and singular drop of 8%. However, because I am running the regression at the month level (and not the quarter level) to highlight the immediacy of the effect, I expect a higher potential for spurious significant effects, which this may be given the lack of consistency with adjacent indicators.

After the journalist discloses the findings of the autopsy report, low-level arrests declined by 7% that month. The subsequent pre-treatment period indicators mostly show smaller and insignificant reductions. It is possible that this significant drop is statistical noise, as there is no broad awareness of the incident at this time (according to Google Trends) and the effect size does not substantially deviate from adjacent estimates, although the coefficients are consistently negative for the most part. Finally, after the dash cam footage was released, there was the spike in community awareness, as shown in Figure B.18. Low-level arrests discontinuously fell by 11% that month, despite it occurring in the last week of November. In the ensuing year, these low-level arrests faced precipitous and growing declines of 22–36%, mirroring the results in the national analysis.[44] In contrast with the national analysis, the reduction in low-level arrests is far more pronounced in Black suspects Figure B.19.

In totality, the sharp and persistent reductions in low-level arrests coinciding precisely with community awareness—despite a year passing since McDonald's death—provides cogent evidence that further substantiates public scrutiny as the key mechanism. The case study also helps negate other potential explanations, such as a police response to the incident itself absent any public scrutiny. However, other possible confounders exist: As a consequence of the public outcry following the video release of Laquan McDonald's murder, the federal Department of Justice initiated a pattern-or-practice (PoP) investigation into CPD that began in January 2016. That month also happened to coincide with the start of an unrelated police monitoring agreement with the ACLU. The sharp drop in arrests in November and December 2015 (months 0 and 1, Figure 9a), before the investigation or the monitoring took place, suggests that public scrutiny plays a pivotal role in the reduction of policing effort—separate from these outside monitoring/investigative initiatives.

Shown in Figure 9b, there is also a sharp reduction in pedestrian stops. The impetus for the decline has been attributed to the ACLU monitoring agreement (Cassell and Fowles, 2018), which made pedestrian stops far more labor intensive by forcing officers to fill out a new two-page investigatory stop report. These additional labor costs were imposed in

[44] As the difference in log points grows, the approximation of percent change is less precise. Even though decline in low-level arrests reaches 0.45 log points, the percent change is 36%, using the formula $\% = 100 \cdot [exp(\hat{\beta}) - 1]$.

January 2016, and they likely explain a large extent of the persistent over 83% drop.[45] However, pedestrian stops significantly fall for two months *before* this change was imposed, resulting in reductions of up 44%. As Cassell and Fowles (2018) point out, CPD officers substitute into making vehicular stops instead Figure 9c, since these stops required less documentation. Notably, these stops also significantly decline, with decreases up to 12%, in the two months after the video of McDonald's death was released—before the incentives to switch from pedestrian to vehicular stops existed. Theoretically, a larger reduction in pedestrian stops is consistent with these street-level interactions likely being more affected by public scrutiny. Consistent with the low-level arrests findings, these changes in stops are larger for Black civilians than White ones (Figure B.20). Although the additional reporting costs in January 2016 results in a dramatic shift from pedestrian stops to vehicular stops—with eventual increases of nearly 170%—there is a large overall reduction in police-civilian interactions measured by stops (Figure B.21).

This scandal also coincides with significant increases in murders of 5–14% (Figure 9d), which is consistent with the results from the national analysis—although the increase in Chicago does not occur immediately.[46] However, these results also show a significant but smaller rise occurring 2–4 months before the video was released. The increase in murders largely afflicts Black victims (Figure B.22), especially when considering the unadjusted numbers (Figure B.23), though there are more muted increases in murders of White victims.

To tease out potential mechanisms, I explore heterogeneity in the increases in murder by victim-offender relationship and location. If the increase in murders is a response to a visible reduction in police effort, measured by low-level arrests, then perhaps the murder rise should be more prominent in more pre-meditated crimes than "crimes of passion." I test this by exploring whether the increase is being driven by murders where the offender and victim do not have a "domestic" relationship, defined in Illinois as violence that occurs between family or members of a household. Figure B.24a show that the increases are entirely driven by these non-domestic murders, though an important caveat may be that the share of domestic murders is quite small during this time period (6%, or 142 incidents).

Another test may be exploring whether these murders occurred in public spaces, where offenders may be more plausibly affected by police effort and/or presence. I define the murder as occurring in a public space if the location is a street, sidewalk, alley, parking lot, or vacant lot (64% of sample). Figure B.24b shows the results for this heterogeneity test are quite messy, with many significant indicators during the pre-treatment period. Despite that, it appears that the increases in murder occur in both public and private spaces, with suggestively larger increases for public spaces. These tests provide suggestive evidence for the

[45] See previous footnote to clear up confusion about how a decline of 2.1 log points results in a just over 83% drop.

[46] I focus specifically on murders since crime reporting may change drastically with the ACLU monitoring and the PoP investigation in addition to McDonald's death, rendering estimates of other crimes especially noisy.

theory that reductions in policing effort may be visible to offenders, driving larger increases in murders that may be more sensitive to police presence.

On the other hand, Figure 9d shows that the largest month-to-month increases in murders do not align with the steepest drops in low-level arrests or stops. The sustained significant rise in murder occurs in May 2016 (6 months after the video release), which to my knowledge cannot be explained by any developments in the criminal trial related to McDonald's death, any changes in the ACLU monitoring agreement, or the DOJ PoP investigation. I posit two non-mutually-exclusive theories that match the effect dynamics: (1) Lingering outrage from McDonald's death triggers additional public scrutiny after the high-profile police killing of Pierre Loury, a 16 year old Black teen, in April 2016. This incident generates over 3,500 news articles and results in hundreds of people protesting in Chicago. (2) The Chicago Police Accountability Task Force (PATF), set up to review CPD after McDonald's death, issues their report in April 2016, which plainly states that the "community's lack of trust in CPD is justified." The report provides over 76 recommendations across accountability, oversight, and training—nearly a third of which are immediately adopted by the Mayor and head of CPD despite noticeable pushback from the head of the union. These two events provide suggestive evidence that community outrage over high-profile police killings (and their subsequent revelations) may have propelled the lasting increase in murders. This may also provide an explanation for why the rise in murders occurs in private locations as well as public ones. Regardless of the exact driver, it is clear that the killing of McDonald and its subsequent cover-up contributed to a significant change in officer and offender behavior, imposing significant crime costs upon Chicago through sustained increases in murders—a cost borne nearly entirely by Black Chicagoans (Figure B.23).

6 Discussion

These results broadly show reductions in arrests for less serious offenses without changes in arrests for more serious ones. Conversely, there are large increases in violent crime and more moderate increases for property crime. The results seem to be driven by the differential timing of high-profile police killings in treated jurisdictions, as they are similar when the sample is limited to treated departments only (Figure B.12, Figure B.13). To understand how the impacts vary across municipality size, the empirical analysis can be conducted without the population control, weighting the results by municipality population size. The results are qualitatively similar, although there seems to be much larger variation in effects of low-level arrests in larger cities as evidenced by the larger confidence intervals and the arrests for property crime seem to be driven by larger municipalities (Figure B.30). The arrests for theft and marijuana possession slump more rapidly in larger cities, reaching reductions of up to 20% and 33%, respectively. The crime results are attenuated, finding significant yet smaller effects for violent crime (Figure B.31). This difference suggests that the increase in murder may be larger in relatively smaller jurisdictions, although the results are noisier generally.

To reconcile how to interpret the arrest results with increases in reported crime, I run two empirical tests that attempt to control intermediary changes in offending behavior and civilian reporting, providing an avenue to assess whether declines are occurring because of community cooperation and/or public scrutiny.[47] One preliminary check may be investigating whether the ratio of arrests to crime, or the clearance rate, is significantly reduced during this time period (Section B.5).[48] Naturally, given the significant increases in murders and the relative stability of those arrests, there is a significant reduction in the clearance rate, essentially producing the inverse of the crime figure (Figure B.25b). For similar reasons, there is a suggestive decline in the robbery clearance rate, and the decline in the theft clearance rate is more accentuated than the reduction seen in arrests. The drop in the murder clearance rate, which involves high marginal benefit arrests, suggests a potential decline in community cooperation with police (Equation 5). However, because the outcome is the log ratio of arrests to crime, this analysis assumes that (1) for a given rise in crime, there should be an equally proportional increase in arrests across offenses, and (2) that the proportion is one (e.g., a 1% increase in thefts results in a 1% increase in theft arrests, and the same is true for murders). However, if I more flexibly estimate this relationship by regressing log arrests on the event time indicators with log crime as a control, this relationship does not seem to hold, judging from the differing estimates on the crime control—none of which are one.

This is the preferred test of assessing differences in arrests amidst changes in reported crime. Section B.6 provides the arrest event study figures with controls for log reported crime, closing the intermediary channels of offending behavior or crime reporting affecting the arrest results.[49] The results with the crime controls demonstrate a qualitatively similar story as the main event study findings. Despite the increase in murders after a high-profile police killing, the arrests for murders do not significantly decline (after controlling for changes in reported crime, including murders), though the estimates are expectantly lower when controlling for crime than not. This case is broadly representative of a lack of significant reductions in more serious arrests, ultimately suggesting that significant changes in community cooperation with police are not causing changes in arrests. Given that reductions occur in less serious arrests, predominantly the low-level ones, the findings are consistent with the model's prediction

[47]Assuming arrests are (weakly) monotonically increasing with the amount of reported crime, it would still be a decline in policing effort if arrests are unchanged after rise in crime—all else constant.

[48]In these sets of regressions, I regress the clearance rate (arrests/crime) on the event time indicators, log population control, month-of-sample fixed effects, department fixed effects, and county-specific linear trends.

[49]An important caveat when interpreting these results is that crime is an endogenous control and may cause selection bias. Because the purpose of this empirical exercise is a mechanism analysis, I am less concerned about bias that limits causal interpretation. Specifically, I run Equation 11 with the added controls of log violent crime, property crime, and for Part I arrests, the specific offense. 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, which taken together provide the reported crime measures.

for public scrutiny being the causal mechanism (Equation 3). The model illustrates that the additional public 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 consequently, arrests fall for offenses that have low monetary, personal, and/or professional return.

Given recent legalization of the possession and (regulated) sale of marijuana in a handful of states, these accentuated estimates are internally consistent with less reward to the officer because of less cost to society. Importantly, none of the results are being driven by changes in state law, as the findings are unchanged after interacting the month-of-sample fixed effects with the state (Section B.2.2). Further, there are other low-level 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 have the least incentives to make arrest. Naturally, these are also the cases that are the most sensitive to officer effort because they 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-level arrests—where the effects are most concentrated—is instrumental in disentangling the mechanism as public scrutiny.

Moderate reductions in theft arrests—the least costly of Part I crimes (Table B.1)—could still be consistent with the theoretical prediction that public scrutiny drives declines in lower-level, more discretionary arrests. Similar to the low-level arrests, I find that theft arrests are primarily on-view, using data from the FBI's National Incident-Based Report System. Besides an on-view arrest, the next most common law enforcement response to a theft offense is a citation or summons to court (i.e., no custodial arrest)—as the far majority of offenses are shoplifting—highlighting their often less serious nature and the discretion an officer possesses. These aspects may explain why theft has a substantially lower cost of crime on average (Table B.1), and suggest that it is unlikely that civilian reporting or community cooperation are driving differences. Moreover, there are no significant drops in burglary, robbery, or motor vehicle theft arrests (Figure B.5)—all of which are more commonly driven by previous incident reports or warrants (i.e., civilian reporting and community cooperation) and have substantially higher social cost per crime. There is also no evidence of officers reducing arrests for the sale of heroine/cocaine and only suggestive declines in weapon violation arrests—two Part II offenses that have a higher return to the officer and more 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 public scrutiny being the relevant mechanism.

This notion is further bolstered after examining at the effect heterogeneity by media coverage by modifying the threshold of what fatalities are considered *high profile*. Figure B.32 illustrates that the general pattern of effects is robust across thresholds: no reduction in violent crime arrests, moderate decreases in property crime arrests (driven by declines in theft arrests), and sharp drops in low-level arrests. As the threshold for high-profile incidents increases from 1,000 articles of news coverage to 2,500 or 5,000, police killings with more public scrutiny drive larger declines in property and low-level arrests, where reductions from the 5,000 article threshold are twice that of the 1,000 article threshold. Though the effects are significantly larger for higher-profile OFs, they do not seem to be driven by any single incident, as the results do not noticeably change when excluding the two highest-profile incidents in the sample, which I discuss further below.

In the national analysis, the declines in effort are present across arrests of Black and White suspects, suggesting that public scrutiny generally increases the marginal cost of effort for policing (Figure B.34, Figure B.35, Figure B.36). These results provide evidence of a broader pullback of policing activity, rather than one solely reducing interactions with Black suspects or engagement in majority-Black neighborhoods. However, for theft and marijuana possession arrests, the effects for Black suspects are suggestively larger than White suspects (Figure 5). Notably, analyzing incident-level data from Chicago after the killing of Laquan McDonald, there are sharper changes in arrests and stops of Black civilians (Figure B.19, Figure B.20). The motivating anecdote of the Ferguson Effect may explain the race results: Officers may not want to make public, street-level arrests of Black individuals, potentially substituting their time to patrolling in their vehicle.

The focus on street-level clearances may also provide clarity in understanding the persistence of the decline in certain low-level arrests, such as marijuana possession or liquor violations (Figure 4). 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 3). The prolonged nature of the effects contradict the theory that the declines in low-level arrests is caused by the reassignment of police officers for crowd control at protests. The sustained transition to a lower equilibrium effort may be a result of officers enduring persistent increases in psychic costs or perceived-error costs in street-level interactions for these low-level clearances. If these factors become too costly, they could also cause increases in retirements or resignations, as officers leave the police department involved in the high-profile police killing—an anecdotal phenomenon described in news coverage (Westervelt, 2021).

When examining DUI, marijuana sale, or theft arrests, the low point in officer effort is around Q4, before the effect abates toward the original equilibrium. The low point occurring at 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 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, not to mention convicted (Stinson, 2021)—the sample is unique in that I subset to only high-profile events. After these updates on court proceedings, public scrutiny may intensify to its fever pitch before subsiding, eventually resulting in a partial return to equilibrium effort for these relatively lower-level, non-violent offenses with higher social costs than other low-level arrests. Moreover, it is plausible that scrutiny pushes officers to leave the involved department, as officers opt for environments where they face less costs (Mourtgos, Adams, and Nix, 2022). Correcting this recruitment and retention issue likely takes time. Finally, the effect abatement for these offenses could partially represent a shift from more public pedestrian encounters to vehicular stops—as was the case in Chicago after the killing of Laquan McDonald (see Section 5.6 for more details). The theory of substitution from pedestrian to vehicular stops may also be consistent with why there is no reduction in motor vehicle theft arrests and may potentially even explain the (statistically insignificant) declines in weapon violation arrests (Figure B.5).

To further explore the dynamics of public scrutiny on police and offending behavior, I consider the case study of Laquan McDonald's death, where there is a the year gap between the incident and community awareness (discussed in detail in Section 5.6). Given the initial misinformation about the case, the use of force was considered justified and there was a lack of awareness (and scrutiny) of the event as evidenced by Google Trends (Figure B.18). Subsequently, there was minimal change in low-level arrests (Figure 9a). However, after the Chicago Police Department was forced to release a video of the incident, there was a surge in public scrutiny—measured by local Google searches of the incident, protests, and news articles. Low-level arrests discontinuously fell by 11% that month, despite it occurring in the last week of November. In the ensuing year, low-level arrests faced precipitous and growing declines of 22–36%, mirroring the results in the national analysis. Similarly, there are significant reductions in pedestrian and vehicular stops of up to 44% and 12% respectively right after the video release (Figure 9b, Figure 9c). Theoretically, a larger reduction in pedestrian stops is consistent with these street-level interactions likely being more affected by public scrutiny. In totality, the sharp and persistent reductions in low-level arrests and overall stops coinciding precisely with community awareness—despite a year passing since McDonald's death—provides cogent evidence that further substantiates public scrutiny as the key mechanism. The case study also helps negate other potential explanations, such as a police response to the incident itself absent any public scrutiny.

Nevertheless, recent papers have attributed reductions in policing effort in municipalities such as Chicago are the result of oversight of ACLU monitoring or Department of Justice pattern-or-practice (PoP) investigations, rather than solely public scrutiny (Cassell and Fowles, 2018; Devi and Fryer, 2020). Naturally, this is difficult to test because public scrutiny of policing—occasionally in response to high-profile police killings—often play a role in instigating PoP investigations or court-ordered monitoring agreements.[50] Of the 30 PoP

[50] Although these legal interventions arguably are an extension or culmination of public scrutiny, it is still valuable to consider whether they are driving the stated effects.

investigations between 2005–2016, six were plausibly induced by high-profile OFs,[51] including the two highest-profile incidents in the analysis—the deaths of Freddie Gray in Baltimore and Michael Brown in Ferguson.[52] After removing the departments that experienced PoP investigations potentially as a result of high-profile police killing, the main results are remarkably similar (Section B.10.1). To be more restrictive, I also test robustness to excluding all agencies that experience any PoP investigation from 2004–2016, and again find results that match the main findings (Section B.10.2).[53] Moreover, as a consequence of the public outcry following the video release of Laquan McDonald's death (November 2015), there was a PoP investigation into the Chicago Police Department that began in January 2016. That month also happened to coincide with the start of an unrelated police monitoring agreement with the ACLU. However, there is a significant drop in low-level arrests, and pedestrian and vehicular stops in November and December 2015 (months 0 and 1, Figure 9), *before* the investigation or the monitoring took place. This suggests that public scrutiny plays a pivotal role in the reduction of policing effort—separate from these outside monitoring/investigative initiatives.

Despite arrest reductions primarily occurring in less serious crimes, the national analysis reveals substantial increases in *serious offenses* in the aftermath of these high-profile, officer-involved fatalities (Figure 6): For violent crimes, Figure 7 documents a generally statistically significant rise of 11–18% in robberies and murders for Q0–Q4. There is also evidence of smaller (and occasionally significant) increases in property crime, driven by theft (Figure 8). Figure B.33 provides evidence that these results are robust to increasing the threshold for the "high-profile" treatment inclusion. The highest-profile deaths—ones that generate at least 5,000 articles of coverage—draw even greater increases of 31% in murder, 12% in aggravated assault, and 13% in burglary (Table B.15, Table B.16). This increase in crime does not seem to be driven by any single death, nor does the decline in arrests: When I narrow the sample to agencies who never experience a PoP investigation (Section B.10.1), which removes the two highest-profile deaths from the analysis sample (Michael Brown and Freddie Gray), the results are quite similar.

The killing of Laquan McDonald also resulted in increases in murders of 5–14% in Chicago (Figure 9d), but the effects were not immediate—unlike what was witnessed for low-level arrests or stops in Chicago (or for murder in the national analysis). This lag makes it especially tough to separate whether the increase in murder is related to the cover-up around McDonald's killing, the PoP investigation, or the ACLU monitoring. However, the Chicago

[51] The Department of Justice typically does not comment on what prompts their investigations.

[52] PBS Frontline provides data on PoP investigations here, including the allegations against the department, the start date, and the outcome of the investigation. Because of a lack of consistent reporting of arrest and crime data to the FBI, only two of the six police departments whose high-profile police killings likely induced PoP investigations are in the analysis sample.

[53] The results using the event study specification are also robust to excluding the PoP agencies, and are even more similar than the DD results, likely because the event study design has more police killings in the treatment sample. Only the DD results are shown for brevity.

murder findings are quantitatively similar to results from the national analysis—the latter of which is robust to excluding all cities involved in PoP investigations, suggesting it is likely that the increase at least partially stems from McDonald's death. The increase in murders is nearly exclusively in Black victims (Figure B.23). This concentration suggests that the Black community is doubly burdened by being overrepresented in high-profile profile killings at 74% (Table 2) as well as the recipient of much of the crime consequences.

The additional crimes, especially the rise in murder, impose tremendous costs on affected jurisdictions (Table B.1). This concern is heightened when considering the possibility that other offenses, although less serious, may be underreported after a high-profile police killing (Desmond, Papachristos, and Kirk, 2016). However, because community awareness of the incident and the subsequent public scrutiny drive the reductions of officer effort, it is difficult to identify the determinant of the offending response: is it a response to the reduction in policing effort or outrage at the perceived unjust nature of the police killing itself? For the former theory, decreases in arrests for primarily lower-level offenses—many of which involve public, street-level interactions—would be internalized by marginal offenders as a signal of lower apprehension risk, decreasing the expected costs from engaging in crime (Becker, 1968). Alternatively, marginal offenders may be reacting to the incident itself and subsequent (lack of) legal proceedings, where diminished perceptions of police legitimacy and procedural justice lead to increases in legal cynicism and estrangement—resulting in people *taking the law into their own hands* (Persico, 2002; Kirk and Papachristos, 2011; Leovy, 2015; Bell, 2017; Rosenfeld and Wallman, 2019).

I find suggestive evidence that marginal offenders may be responding to policing effort. This theory becomes more plausible if low-level arrests signify police presence to marginal offenders or is at least correlated with it, which has been shown to be effective in reducing crime (Weisburd, 2021). Plausibly substantiating that theory, the decrease in stops coincides with the drop in low-level arrests *right after* the video release of the murder of Laquan McDonald. The results from the national analysis are quite consistent with Evans and Owens (2007), Chalfin and McCrary (2018), and Mello (2019), which find that violent crime—particularly robbery and murder—is more sensitive to fluctuations in policing than property crime. In Section B.8, I provide evidence that municipalities that experience higher-profile police killings also draw larger reductions in low-level arrests as well as larger increases in violent crime. (Although, this artifact could alternatively be explained by stating that higher-profile incidents draw more legal cynicism, and the reductions in effort are incidental.) Moreover, the increase in murders in Chicago after the McDonald video release were arguably concentrated in situations that may be more affected by police effort and presence, such as increases in public spaces more than privates ones (Figure B.24a, Figure B.24b).

However, in the national analysis, Figure 7 shows discontinuous increases in robberies and murders that occur immediately after the police killing becomes salient to the local community, whereas low-level arrests are continually (yet steeply) reduced each quarter from the death (Figure 2). In fact, low-level arrests are not even significant reduced in Q0. Moreover, the rise in murders abates after five quarters from the incident, where low-level arrests reach their lowest level. Since changes in murder do not exactly coincide with decreases in police effort, this evidence may suggest that outrage and legal cynicism from the incident does play a role in contributing to increases in (violent) offending. This is also bolstered by evidence from the McDonald case study, which find that the largest month-to-month increases in murders do not align with the steepest drops in low-level arrests or stops. Additionally, the delayed increase in murder was likely caused by another high-profile police killing in Chicago after which police effort did not substantially change (see end of Section 5.6), and this increase occurred in private spaces as well, although to a lesser extent.

7 Conclusion

Amidst an unprecedented rise in civil unrest touched off by the murder of George Floyd (Buchanan, Bui, and Patel, 2020), the national discourse has turned a spotlight on issues in policing, and in particular, police use of force. There are over 1,100 officer-involved fatalities per year, meaning that 7–9% of homicides (broadly defined) in the US involve the police—a few of which have sparked the largest demonstrations of civil unrest in recent history. Coinciding with these protests (among many other factors related to the COVID-19 pandemic), the US experienced the largest recorded increase in murders (Asher, 2021), creating a pressing need to understand if there is any connection between these patterns.

This paper provides systematic evidence of how high-profile police killings affect the arresting patterns for the involved police department and crime in that jurisdiction. To credibly identify the causal effect of public scrutiny on police officer effort, I provide the first national analysis that exploits the plausibly exogenous timing of these fatalities. I measure scrutiny through community awareness, media coverage, and local protests of OFs. 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 several different mechanisms affecting arrests, such as (1) greater scrutiny of police, (2) reduced community cooperation in the clearance of 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 of 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.

For the empirical estimation, 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 police killings, their associated media coverage, and the timing of community awareness. The empirical strategy directly leverages aspects of scrutiny by only including incidents that reach over 1,000 articles of media coverage ("high profile") and adjusting treatment timing to when the local community first searches for the incident, as measured by Google Trends. To further validate that this "high profile" measure is related to other aspects of public scrutiny, I show that treated municipalities experience larger and more frequent protests.

I complement the national analysis with a detailed analysis of how killing of Laquan McDonald and subsequent cover-up by the Chicago Police Department affected the arrests, stops, and crimes in Chicago. McDonald's death is one of the highest-profile police killings in the sample, and the incident is notable since there is a year gap between his death and broad community awareness.

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 increased public scrutiny, officers experience higher marginal cost of effort and curtail policing activity, yielding fewer arrests for lower-level offenses where officers typically have more discretion. Theft arrests experience temporary reductions of 3–10%, while there are persistent declines of up to 21% in low-level arrests. The most marked change is in marijuana possession arrests, which experience a drop of up to 28%. Similarly, sharp reductions can be seen across a handful of low-level arrests, including disorderly conduct, liquor violation, and marijuana sale arrests. These low-level arrests are the most sensitive to officer effort because they usually result from officer-initiated stops, where they have the most discretion in determining whether to investigate and/or intervene. The sustained reduction of at least 1.5 years likely negates the possibility that these effects are driven by officers being re-assigned to crowd control duties for protests. Notably, these reductions are not present with more serious arrests, demonstrated by the insignificant changes in violent crime or more serious property crime clearances. After higher-profile police killings, the drops in policing activity are significantly larger. These results are corroborated in the McDonald case study, where the incident-level data reveal that the Chicago Police Department had discontinuous reductions in low-level arrests and stops *only after* the incident began broadly aware to the public. This finding excludes the possibility that the changes are driven by Chicago Police Department addressing the incident itself internally absent public scrutiny.

These results provide evidence of a broader pullback of policing activity, rather than one solely focused on reducing interactions with Black suspects or engagement in majority-Black neighborhoods. 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 suspects. In Chicago, after the killing of McDonald, there are sharper changes in the arrests and stops of Black civilians.

Despite reductions primarily occurring in less serious arrests, the national analysis reveals substantial increases in *serious offending.* Most notably, there is a significant rise of 11–18% in murders and robberies. There are also smaller, suggestive increases in motor vehicle theft and theft of up to 7% and 5% respectively. The violent crime effects, particularly for murder, are more prominent for the most publicized deaths. The spike in crime imposes significant crime costs on the affected jurisdictions. In Chicago, I show that the increase in murders almost entirely afflicts Black victims. Using heterogeneity tests from both the national analysis and the McDonald case study, I find suggestive evidence that marginal offenders are *both* responding to a reduction in police effort and reacting to the high-profile police killing itself.

The arrest and crime findings are robust to numerous changes in empirical specification, transformations of the dependent variable, varying levels of fixed effects that control for changes in state law and treatment spillovers, and increasing the high-profile threshold. I find that the results are not driven by a single high-profile police killing. I also do not find that these high-profile incidents result in sizable spillover effects *on average*, although there still may be spillovers for the highest-profile cases, as was found for the death of Michael Brown (Cheng and Long, 2018). Finally, I present evidence that suggests these effects are not driven by a policy mechanism such as a pattern-or-practice investigation or a court-mandated monitoring agreement.

These high-profile police killings are substantial enough events that involved agencies face significant increases public scrutiny that reduce more discretionary police behavior in the form of low-level arrests and stops. These changes in police effort are internalized by marginal offenders, resulting in part of the witnessed increase in violent crime. Law enforcement agencies should consider research-based interventions that mitigate the additional marginal costs imposed by public scrutiny, particularly the psychic costs associated with loss of morale. Although these field experiments did not have media coverage and protests of policing, Linos, Ruffini, and Wilcoxen (2021) and Linos and Harney (2022) show that low-cost, behavioral interventions such as nudging 911 dispatchers and correctional officers to anonymously share their experience and advice in an online platform helps reduce burnout and PTSD. This increased belonging improves service delivery and reduces resignations. To the extent that police staffing is contributing, law enforcement agencies may also turn to cash incentives to recruit and retain officers to alleviate the impact of departures—as documented in a few municipalities (Westervelt, 2021). It is possible that these types of interventions may attenuate the reductions in policing effort, thereby limiting any increases in violent crime.

However, the spike in violent offending is also driven by a reaction to these high-profile police killings themselves—likely as a result of their perceived unjust nature, where a majority of incidents involve an unarmed civilian. As communities turn away from law enforcement as a means of protection and an arbiter to address conflict, murders may increase as people resort to vigilantism (Kirk and Papachristos, 2011; Leovy, 2015; Bell, 2017). Thus, interventions aimed at insulating officers from some of the negative effects of public scrutiny will not fully resolve this phenomenon. Consequently, law enforcement agencies should especially prioritize policies that reduce police use of force, mechanically reducing the probability of these high-profile incidents occurring in the first place. Recent literature have highlighted a handful of promising avenues, such as carefully selecting field training officers (Adger, Ross, Sloan, 2022), diversifying the police department (Ba et al., 2021; Hoekstra and Sloan, 2022), increasing the minimum age of officers (Ridgeway, 2020), and procedural justice training (Owens et al., 2018; Woods, Tyler, Papachristos, 2020).

References

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.

Adger, Chandon, Matthew Ross, and CarlyWill Sloan. 2022. "The Effect of Field Training Officers on Police Use of Force." Working Paper, January.

Ang, Desmond. 2021. "The Effects of Police Violence on Inner-City Students." *The Quarterly Journal of Economics* 136 (1): 115–68. https://doi.org/10.1093/qje/qjaa027.

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.

Asher, Jeff. 2021. "Murder Rose by Almost 30% in 2020. Its Rising at a Slower Rate in 2021." *The New York Times*, September 22, 2021, sec. The Upshot. https://www.nytimes.com/2021/09/22/upshot/murder-rise-2020.html.

Ba, Bocar A., Dean Knox, Jonathan Mummolo, and Roman Rivera. 2021. "The Role of Officer Race and Gender in Police-Civilian Interactions in Chicago." *Science* (New York, N.Y.) 371 (6530): 696–702. https://doi.org/10.1126/science.abd8694.

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.

Bell, Monica C. 2017. "Police Reform and the Dismantling of Legal Estrangement." *The Yale Law Journal* 126: 2054–2150.

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.

Buchanan, Larry, Quoctrung Bui, and Jugal K. Patel. 2020. "Black Lives Matter May Be the Largest Movement in U.S. History." *The New York Times*, July 3, 2020, sec. U.S. https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html.

Cassell, Paul G., and Richard Fowles. 2018. "What Caused the 2016 Chicago Homicide Spike: An Empirical Examination of the ACLU Effect and the Role of Stop and Frisks in Preventing Gun Violence." *University of Illinois Law Review* 2018 (5): 1581–1684.

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

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.

Chicago Police Department. "Public Arrest Data- Chicago Police Department." Accessed August 2, 2019. https://home.chicagopolice.org/statistics-data/public-arrest-data/.

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.

Comey, James. 2015. "Law Enforcement and the Communities We Serve: Bending the Lines Toward Safety and Justice." Speech. Federal Bureau of Investigation, October 23, 2015. https://www.fbi.gov/news/speeches/law-enforcement-and-the-communities-we-serve-bending-the-lines-toward-safety-and-justice.

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.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96. https://doi.org/10.1257/aer.20181169.

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.

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.

Edwards, Frank, Hedwig Lee, and Michael Esposito. 2019. "Risk of Being Killed by Police Use of Force in the United States by Age, Race-Ethnicity, and Sex." *Proceedings of the National Academy of Sciences* 116 (34): 16793–98. https://doi.org/10.1073/pnas.1821204116.

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

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.

———. 2020. "A Response to Steven Durlauf and James Heckman." *Journal of Political Economy*, July. https://doi.org/10.1086/710977.

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

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

Goodman-Bacon, Andrew. 2018. "Difference-in-Differences with Variation in Treatment Timing." Working Paper 25018. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w25018.

Grunwald, Ben, and John Rappaport. 2020. "The Wandering Officer." *Yale Law Journal*

129 (6): 1600–1945.

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.

Heckman, James J., and Steven N. Durlauf. 2020. "Comment on 'An Empirical Analysis of Racial Differences in Police Use of Force' by Roland G. Fryer Jr." *Journal of Political Economy*, July. https://doi.org/10.1086/710976.

Hoekstra, Mark, and CarlyWill Sloan. 2022. "Does Race Matter for Police Use of Force? Evidence from 911 Calls." *American Economic Review* 112 (3): 827–60. https://doi.org/10.1257/aer.20201292.

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.

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.

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/.

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.

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.

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.

Leovy, Jill. 2015. *Ghettoside: A True Story of Murder in America*. Reprint Edition. One World.

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.

Linos, Elizabeth, and Jessie Harney. 2022. "Understanding Burnout in Correctional Officers." Working Paper, March.

Linos, Elizabeth, Krista Ruffini, and Stephanie Wilcoxen. 2021. "Reducing Burnout and Resignations among Frontline Workers: A Field Experiment." Working Paper, July.

https://doi.org/10.2139/ssrn.3846860.

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.

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.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do Political Protests Matter? Evidence from the Tea Party Movement." The Quarterly Journal of Economics 128 (4): 1633–85. https://doi.org/10.1093/qje/qjt021.

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.

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.

Mourtgos, Scott M., Ian T. Adams, and Justin Nix. 2022. "Elevated Police Turnover

Following the Summer of George Floyd Protests: A Synthetic Control Study." *Criminology & Public Policy* 21 (1): 9–33. https://doi.org/10.1111/1745-9133.12556.

Murgado, Amaury. 2012. "Handling DUI Stops." *Police Magazine*. February 17, 2012. https://www.policemag.com/340705/handling-dui-stops.

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.

Ornaghi, Arianna. 2019. "Civil Service Reforms: Evidence from U.S. Police Departments." Working Paper, July, 55.

Ouss, Aurlie, 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, and Bocar Ba. 2021. "The Economics of Policing and Public Safety." *Journal of Economic Perspectives* 35 (4): 3–28. https://doi.org/10.1257/jep.35.4.3.

Owens, Emily, David Weisburd, Karen L. Amendola, and Geoffrey P. Alpert. 2018. "Can You Build a Better Cop?" Criminology & Public Policy 17 (1): 41–87. https://doi.org/10.1111/1745-9133.12337.

PBS Frontline. "Fixing the Force." Accessed November 10, 2020. https://www.pbs.org/wgbh/frontline/interactive/fixingtheforce/.

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

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.

Pyrooz, David C., Scott H. Decker, Scott E. Wolfe, and John A. Shjarback. 2016. "Was There a Ferguson Effect on Crime Rates in Large U.S. Cities?" *Journal of Criminal Justice* 46 (September): 1–8. https://doi.org/10.1016/j.jcrimjus.2016.01.001.

Ridgeway, Greg. 2020. “The Role of Individual Officer Characteristics in Police Shootings.” *The ANNALS of the American Academy of Political and Social Science* 687 (January): 58–66. https://doi.org/10.1177/0002716219896553.

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

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.

Rushin, Stephen. 2020. “Police Arbitration.” SSRN Scholarly Paper ID 3654483. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=3654483.

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

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.

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

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.

Stinson, Philip. 2021. C-SPAN.org: Philip Stinson on His Research on Police Misconduct and Prosecutions. https://www.c-span.org/video/?510567-5/washington-journal-philip-stinson-discusses-research-police-misconduct-prosecutions.

Sun, Liyang, and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics*, December. https://doi.org/10.1016/j.jeconom.2020.09.006.

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.

Weisburd, Sarit. 2021. "Police Presence, Rapid Response Rates, and Crime Prevention." *The Review of Economics and Statistics* 103 (2): 280–93. Link.

Westervelt, Eric. 2021. "Cops Say Low Morale And Department Scrutiny Are Driving Them Away From The Job." NPR, June 24, 2021, sec. Throughline. Link.

Wolfers, Justin. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96 (5): 1802–20. https://doi.org/10.1257/aer.96.5.1802.

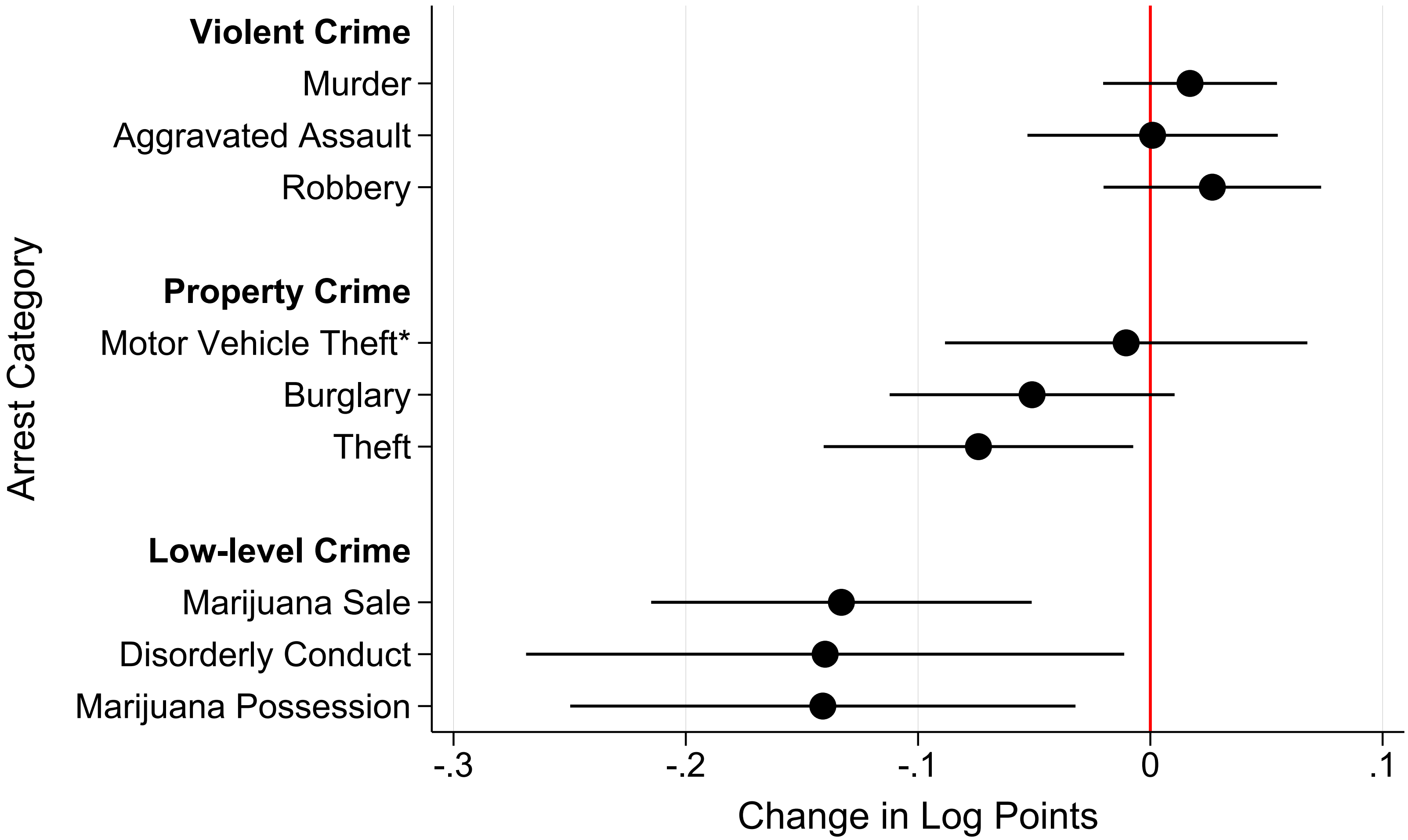
Wood, George, Tom R. Tyler, and Andrew V. Papachristos. 2020. "Procedural Justice

Training Reduces Police Use of Force and Complaints against Officers." *Proceedings of the National Academy of Sciences* 117 (18): 9815–21. https://doi.org/10.1073/pnas.1920671117.

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.

8 Figures

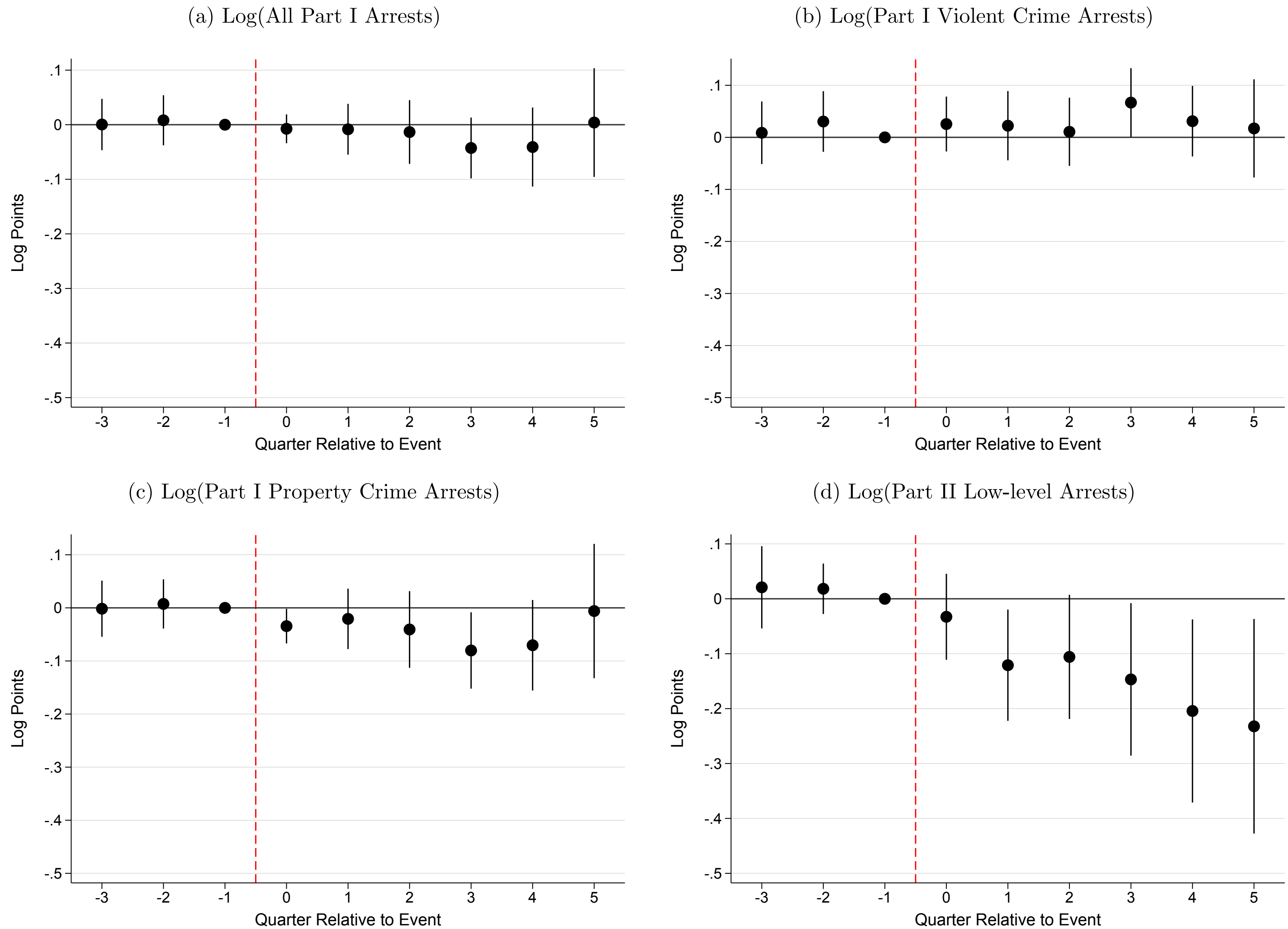
Figure 1: Difference-in-differences (DD) Coefficient Estimates by Arrest Category



*There is a significant difference between the average pre-treatment trends of control and treated departments for motor vehicle theft arrests, but no others.

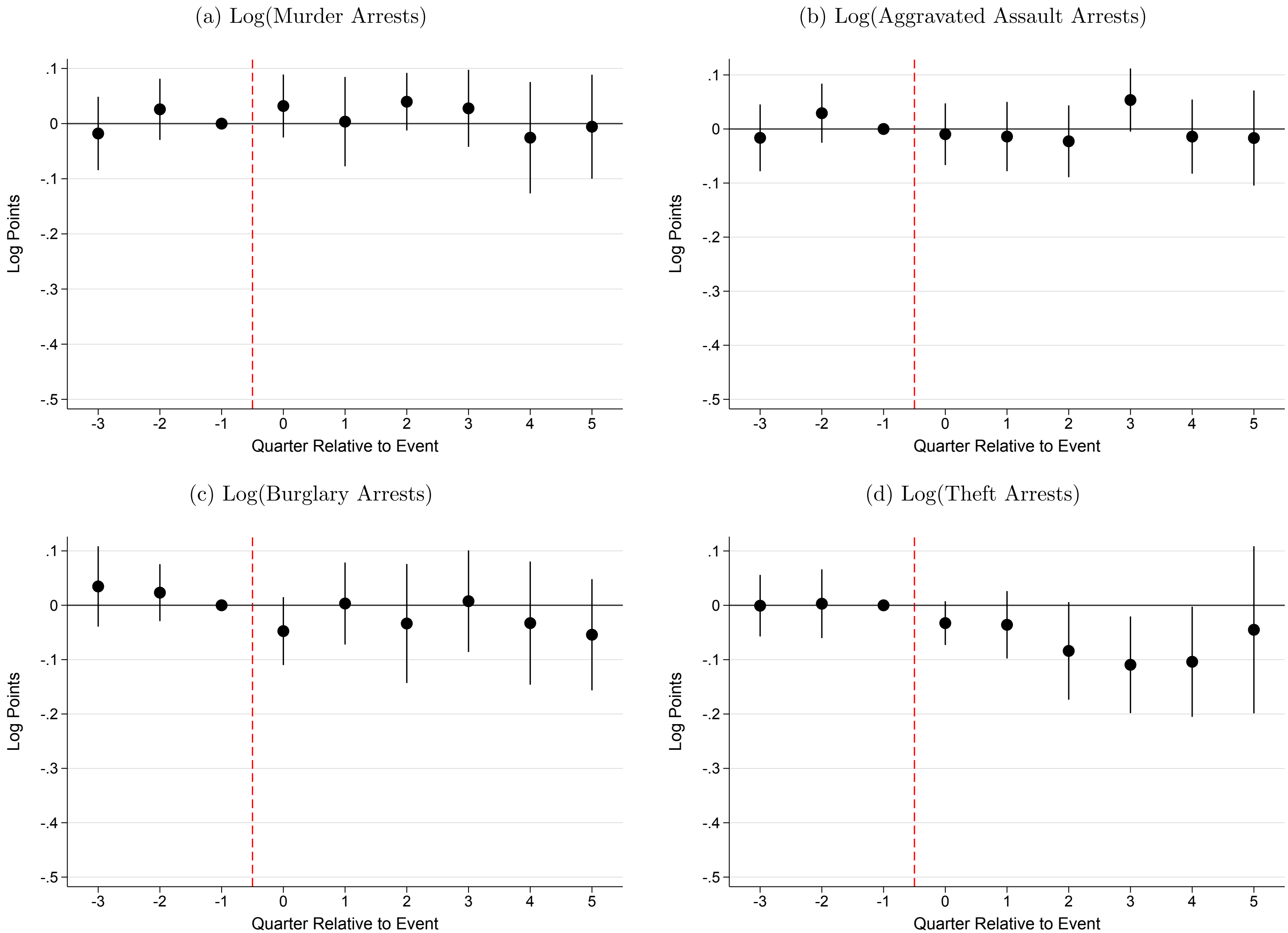
Circles display DD coefficients from separate regressions—in descending order of the social cost of crime—using a sample of city police departments with fewer than 9 outliers and a population greater 10,000. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. The arrest DD tables begin with Table B.2. [740,838 observations; 52 treated; 2,687 control agencies]

Figure 2: Effect of High-Profile Police Killing on Arrest Categories



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. [598,788 observations; 52 treated; 2,687 control agencies]

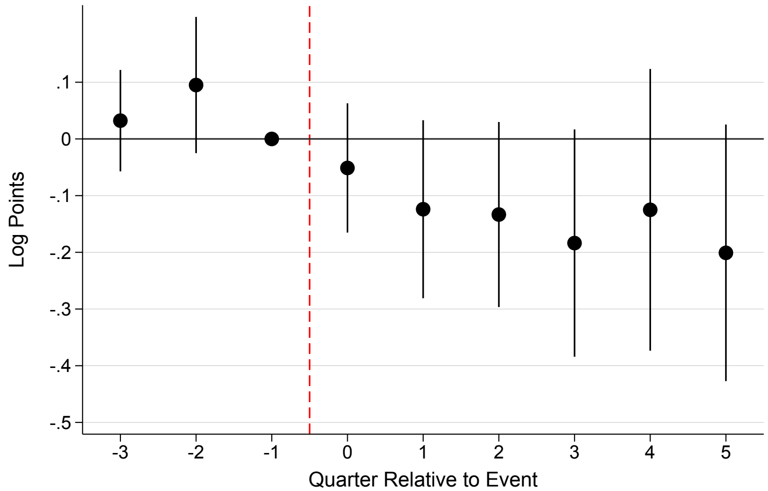
Figure 3: Effect of High-Profile Police Killing on Violent and Property Crime Arrests



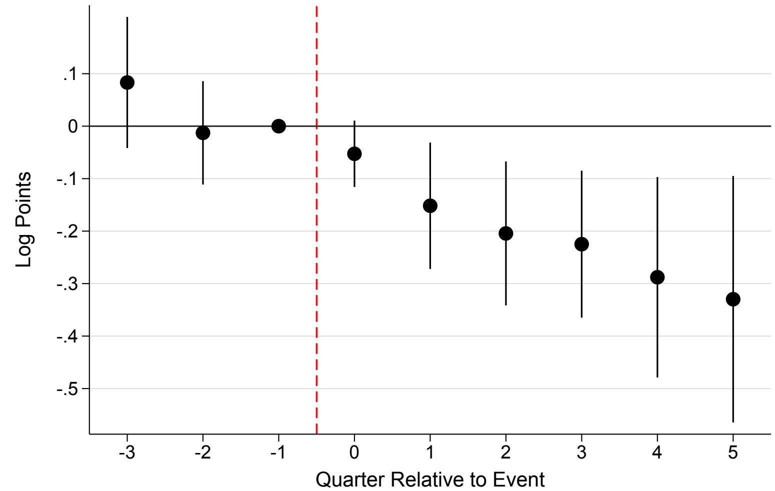
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard 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. [598,788 observations; 52 treated; 2,687 control agencies]

Figure 4: Effect of High-Profile Police Killing on Low-level Arrests

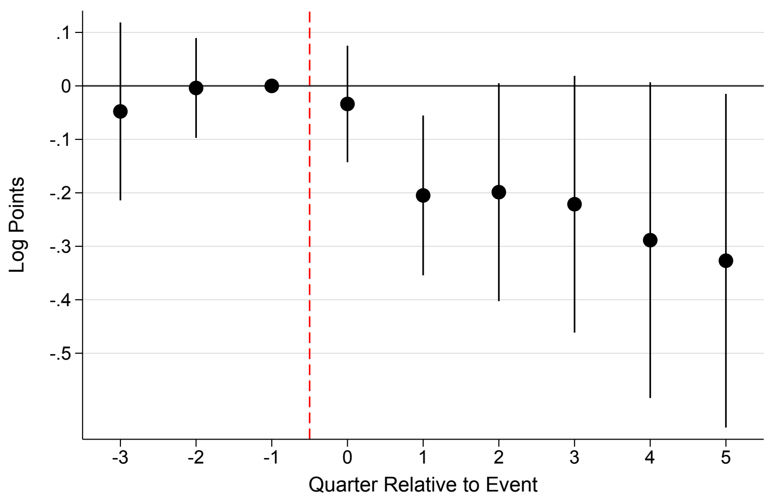
(a) Log(Disorderly Conduct Arrests)



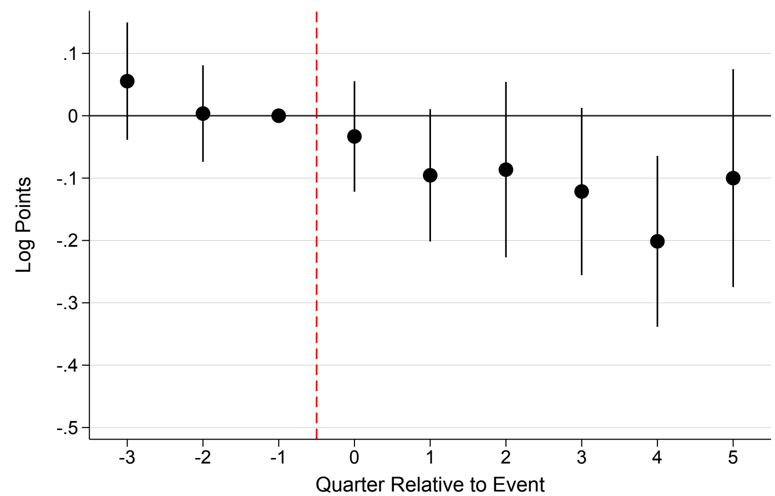
(b) Log(Marijuana Possession Arrests)



(c) Log(Liquor Law Violation Arrests)



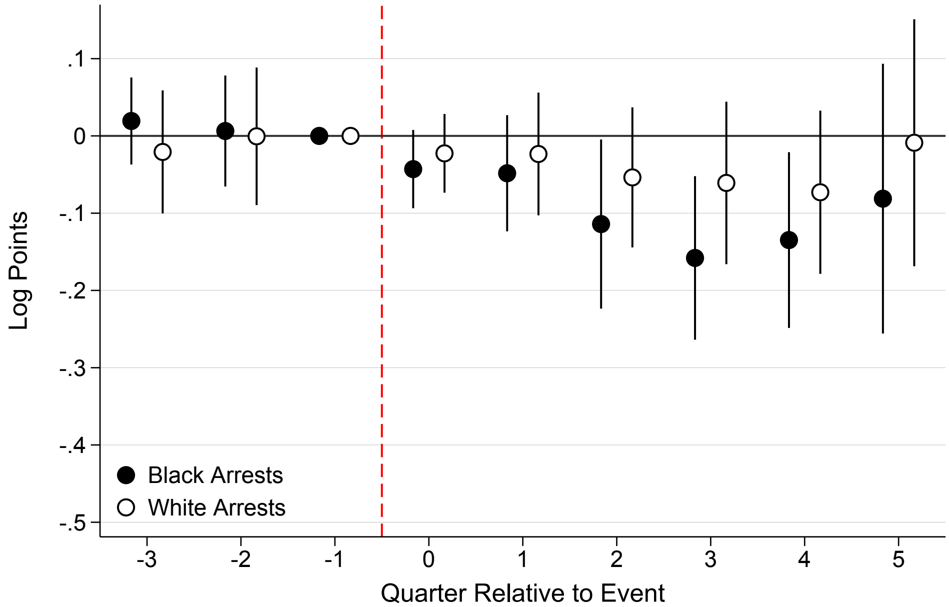
(d) Log(Marijuana Sale Arrests)



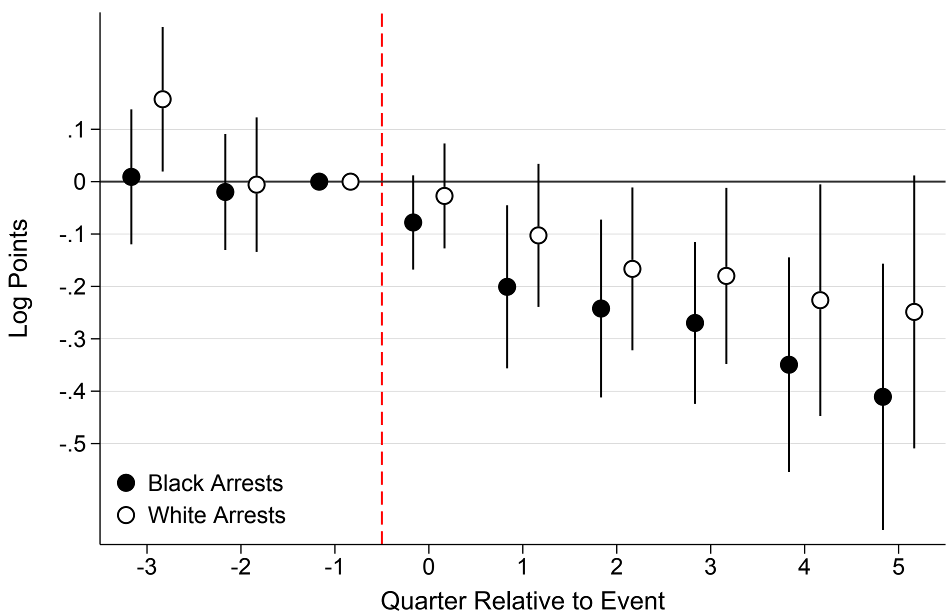
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,964 observations; 52 treated; 2,687 control agencies]

Figure 5: Effect of High-Profile Police Killing on Race-Specific Arrests

(a) Log(Theft Arrests)

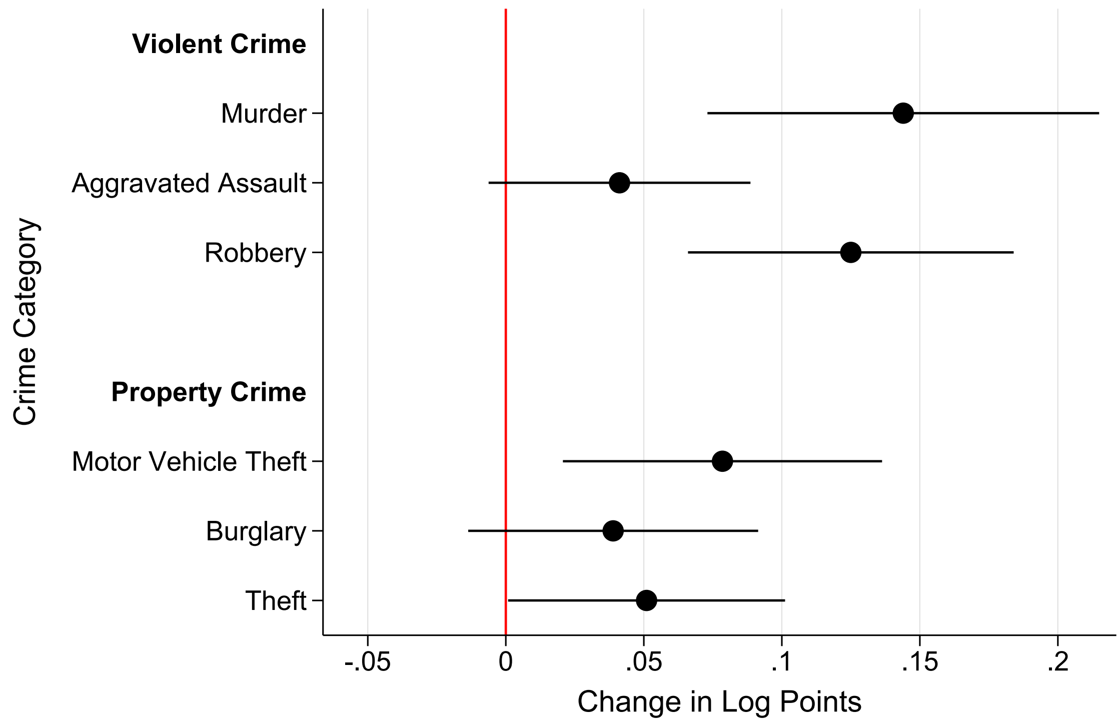


(b) Log(Marijuana Possession Arrests)



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. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which are interacted with race indicators. [598,788 observations; 52 treated; 2,687 control agencies]

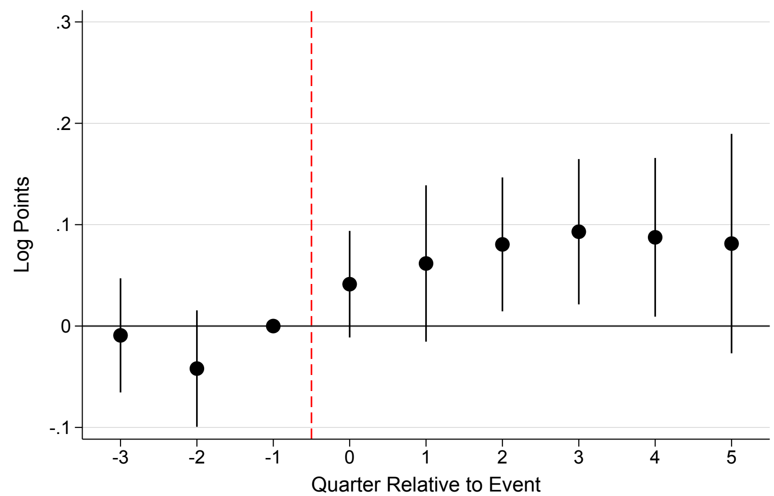
Figure 6: Crime Analysis: Difference-in-differences (DD) Coefficients by Crime Category



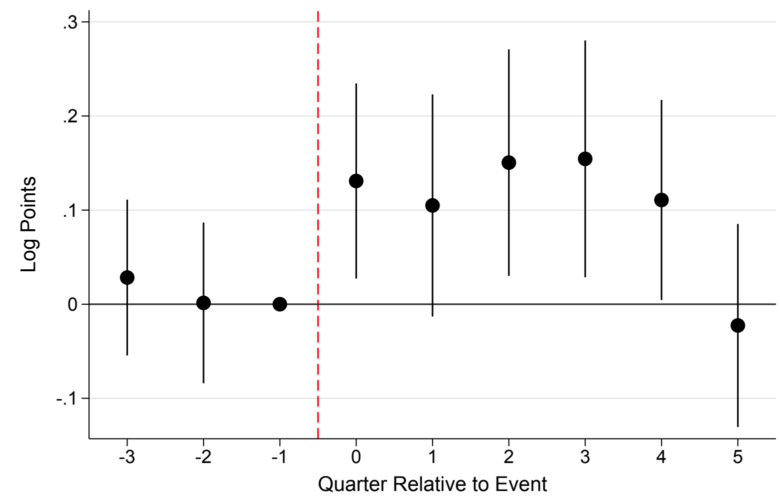
Circles display DD coefficients from separate regressions—in descending order of the social cost of crime—using a sample of city police departments with fewer than 9 outliers and a population greater 10,000. There are no significant differences between the average pre-treatment trends of control and 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. The crime DD tables begin with Table B.5. [740,838 observations; 52 treated; 2,687 control agencies]

Figure 7: Crime Analysis: Effect of High-Profile Police Killing on Violent Crime

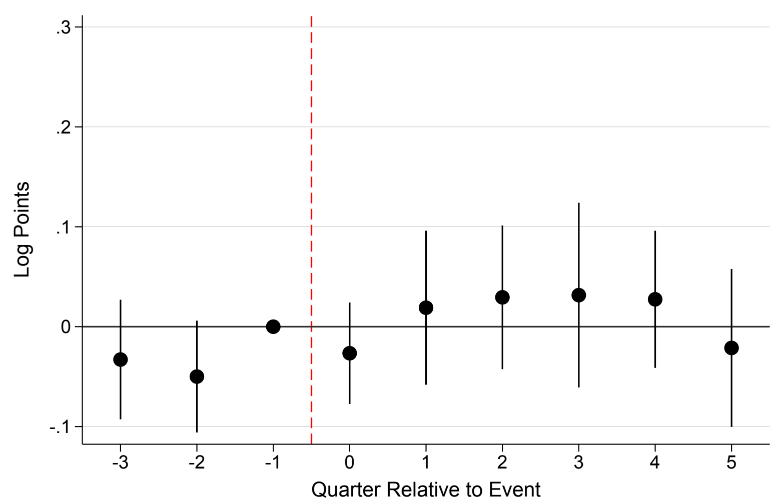
(a) Log(Violent Crime)



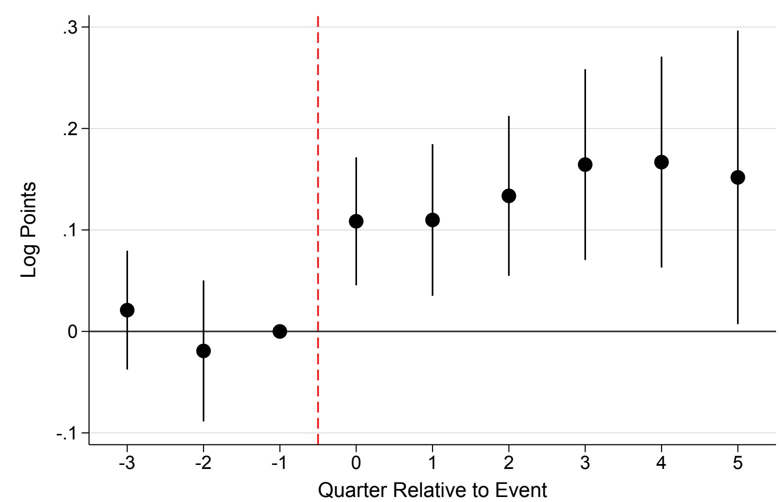
(b) Log(Murder)



(c) Log(Aggravated Assault)

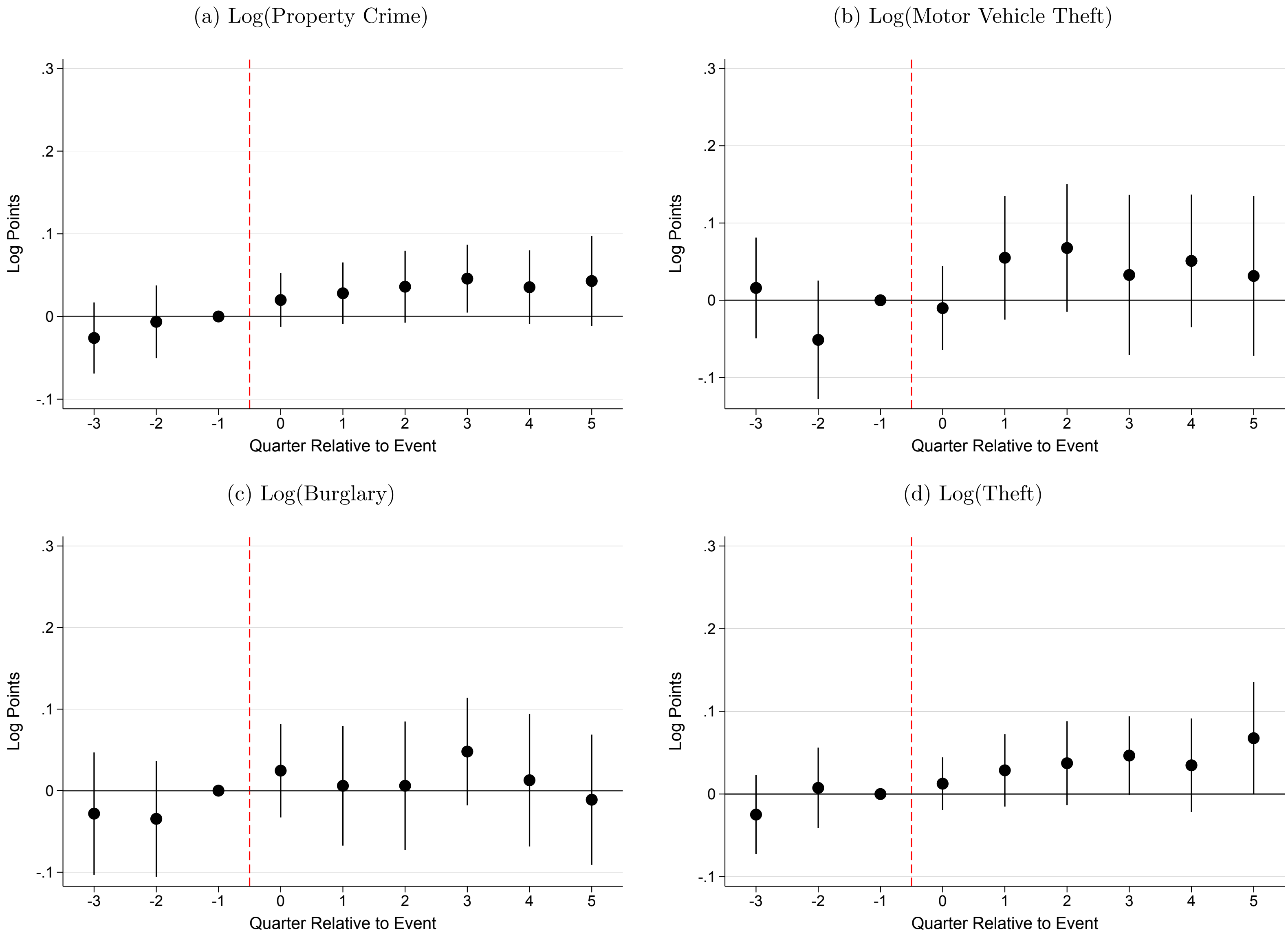


(d) Log(Robbery)



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. [598,788 observations; 52 treated; 2,687 control agencies]

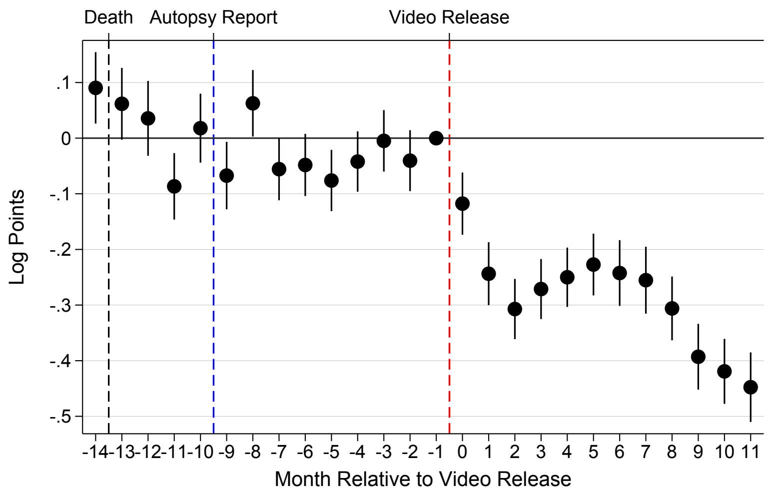
Figure 8: Crime Analysis: Effect of High-Profile Police Killing on Property Crime



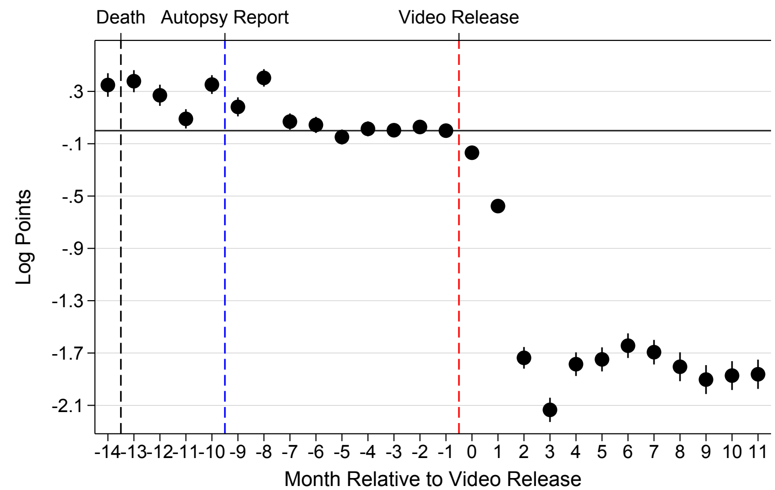
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in crime after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard 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. [598,788 observations; 52 treated; 2,687 control agencies]

Figure 9: Effect of Laquan McDonald's Death on Low-level Arrests, Stops, and Murders

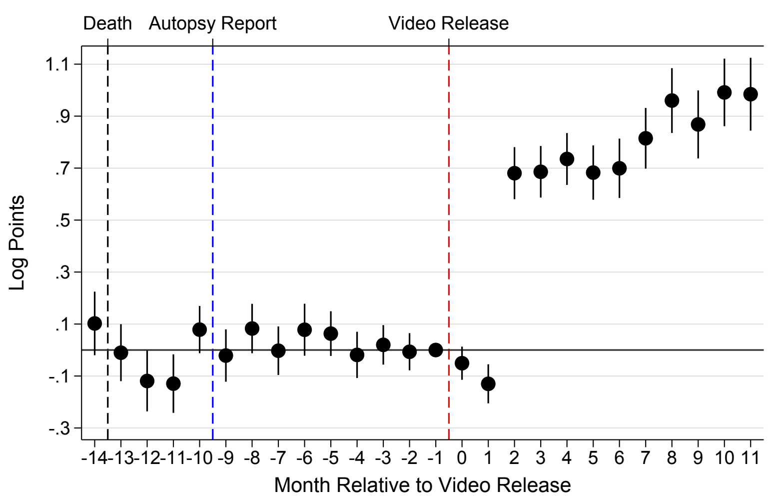
(a) Log(Low-level Arrests)



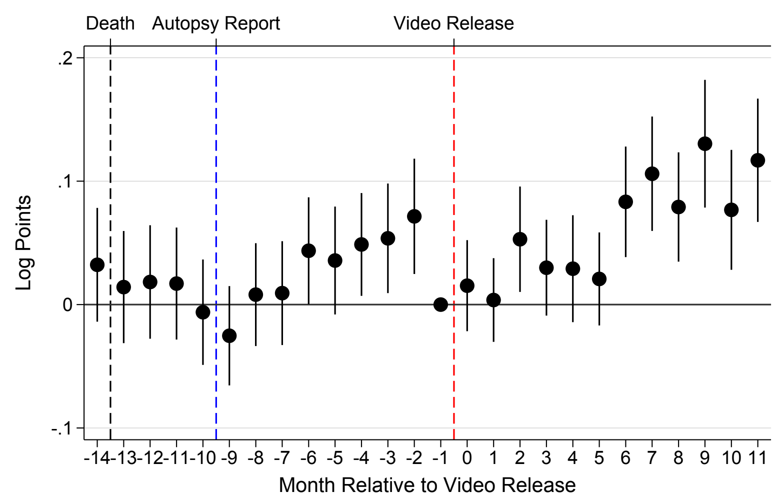
(b) Log(Pedestrian Stops)



(c) Log(Vehicular Stops)



(d) Log(Murders)



The black line is the murder of Laquan McDonald (Oct 2014). The blue line is when the autopsy report results became public (Feb 2015). The red line is the public release of the video of McDonald's death—first broad community awareness of the incident (Nov 2015). In part (a–c), circles display monthly event time coefficients from a regression of log low-level arrests, pedestrian stops, or vehicular stops at the beat-month-race level in Chicago from 2014–2017, while part(d) involves a regression of log murders at the beat-month level. Lines represent the 95% confidence interval using standard errors clustered at the beat level. I control for beat-level fixed events and linear trends.

9 Tables

Table 1: Mean (S.D.) of Municipality Arrest, Characteristics, and Protest Data

	Pure Controls	Treated	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-level Arrests	188.2	184.7	188.1
	(489.6)	(168.8)	(480.2)
Black Low-level 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 agencies have had at least one OF. There are 105 high-profile OFs with a maximum of 6 in one jurisdiction, Los Angeles.

Table 2: Fatalities by Civilian & Incident Characteristics with Mean News Articles

	Share	**Mean News Articles**	**Number**
Gender			
Female	14%	2,234	10
Male	86%	18,044	62
Race/Ethnicity			
Asian	1%	1,610	1
Black	74%	20,575	53
Hispanic	11%	2,740	8
Native American	3%	1,385	2
White	11%	3,034	8
Cause of Death			
Asphyxiation/Restrained	3%	2,160	2
Blunt Force	7%	52,442	5
Gunshot	90%	13,454	65
Armed?			
Allegedly Armed	40%	4,230	21
Unarmed	60%	27,309	32
Behavioral Health Issue?			
No	76%	22,485	44
Yes, Alcohol/Drug Use	2%	6,340	1
Yes, Mental Illness	22%	8,012	13

Source: Fatal Encounters (2005–2016); 72 observations; ASR Sample

Appendices

A Appendix: Data

A.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). Though this reporting framework suggests that the data could be an undercount, the National Incident-Based Reporting System (NIBRS)—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.2 Cleaning the Uniform Crime Report

For the primary outcomes of interest, I use the FBI's "Offenses Known and Clearances by Arrest" and the "Arrests by Age, Sex, Race" dataset, cleaned and formatted by Jacob Kaplan (Kaplan, 2018; Kaplan, 2019). Then, I extensively clean the data myself. I further allay concerns about the reporting quality by narrowing the analysis sample to police departments that report at least six years of data between 2005–2016 to both FBI datasets. Additionally, I only keep departments that report each month consistently across years, excluding 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. Moreover, 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 jumps 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 the analysis. Finally, 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, and I have monthly data as to opposed yearly data.

A.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 murders, 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.4 Modifying Event Time Using Google Trends

A key component in measuring the causal impact of high-profile, officer-involved fatalities on policing and offending behavior is ensuring that the timing aligns with community awareness of the incident, when public scrutiny (as well as other possible mechanisms) can plausibly begin. I input the name of each victim in Google Trends, constraining the results to search trends in the locality where they were killed a year before they died to the end of the sample frame (December 2016).[54] Since some of the victim names are common, often there is "noise" in the Google Trends figures, evidenced by searches for the victim name in the locality prior to their death. I conduct a visual test of the trends, and determine the first month after the date of death in which searches spike beyond the "noise" for that search term. For the fatalities whose selected months are different than the month of death, I work to cross-validate the timing by exploring what prompted the rise in community awareness,

[54] For the more high-profile incidents, a completed search phrase automatically comes up when searching the name. For example, when typing in "Laquan McDonald," Google Trends will allow you examine search trends of his name, but also will suggest the "Murder of Laquan McDonald." In those instances, I compare both trends of strictly the name being searched and the search phrase in the locality to determine the proper community awareness time.

typically by finding a news article that describes revelations about the case (e.g., video of the incident that is uncovered or obtained).

A natural example is to look at the death of Laquan McDonald in Chicago, which is discussed thoroughly in Section 5.6 and has a difference in timing between his death and community awareness Figure B.18. For another example, Derek Williams was killed in Milwaukee in July 2011, and searches for that name can be accessed here. There appears to be noise in the search trends, where that name is searched in Milwaukee intermittently throughout the sample frame, including before his death. In fact, the local maximum before September 2012 is November 2010, eight months before his death. The first spike in searches occurs on September 2012, over a year after his death, when the Milwaukee Journal Sentinel (MJS) obtained and posted the squad car surveillance video of his final moments on their website (link to MJS article here, and link to New York Times article detailing timeline of events). Thus, the time of the event is modified to be September 2012, the first date of widespread community awareness.

B Appendix: Figures and Tables

B.1 Supplemental Findings

Table B.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

Figure B.1: Histogram of News Articles on Officer-Involved Fatalities (2005-2016)

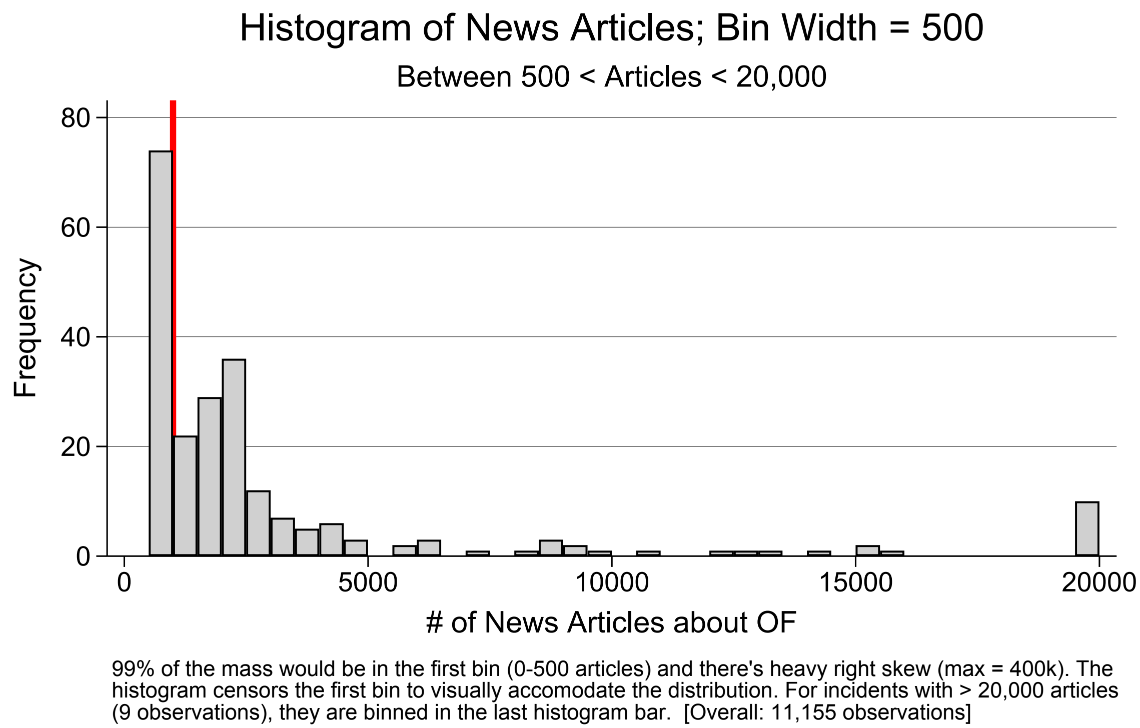
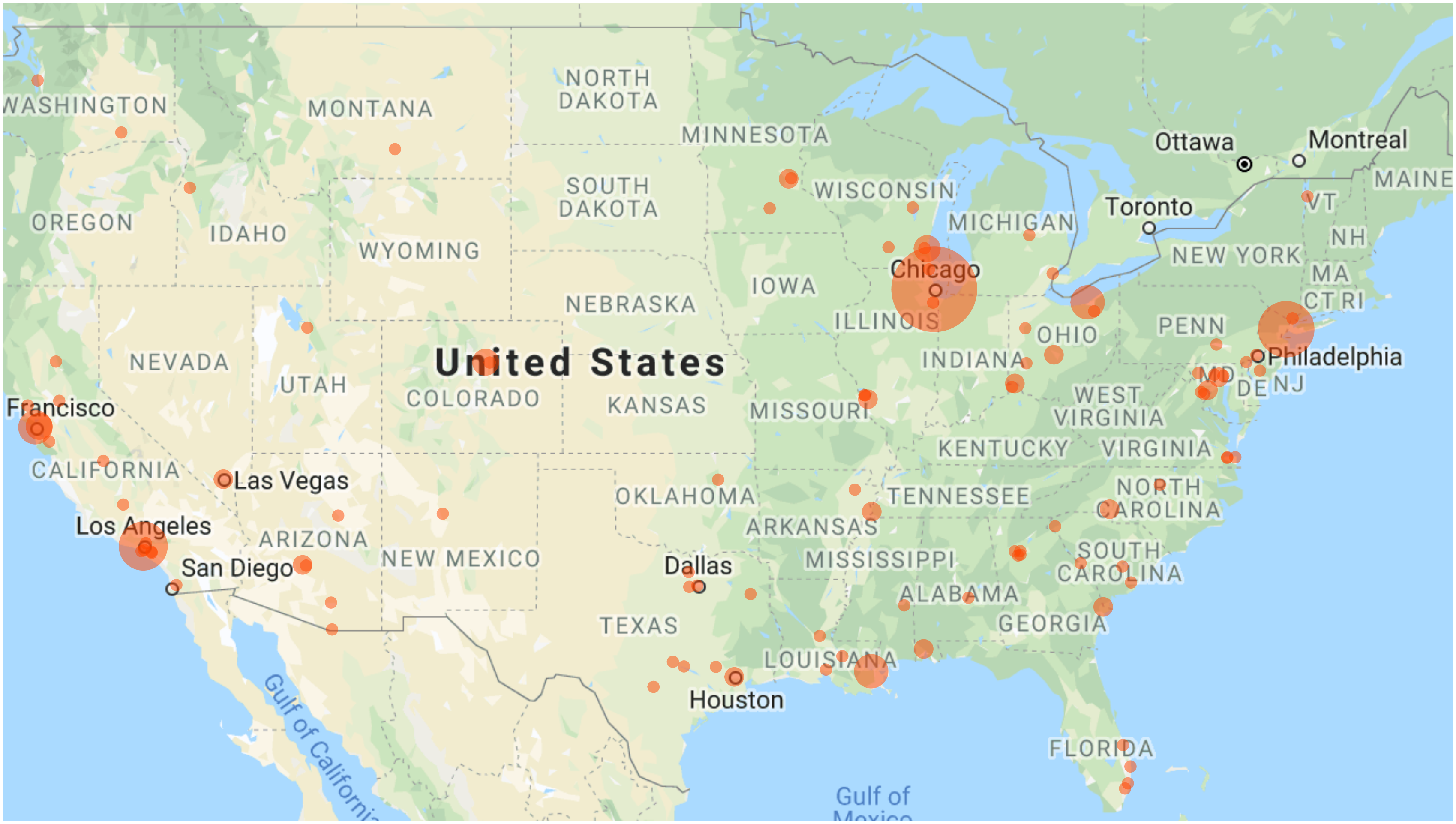
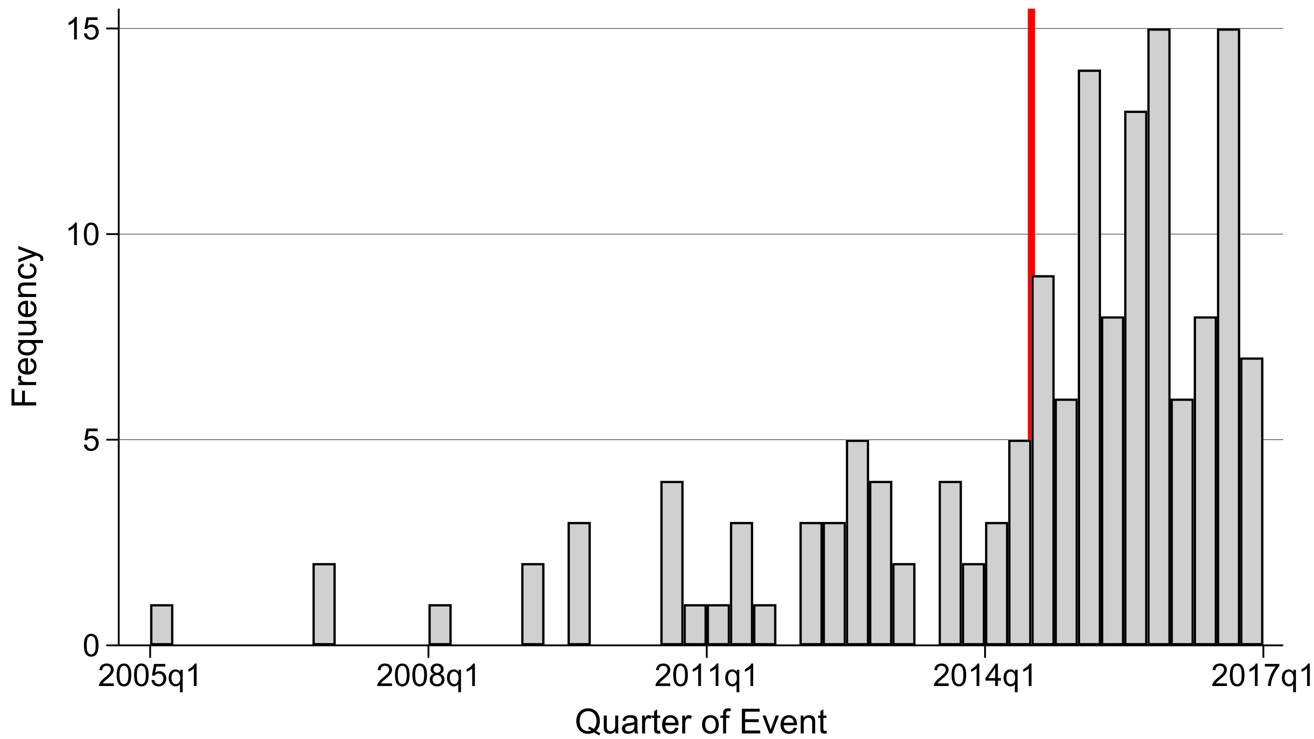


Figure B.2: Map of High-Profile Police Killings (2005–2016)



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 are not used in the analysis sample, because some involved departments poorly reported data to the FBI.

Figure B.3: Timing of High-Profile Police Killings (2005–2016)



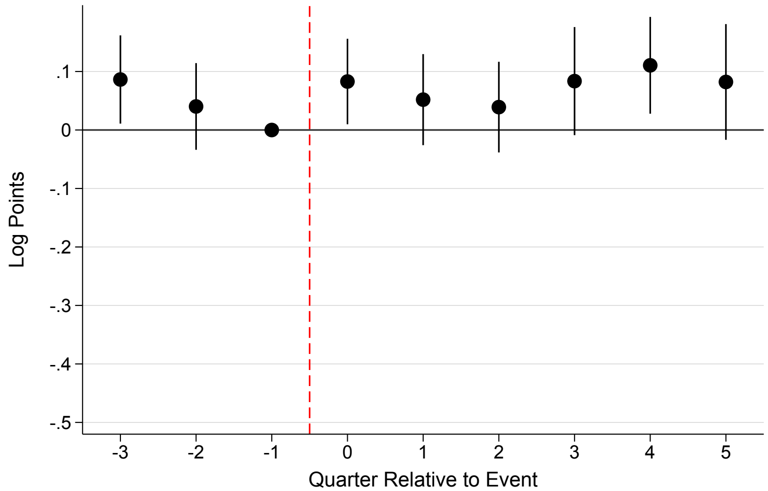
151 fatalities in total from 2005-2016. High profile is defined as ≥ 1,000 news articles. Red line depicts the death of Michael Brown in Ferguson, MO (August 2014). The map includes fatalities that will not be used in the main analysis sample, likely because the involved department poorly reported data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

Figure B.4: Example Integrated Indicator Variable Setup; Event window: -3 to 3 months around event with binned endpoints

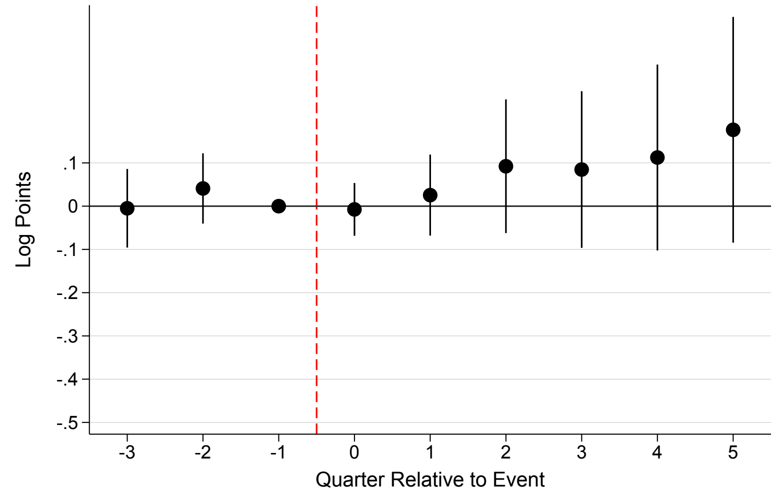
Month Relative to Event	**-4+**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4+**	**Total**
Ferguson: 2005M1	2	0	0	0	.	.	.	.	.	2
Ferguson: 2005M2	2	0	0	0	0	.	.	.	.	2
Ferguson: 2005M3	2	0	0	0	0	0	.	.	.	2
Ferguson: 2005M4	2	0	0	0	0	0	0	.	.	2
Ferguson: 2005M5	2	0	0	0	0	0	0	0	.	2
Ferguson: 2005M6	2	0	0	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 B.5: Effect of Officer-Involved Fatality on Other Arrest Types

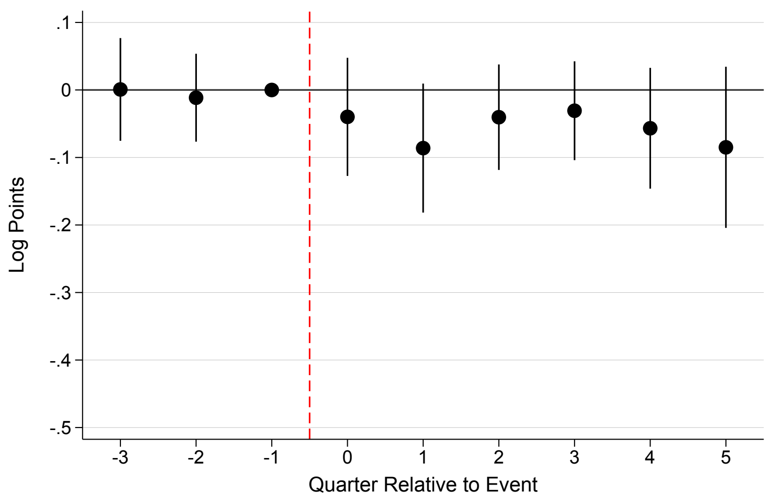
(a) Log(Robbery Arrests)



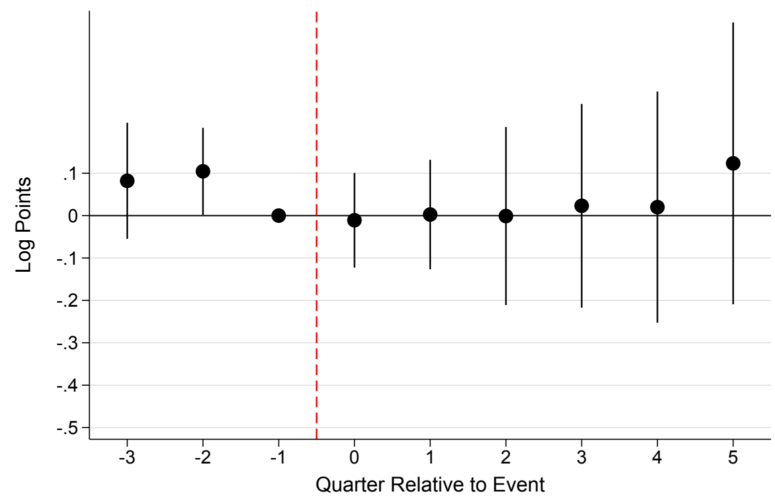
(b) Log(Motor Vehicle Theft Arrests)



(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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

Table B.2: Effect of the Highest-Profile Police Killing on Log(Violent Crime Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Violent Crime Arrests				
Treat*Post	0.029	0.037	0.021	0.028
	(0.026)	(0.030)	(0.031)	(0.038)
Pre-trend Test		0.016		0.015
		(0.029)		(0.041)
Black*Treat*Post			0.017	0.018
			(0.033)	(0.037)
Black Pre-trend Test				0.003
				(0.039)
Panel B: Murder Arrests				
Treat*Post	0.017	0.010	0.018	0.013
	(0.019)	(0.024)	(0.024)	(0.031)
Pre-trend Test		-0.015		-0.009
		(0.029)		(0.036)
Black*Treat*Post			-0.002	-0.008
			(0.038)	(0.049)
Black Pre-trend Test				-0.011
				(0.054)
Panel C: Aggravated Assault Arrests				
Treat*Post	0.001	0.019	-0.007	0.018
	(0.028)	(0.033)	(0.035)	(0.042)
Pre-trend Test		0.035		0.051
		(0.029)		(0.039)
Black*Treat*Post			0.017	0.001
			(0.034)	(0.038)
Black Pre-trend Test				-0.031
				(0.039)
Observations	740,838	740,838	740,838	740,838
Number of Agencies	2,739	2,739	2,739	2,739

Coefficients are from double (DD) and triple difference (DDD) regressions, plotted in Figure 1. 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. 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 B.3: Effect of the Highest-Profile Police Killing on Log(Property Crime Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Property Crime Arrests				
Treat*Post	-0.051**	-0.031	-0.032	-0.001
	(0.025)	(0.028)	(0.030)	(0.036)
Pre-trend Test		0.039**		0.060**
		(0.020)		(0.030)
Black*Treat*Post			-0.039	-0.060*
			(0.025)	(0.033)
Black Pre-trend Test				-0.042
				(0.035)
Panel B: Burglary Arrests				
Treat*Post	-0.051	-0.066*	-0.076**	-0.078*
	(0.031)	(0.036)	(0.035)	(0.043)
Pre-trend Test		-0.029		-0.004
		(0.030)		(0.046)
Black*Treat*Post			0.050	0.024
			(0.038)	(0.052)
Black Pre-trend Test				-0.050
				(0.075)
Panel C: Theft Arrests				
Treat*Post	-0.074**	-0.056	-0.044	-0.021
	(0.034)	(0.038)	(0.038)	(0.045)
Pre-trend Test		0.035		0.045
		(0.022)		(0.029)
Black*Treat*Post			-0.060*	-0.071**
			(0.032)	(0.036)
Black Pre-trend Test				-0.021
				(0.035)
Observations	740,838	740,838	740,838	740,838
Number of Agencies	2,739	2,739	2,739	2,739

Coefficients are from double (DD) and triple difference (DDD) regressions, plotted in Figure 1. 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. 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 B.4: Effect of the Highest-Profile Police Killing on Log(Low-level Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Low-level Arrests				
Treat*Post	-0.095**	-0.116**	-0.082	-0.103*
	(0.048)	(0.051)	(0.054)	(0.058)
Pre-trend Test		-0.042		-0.040
		(0.027)		(0.032)
Black*Treat*Post			-0.025	-0.027
			(0.032)	(0.035)
Black Pre-trend Test				-0.003
				(0.028)
Panel B: Disorderly Conduct Arrests				
Treat*Post	-0.140**	-0.159**	-0.148**	-0.163**
	(0.066)	(0.070)	(0.067)	(0.074)
Pre-trend Test		-0.038		-0.029
		(0.037)		(0.045)
Black*Treat*Post			0.017	0.008
			(0.040)	(0.045)
Black Pre-trend Test				-0.018
				(0.055)
Panel C: Marijuana Possession Arrests				
Treat*Post	-0.141**	-0.165**	-0.148**	-0.180**
	(0.055)	(0.066)	(0.058)	(0.071)
Pre-trend Test		-0.047		-0.063
		(0.043)		(0.052)
Black*Treat*Post			0.014	0.031
			(0.053)	(0.057)
Black Pre-trend Test				0.033
				(0.048)
Observations	741,014	741,014	741,014	741,014
Number of Agencies	2,739	2,739	2,739	2,739

Coefficients are from double (DD) and triple difference (DDD) regressions, plotted in Figure 1. 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. 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 B.5: Effect of the Highest-Profile Police Killing on Log(Violent Crime)

	DD	Pre-trend
	(1)	(2)
Panel A: Violent Crime		
Treat*Post	0.087***	0.096***
	(0.027)	(0.030)
Pre-trend Test		0.018
		(0.025)
Panel B: Murder		
Treat*Post	0.144***	0.142***
	(0.036)	(0.040)
Pre-trend Test		-0.004
		(0.047)
Panel C: Aggravated Assault		
Treat*Post	0.041*	0.059**
	(0.024)	(0.030)
Pre-trend Test		0.034
		(0.032)
Panel D: Robbery		
Treat*Post	0.125***	0.119***
	(0.030)	(0.034)
Pre-trend Test		-0.012
		(0.025)
Observations	740,838	740,838
Number of Agencies	2,739	2,739

Coefficients are from double difference (DD) regressions, plotted in Figure 6. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.6: Effect of the Highest-Profile Police Killing on Log(Property Crime)

	DD	Pre-trend
	(1)	(2)
Panel A: Property Crime		
Treat*Post	0.053**	0.061**
	(0.022)	(0.030)
Pre-trend Test		0.016
		(0.021)
Panel B: Motor Vehicle Theft		
Treat*Post	0.079***	0.099***
	(0.029)	(0.038)
Pre-trend Test		0.040
		(0.032)
Panel C: Burglary		
Treat*Post	0.039	0.046
	(0.027)	(0.034)
Pre-trend Test		0.014
		(0.029)
Panel D: Theft		
Treat*Post	0.051**	0.056
	(0.026)	(0.035)
Pre-trend Test		0.010
		(0.023)
Observations	740,838	740,838
Number of Agencies	2,739	2,739

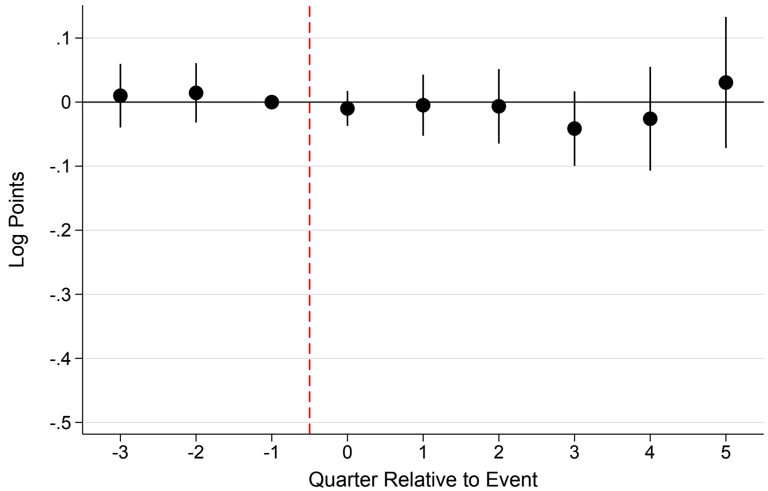
Coefficients are from double difference (DD) regressions, plotted in Figure 6. Standard errors, clustered at the agency level, in parentheses. Uses only city police departments with less than 9 outliers and a population greater 10,000. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Spillover Analysis of High-Profile Police Killings

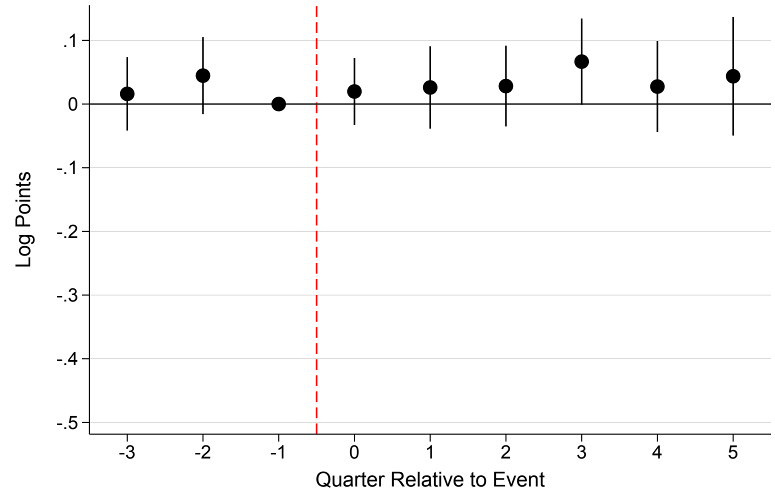
B.2.1 Controlling for Spillovers in Similarly Sized Cities (Population-Group-by-Month-of-Sample Fixed Effects)

Figure B.6: Arrest Results: Estimation with Population-Group-by-Month-of-Sample Fixed Effects

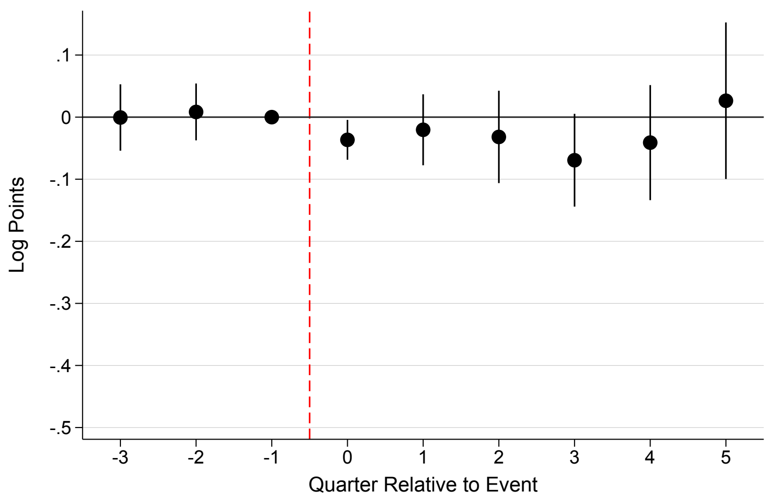
(a) Log(All Part I Arrests)



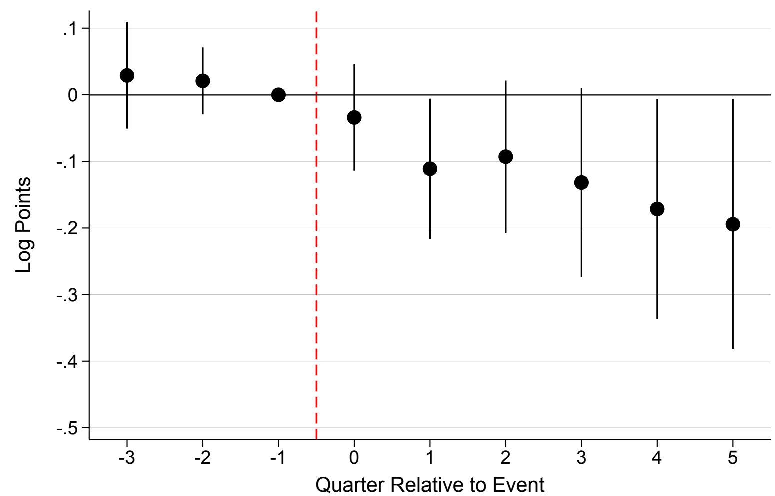
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)

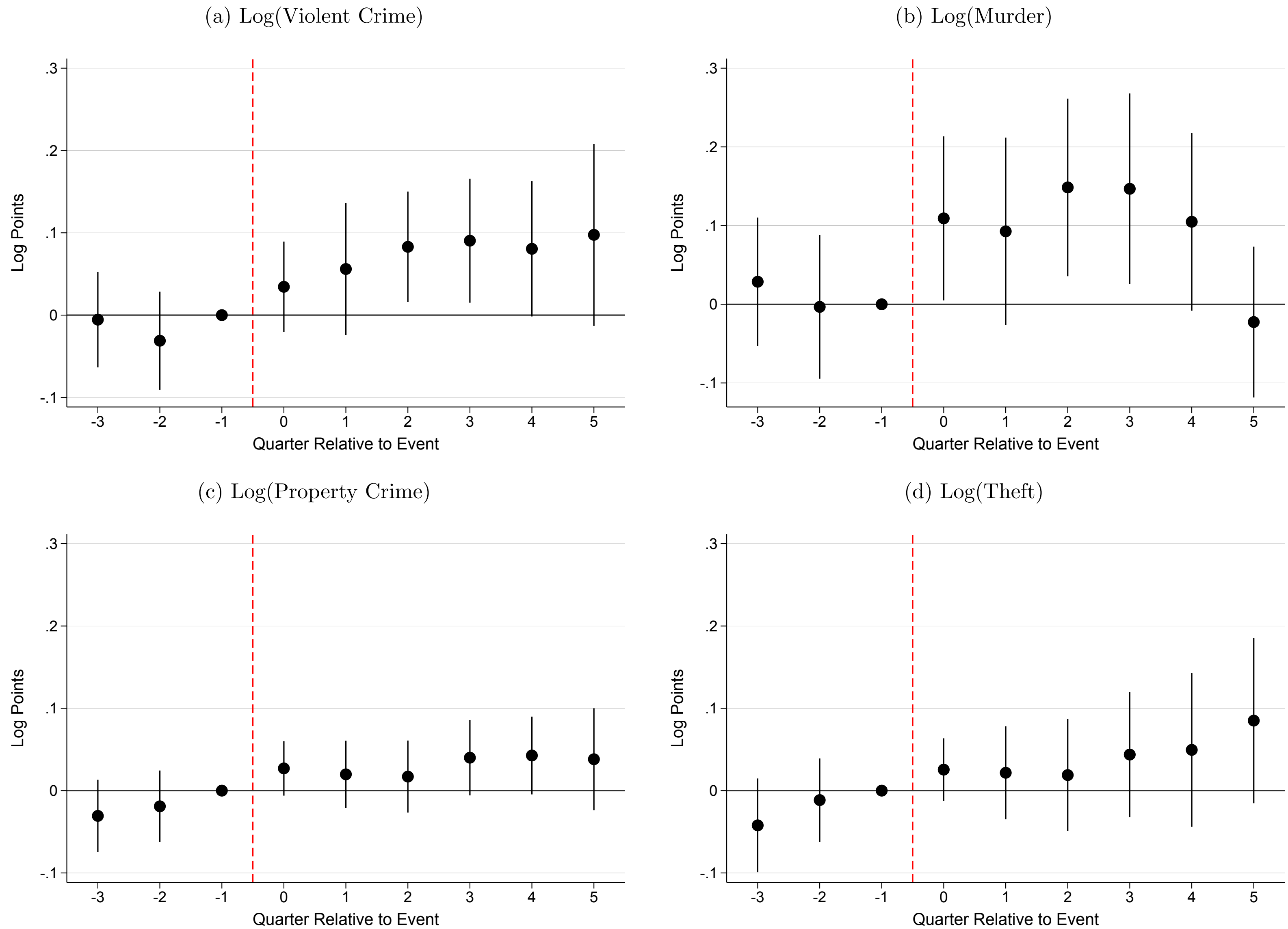


(d) Log(Part II Low-level Arrests)



Using a sample of treated 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 log population. All regressions have month-of-sample interacted with population group fixed effects, department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

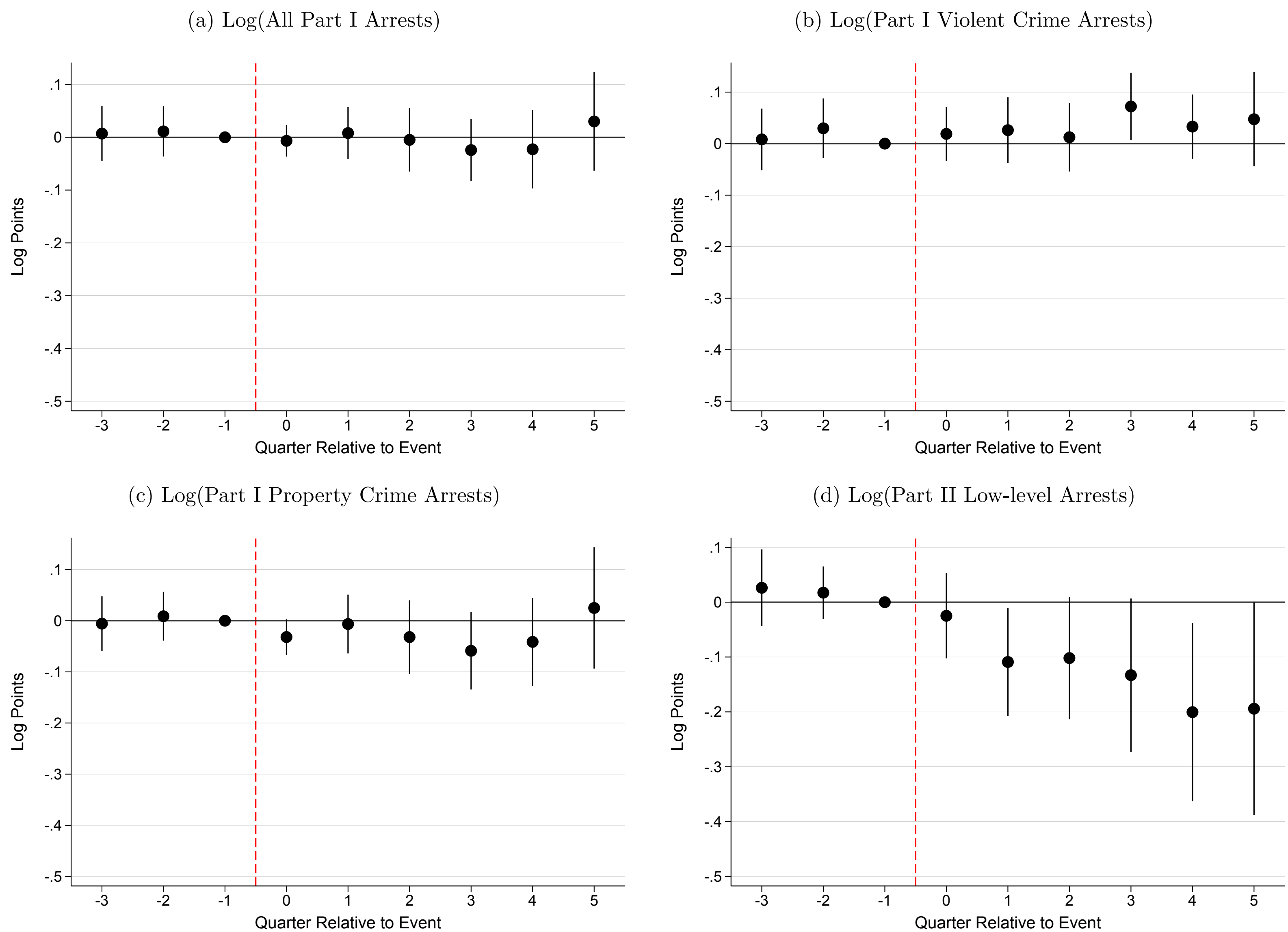
Figure B.7: Crime Results: Estimation with Population-Group-by-Month-of-Sample Fixed Effects



Using a sample of treated 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 log population. All regressions have month-of-sample interacted with population group fixed effects, department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

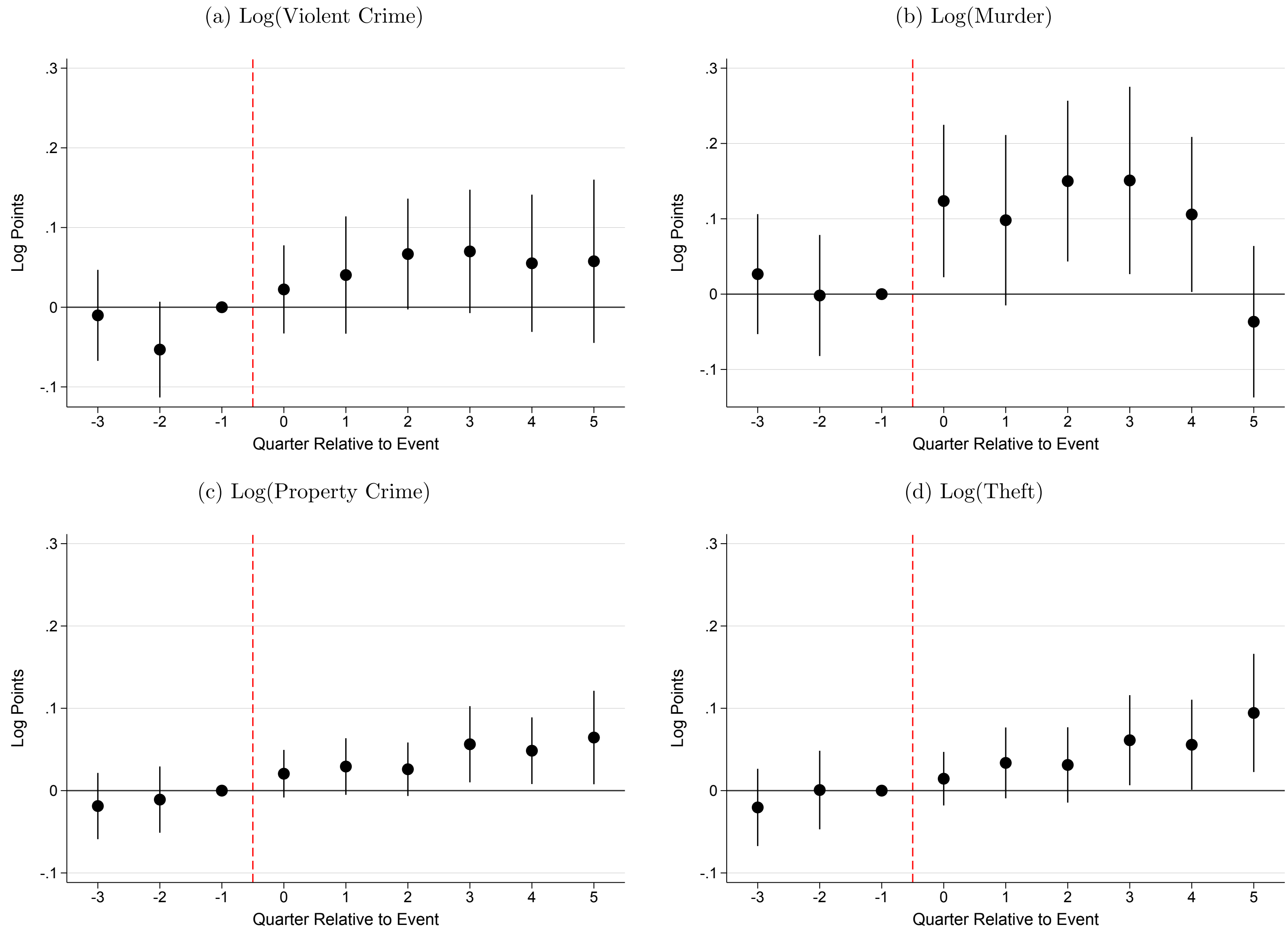
B.2.2 Controlling for Spillovers in Same State (State-by-Month-of-Sample Fixed Effects)

Figure B.8: Arrest Results: Estimation with State-by-Month-of-Sample Fixed Effects



Using a sample of treated 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 log population. All regressions have month-of-sample interacted with state fixed effects, department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

Figure B.9: Crime Results: Estimation with State-by-Month-of-Sample Fixed Effects



Using a sample of treated 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 log population. All regressions have month-of-sample interacted with state fixed effects, department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

B.2.3 Spillover Analysis of Police Departments in the Same County as the Involved Agency

Table B.7: Spillover Analysis of Agencies in the Same County on Log(Violent Crime Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Violent Crime Arrests				
Treat*Post	0.014	0.013	0.040	0.051*
	(0.010)	(0.013)	(0.028)	(0.031)
Pre-trend Test		-0.002		0.023
		(0.011)		(0.028)
Spillover*Treat*Post			-0.035	-0.055*
			(0.028)	(0.032)
Spillover Pre-trend Test				-0.040
				(0.029)
Panel B: Murder Arrests				
Treat*Post	0.002	0.002	0.025	0.016
	(0.004)	(0.004)	(0.022)	(0.024)
Pre-trend Test		-0.002		-0.017
		(0.004)		(0.028)
Spillover*Treat*Post			-0.026	-0.017
			(0.022)	(0.024)
Spillover Pre-trend Test				0.018
				(0.028)
Panel C: Aggravated Assault Arrests				
Treat*Post	0.012	0.017	0.018	0.030
	(0.008)	(0.012)	(0.027)	(0.033)
Pre-trend Test		0.011		0.023
		(0.014)		(0.031)
Spillover*Treat*Post			-0.010	-0.022
			(0.027)	(0.034)
Spillover Pre-trend Test				-0.023
				(0.032)
Observations	741,532	741,532	741,532	741,532
Number of Counties	967	967	967	967

Coefficients are from double (DD) and triple difference (DDD) regressions, where Treat is now defined at the county level and Spillover references departments that reside in the same county as the involved department. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the spillover dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.8: Spillover Analysis of Agencies in the Same County on Log(Property Crime Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Property Crime Arrests				
Treat*Post	-0.010	-0.011	-0.047*	-0.032
	(0.014)	(0.015)	(0.027)	(0.030)
Pre-trend Test		-0.002		0.029
		(0.011)		(0.020)
Spillover*Treat*Post			0.028	0.008
			(0.029)	(0.032)
Spillover Pre-trend Test				-0.038
				(0.023)
Panel B: Burglary Arrests				
Treat*Post	-0.021*	-0.024**	-0.031	-0.045
	(0.011)	(0.011)	(0.034)	(0.039)
Pre-trend Test		-0.006		-0.029
		(0.011)		(0.030)
Spillover*Treat*Post			0.020	0.038
			(0.038)	(0.041)
Spillover Pre-trend Test				0.036
				(0.031)
Panel C: Theft Arrests				
Treat*Post	-0.015	-0.019	-0.087**	-0.077*
	(0.013)	(0.016)	(0.036)	(0.041)
Pre-trend Test		-0.009		0.021
		(0.013)		(0.021)
Spillover*Treat*Post			0.066*	0.046
			(0.038)	(0.042)
Spillover Pre-trend Test				-0.038
				(0.024)
Observations	741,532	741,532	741,532	741,532
Number of Counties	967	967	967	967

Coefficients are from double (DD) and triple difference (DDD) regressions, where Treat is now defined at the county level and Spillover references departments that reside in the same county as the involved department. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the spillover dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.9: Spillover Analysis of Agencies in the Same County on Log(Low-level Arrests)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Low-level Arrests				
Treat*Post	-0.045**	-0.064***	-0.087*	-0.114**
	(0.018)	(0.021)	(0.053)	(0.055)
Pre-trend Test		-0.039***		-0.054**
		(0.015)		(0.027)
Spillover*Treat*Post			0.064	0.081
			(0.054)	(0.054)
Spillover Pre-trend Test				0.034
				(0.029)
Panel B: Disorderly Conduct Arrests				
Treat*Post	-0.029*	-0.045***	-0.129*	-0.145**
	(0.016)	(0.017)	(0.068)	(0.069)
Pre-trend Test		-0.033**		-0.032
		(0.014)		(0.038)
Spillover*Treat*Post			0.106*	0.113*
			(0.062)	(0.065)
Spillover Pre-trend Test				0.014
				(0.039)
Panel C: Marijuana Possession Arrests				
Treat*Post	-0.038	-0.069*	-0.166***	-0.199***
	(0.030)	(0.042)	(0.058)	(0.073)
Pre-trend Test		-0.061**		-0.067
		(0.029)		(0.047)
Spillover*Treat*Post			0.150**	0.158**
			(0.063)	(0.073)
Spillover Pre-trend Test				0.017
				(0.044)
Observations	741,708	741,708	741,708	741,708
Number of Counties	967	967	967	967

Coefficients are from double (DD) and triple difference (DDD) regressions, where Treat is now defined at the county level and Spillover references departments that reside in the same county as the involved department. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the spillover dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.10: Spillover Analysis of Agencies in the Same County on Log(Violent Crime)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Murder				
Treat*Post	0.002	0.002	0.149***	0.149***
	(0.004)	(0.004)	(0.039)	(0.042)
Pre-trend Test		-0.002		0.001
		(0.004)		(0.049)
Spillover*Treat*Post			-0.151***	-0.160***
			(0.039)	(0.041)
Spillover Pre-trend Test				-0.018
				(0.049)
Panel B: Aggravated Assault				
Treat*Post	0.012	0.017	0.069**	0.088***
	(0.008)	(0.012)	(0.027)	(0.032)
Pre-trend Test		0.011		0.037
		(0.014)		(0.030)
Spillover*Treat*Post			-0.051*	-0.077**
			(0.031)	(0.036)
Spillover Pre-trend Test				-0.051
				(0.034)
Panel C: Robbery				
Treat*Post	-0.000	-0.006	0.134***	0.124***
	(0.011)	(0.011)	(0.031)	(0.035)
Pre-trend Test		-0.012		-0.019
		(0.011)		(0.029)
Spillover*Treat*Post			-0.136***	-0.151***
			(0.029)	(0.034)
Spillover Pre-trend Test				-0.030
				(0.030)
Observations	741,532	741,532	741,532	741,532
Number of Counties	967	967	967	967

Coefficients are from double (DD) and triple difference (DDD) regressions, where Treat is now defined at the county level and Spillover references departments that reside in the same county as the involved department. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the spillover dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.11: Spillover Analysis of Agencies in the Same County on Log(Property Crime)

	DD	Pre-trend	DDD	Pre-trend
	(1)	(2)	(3)	(4)
Panel A: Motor Vehicle Theft				
Treat*Post	0.022	0.026	0.097***	0.126***
	(0.017)	(0.017)	(0.033)	(0.041)
Pre-trend Test		0.008		0.058*
		(0.007)		(0.033)
Spillover*Treat*Post			-0.106***	-0.136***
			(0.036)	(0.047)
Spillover Pre-trend Test				-0.058
				(0.037)
Panel B: Burglary				
Treat*Post	-0.021*	-0.024**	0.043	0.048
	(0.011)	(0.011)	(0.030)	(0.036)
Pre-trend Test		-0.006		0.009
		(0.011)		(0.032)
Spillover*Treat*Post			-0.046	-0.066
			(0.033)	(0.041)
Spillover Pre-trend Test				-0.039
				(0.038)
Panel C: Theft				
Treat*Post	-0.015	-0.019	0.028	0.028
	(0.013)	(0.016)	(0.024)	(0.034)
Pre-trend Test		-0.009		-0.000
		(0.013)		(0.024)
Spillover*Treat*Post			-0.029	-0.049
			(0.029)	(0.041)
Spillover Pre-trend Test				-0.039
				(0.037)
Observations	741,532	741,532	741,532	741,532
Number of Counties	967	967	967	967

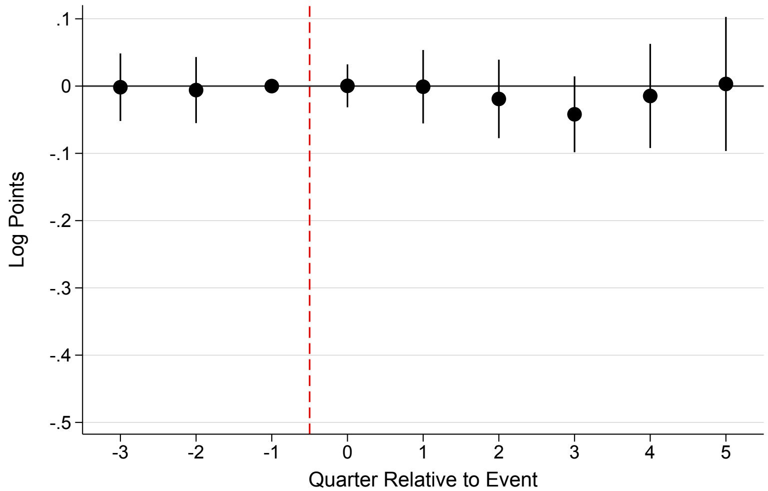
Coefficients are from double (DD) and triple difference (DDD) regressions, where Treat is now defined at the county level and Spillover references departments that reside in the same county as the involved department. Standard errors, clustered at the agency level, in parentheses. Uses only city police departments with less than 9 outliers and a population greater 10,000. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends, which is interacted with the spillover dummy for the DDD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.3 Robustness to Potential Negative Weights in Two-Way Fixed Effects Estimator

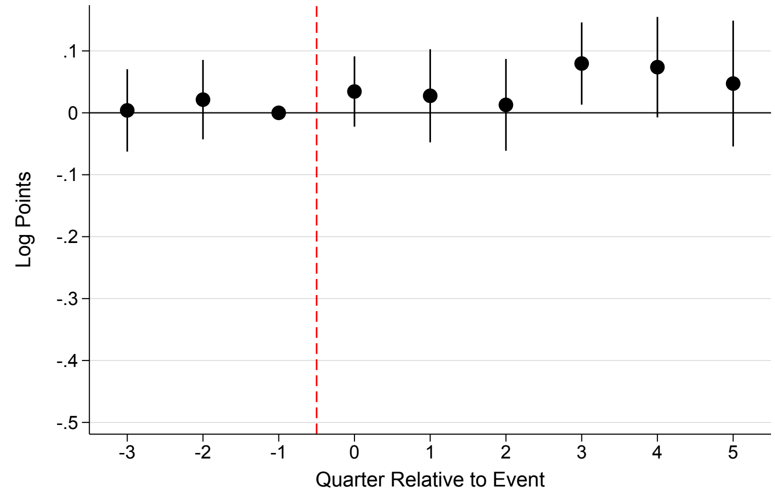
B.3.1 Specification with Police Department Fixed Effects (One-way)

Figure B.10: One-Way Fixed Effects: Effect of High-Profile Police Killing on Arrests

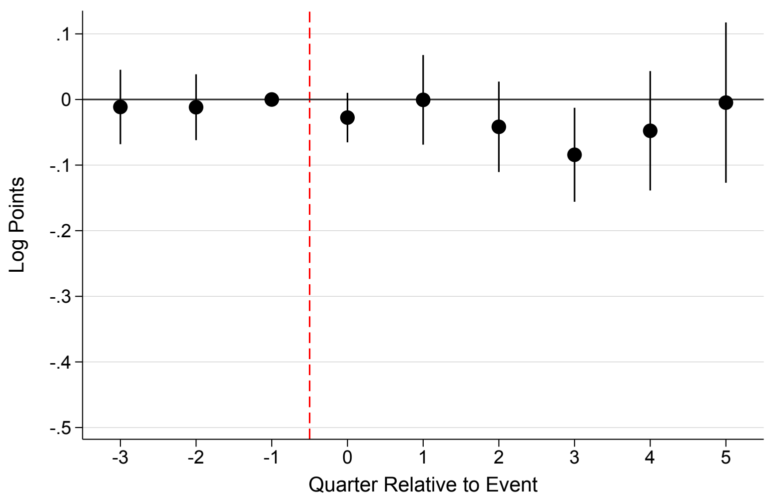
(a) Log(All Part I Arrests)



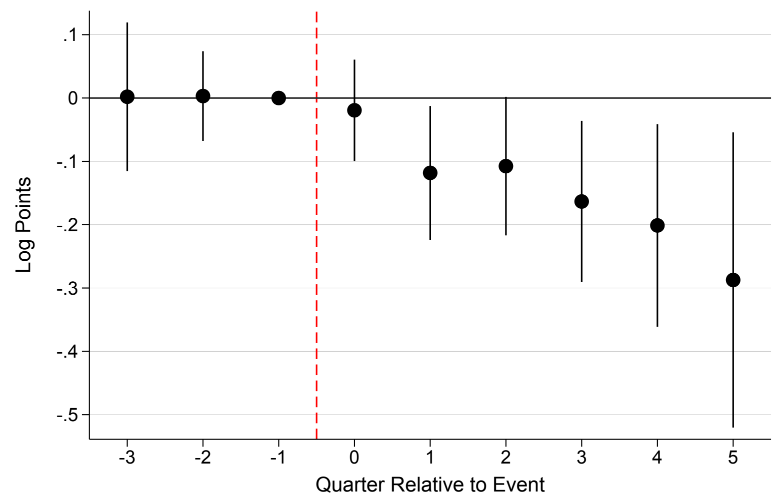
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)



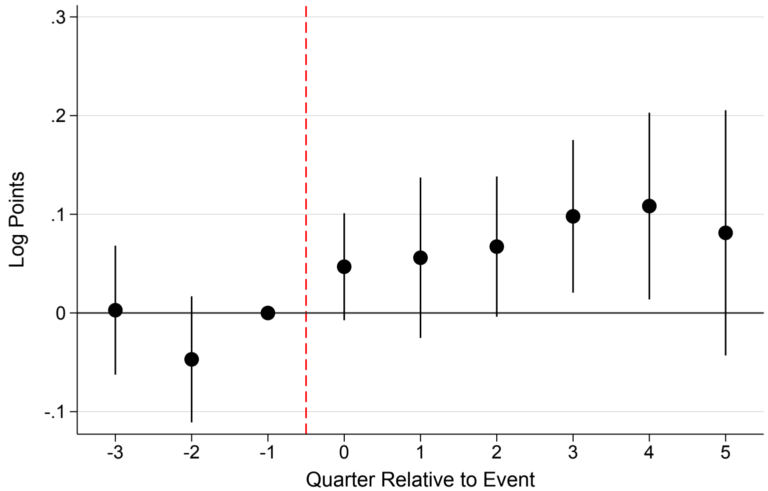
(d) Log(Part II Low-level Arrests)



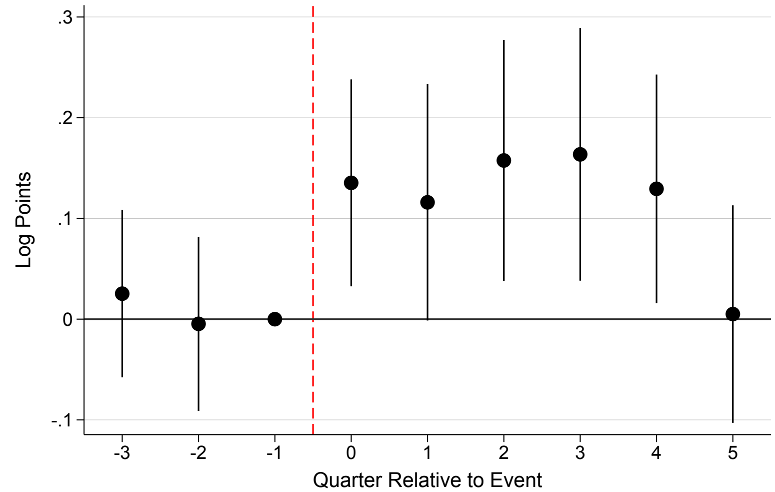
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have department fixed effects. [598,788 observations; 52 treated; 2,687 control agencies]

Figure B.11: One-Way Fixed Effects: Effect of High-Profile Police Killing on Violent and Property Crime

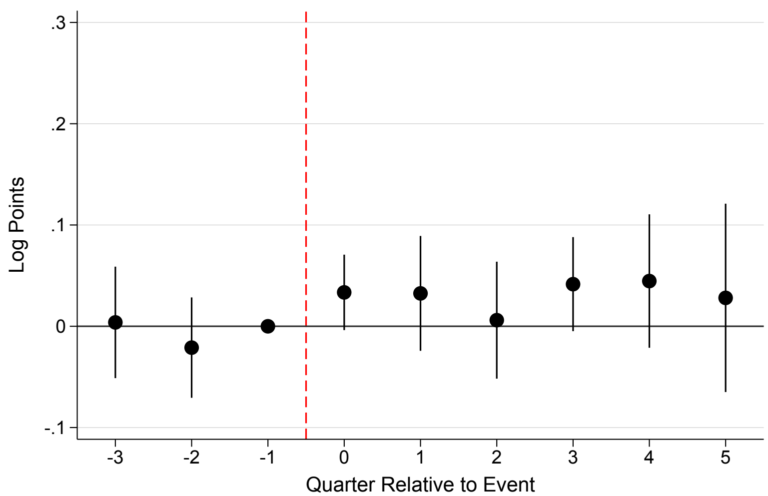
(a) Log(Violent Crime)



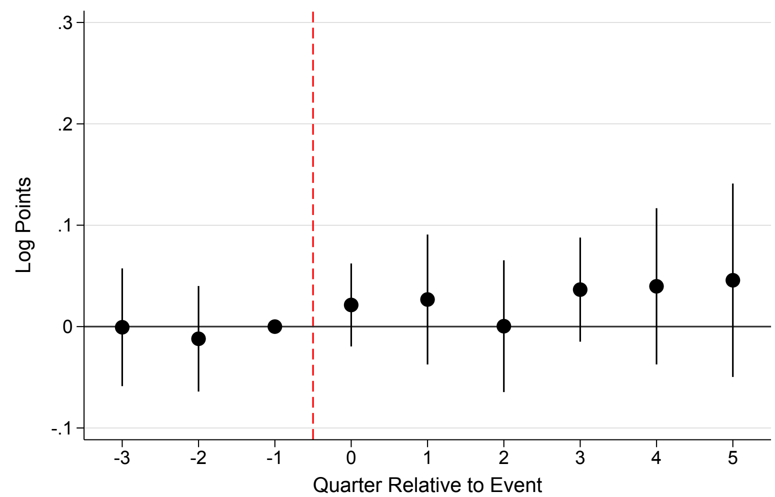
(b) Log(Murder)



(c) Log(Property Crime)



(d) Log(Theft)

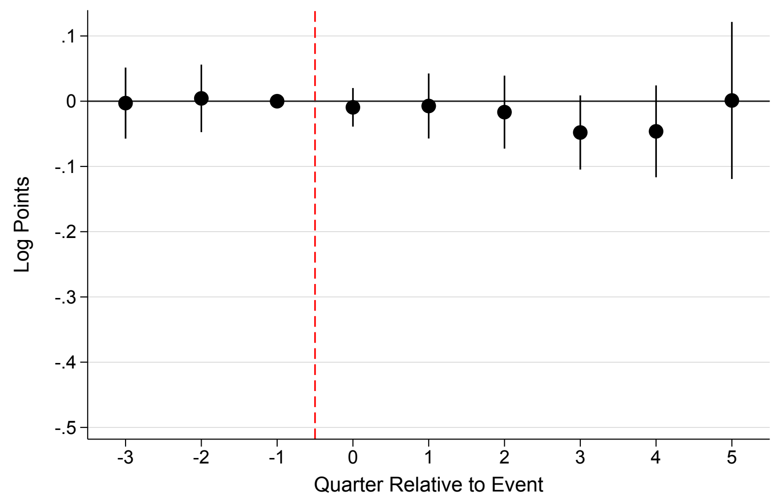


Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in crime after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population. All regressions have department fixed effects. [598,788 observations; 52 treated; 2,687 control agencies]

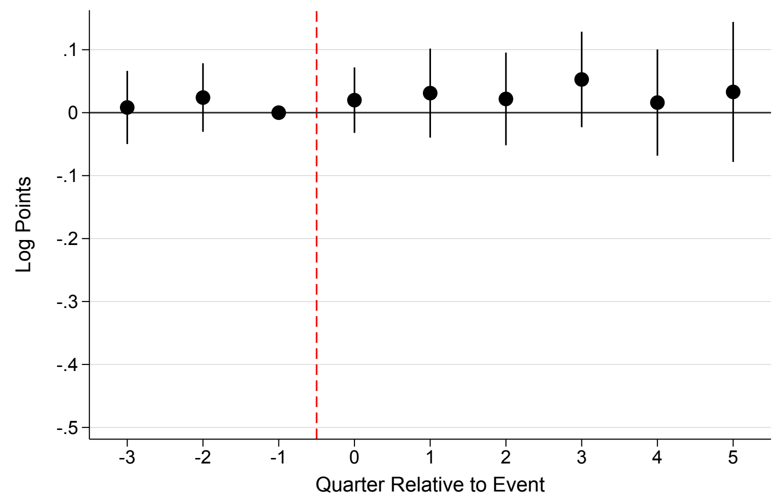
B.3.2 Analysis Sample with Exclusively Treated Police Departments

Figure B.12: Treated Departments Only: Effect of High-Profile Police Killing on Arrests

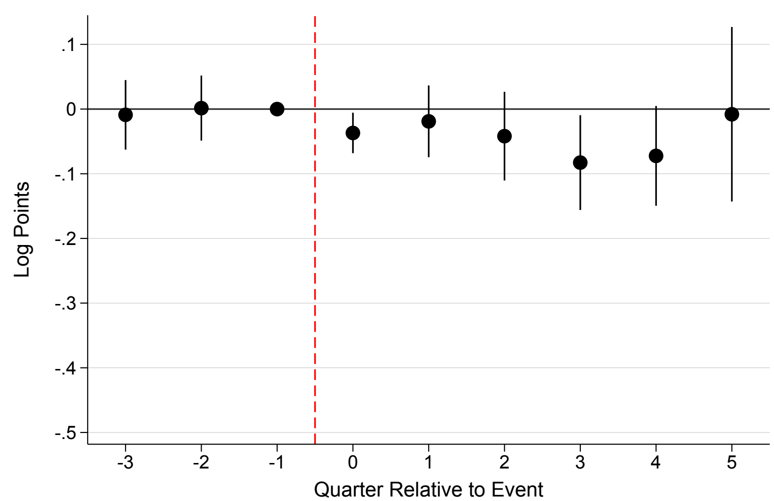
(a) Log(All Part I Arrests)



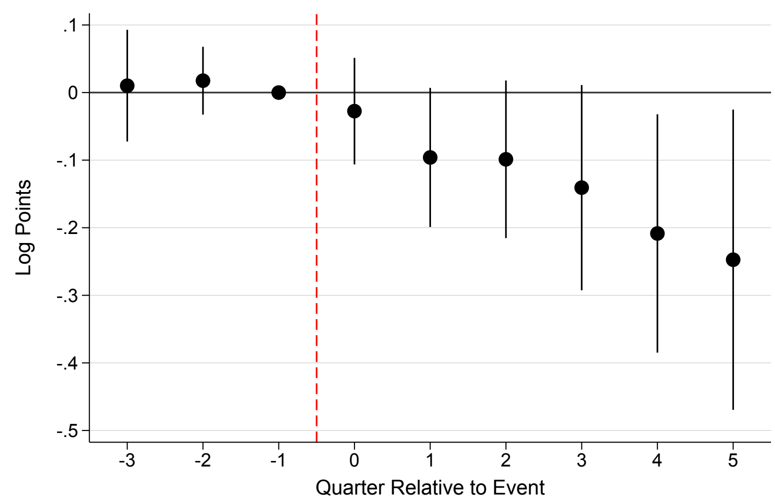
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)

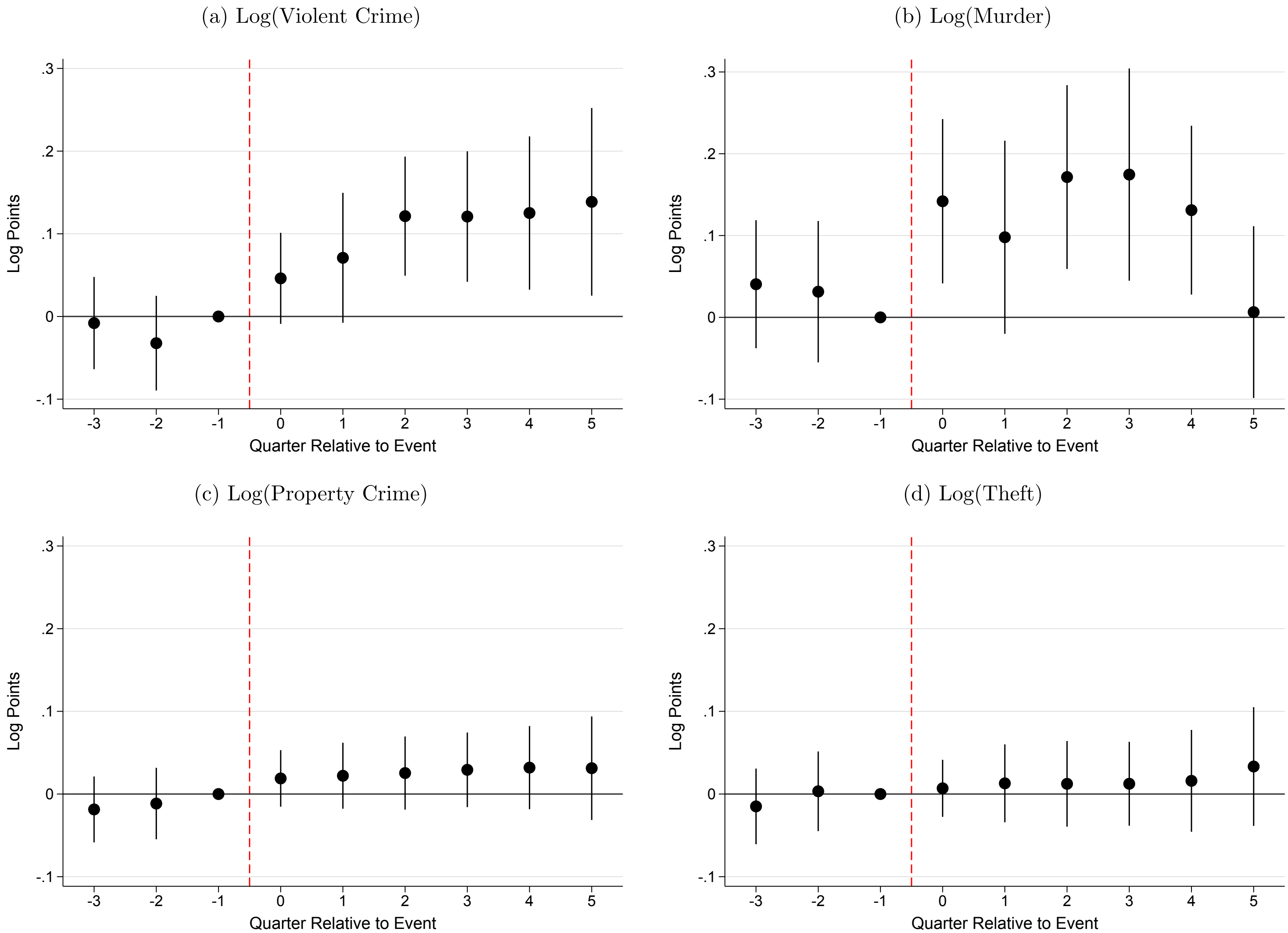


(d) Log(Part II Low-level Arrests)



Using a sample of treated 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. [11,688 observations; 52 treated]

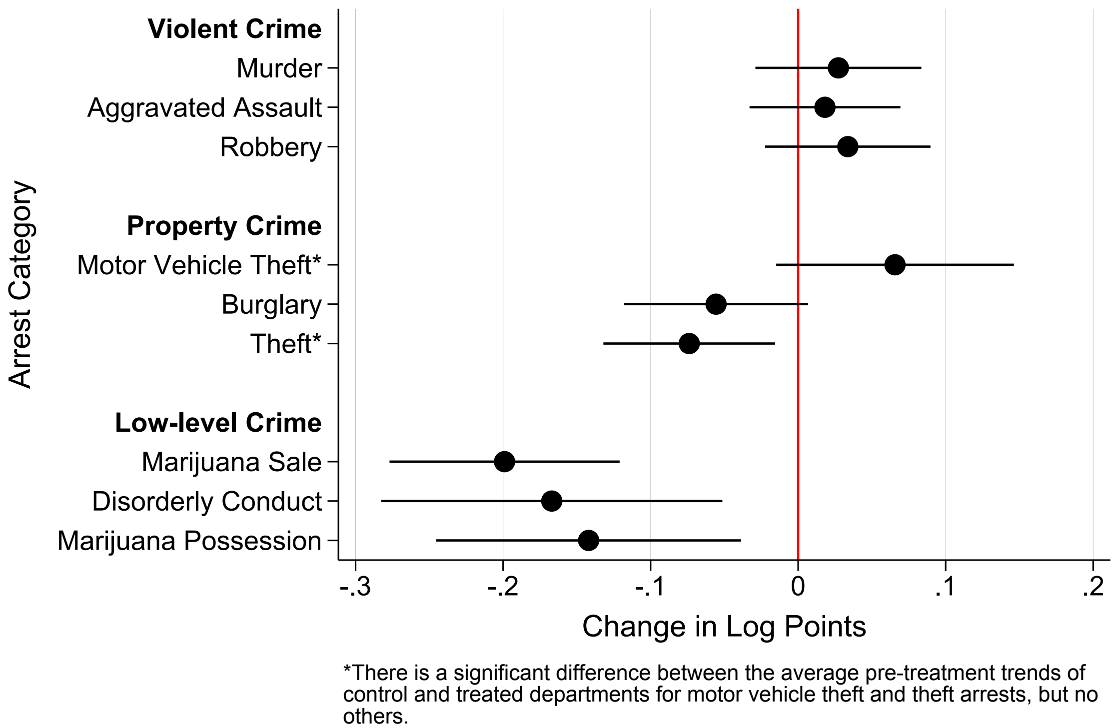
Figure B.13: Treated Departments Only: Effect of High-Profile Police Killing on Violent and Property Crime



Using a sample of treated 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. [11,688 observations; 52 treated]

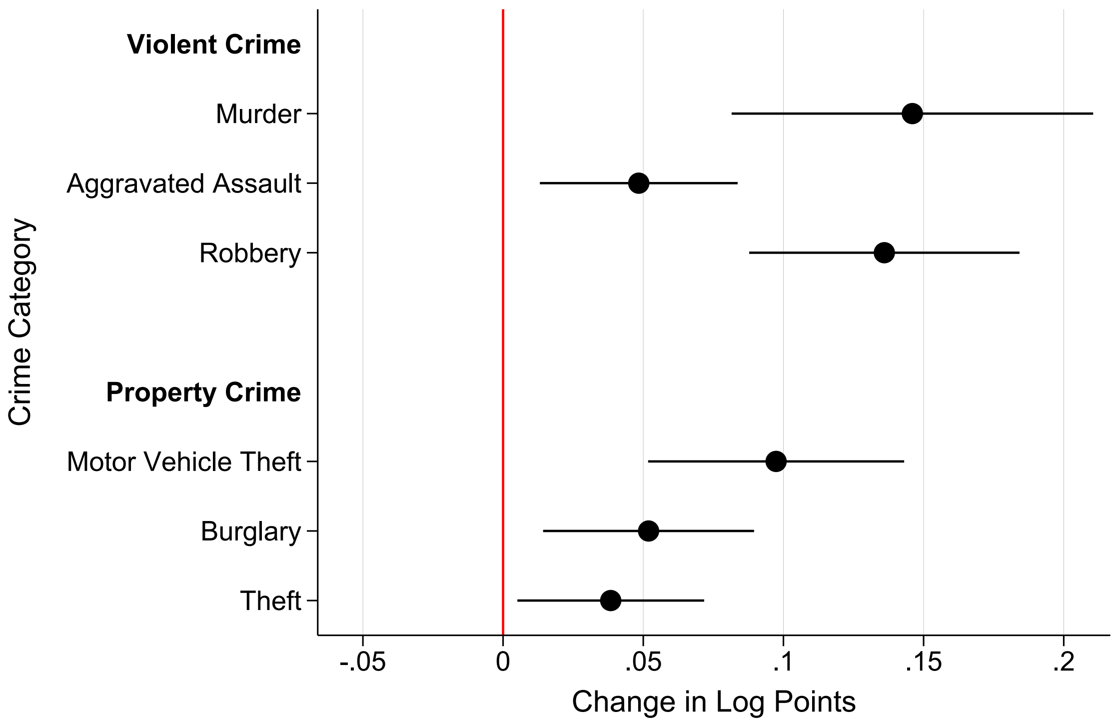
B.3.3 Sun and Abraham (2020): Difference-in-differences (DD) Cohort Interaction-Weighted (IW) Coefficient Estimates

Figure B.14: DD Cohort Interaction-weighted (IW) Coefficient Estimates by Arrest Category



Circles display DD coefficients using the IW estimator from Sun and Abraham (2020). The coefficients are 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. 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. The DD coefficients are similar to the event study coefficients using the IW estimator and only including the highest-profile police killing per jurisdiction.

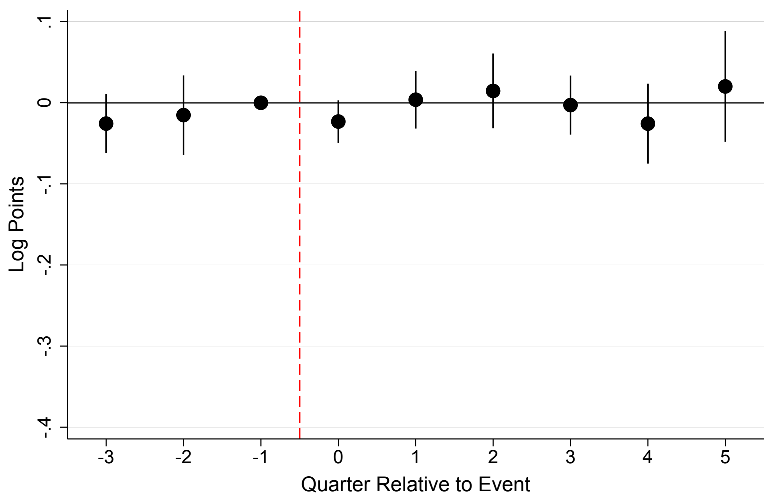
Figure B.15: DD Cohort Interaction-weighted (IW) Coefficient Estimates by Crime Category



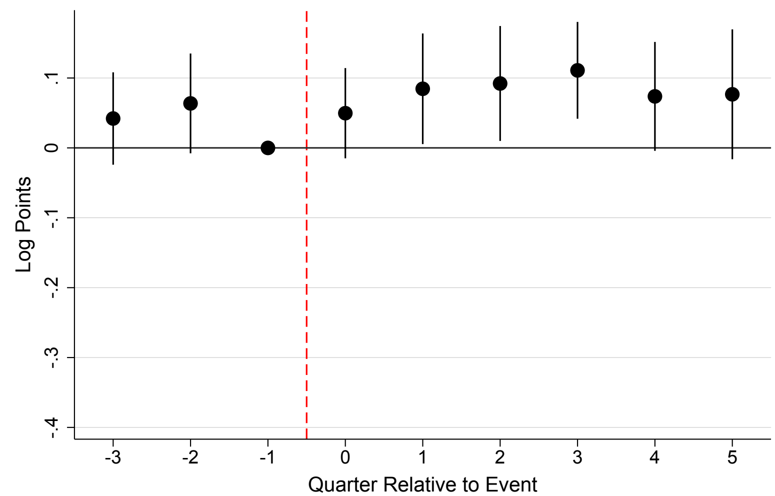
Circles display DD coefficients using the IW estimator from Sun and Abraham (2020). The coefficients are from separate regressions—in descending order of the social cost of crime—using a sample of city police departments with fewer than 9 outliers and a population greater 10,000. Lines represent the 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. The DD coefficients are similar to the event study coefficients using the IW estimator and only including the highest-profile police killing per jurisdiction.

Figure B.16: Cohort Interaction-weighted (IW) Coefficient Estimates for Arrest

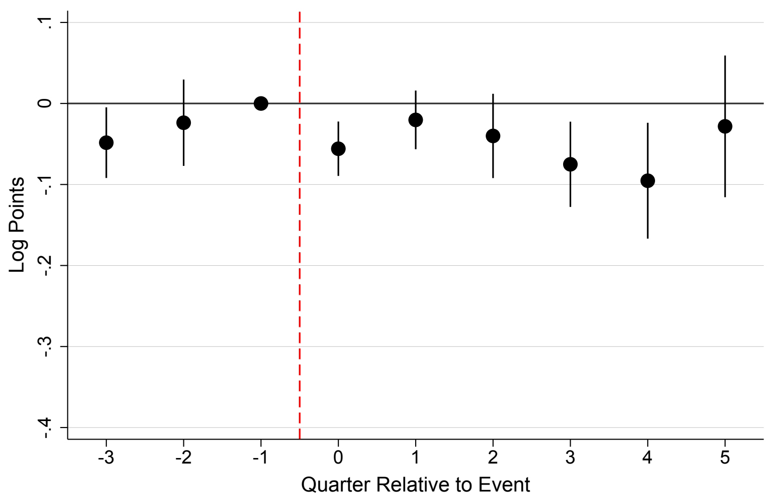
(a) Log(All Part I Arrests)



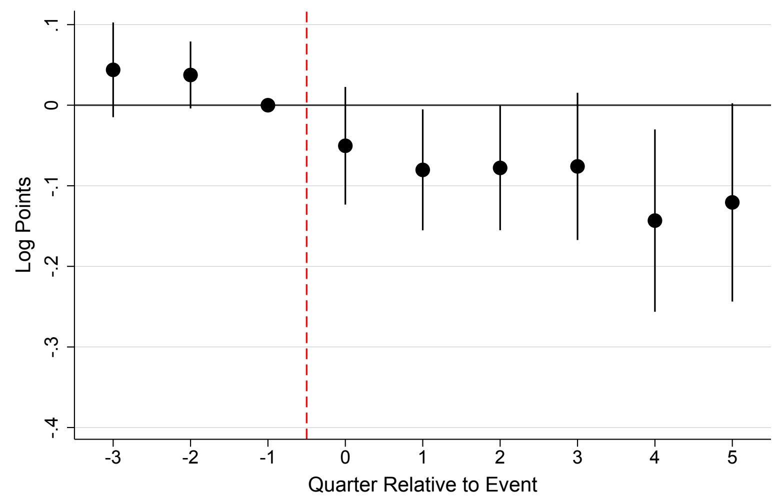
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)



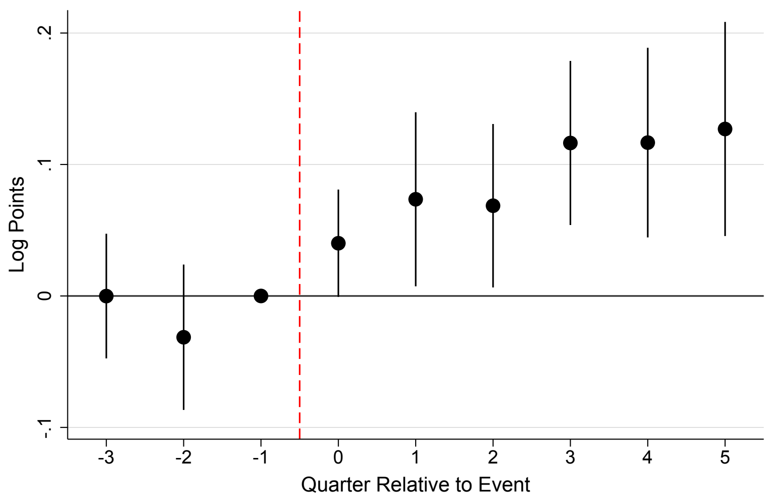
(d) Log(Part II Low-level Arrests)



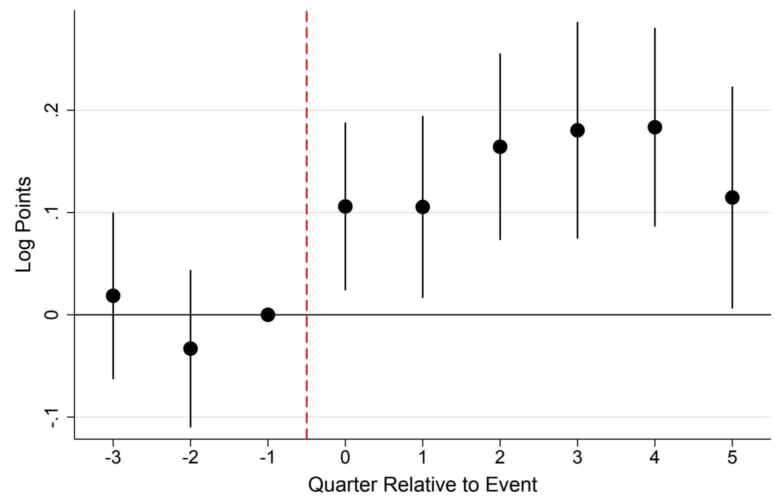
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) use the IW estimator from Sun and Abraham (2020) and represent the change in arrests after the highest-profile police killing per police department. Lines 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.

Figure B.17: Cohort Interaction-weighted (IW) Coefficient Estimates for Violent and Property Crime

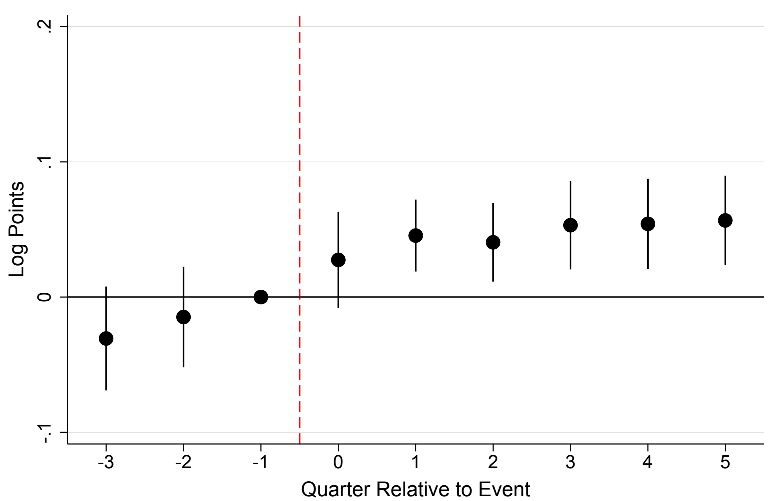
(a) Log(Violent Crime)



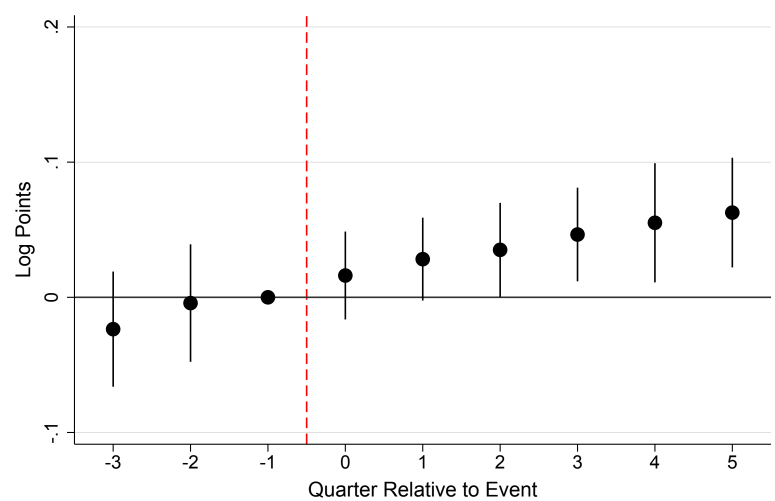
(b) Log(Murder)



(c) Log(Property Crime)



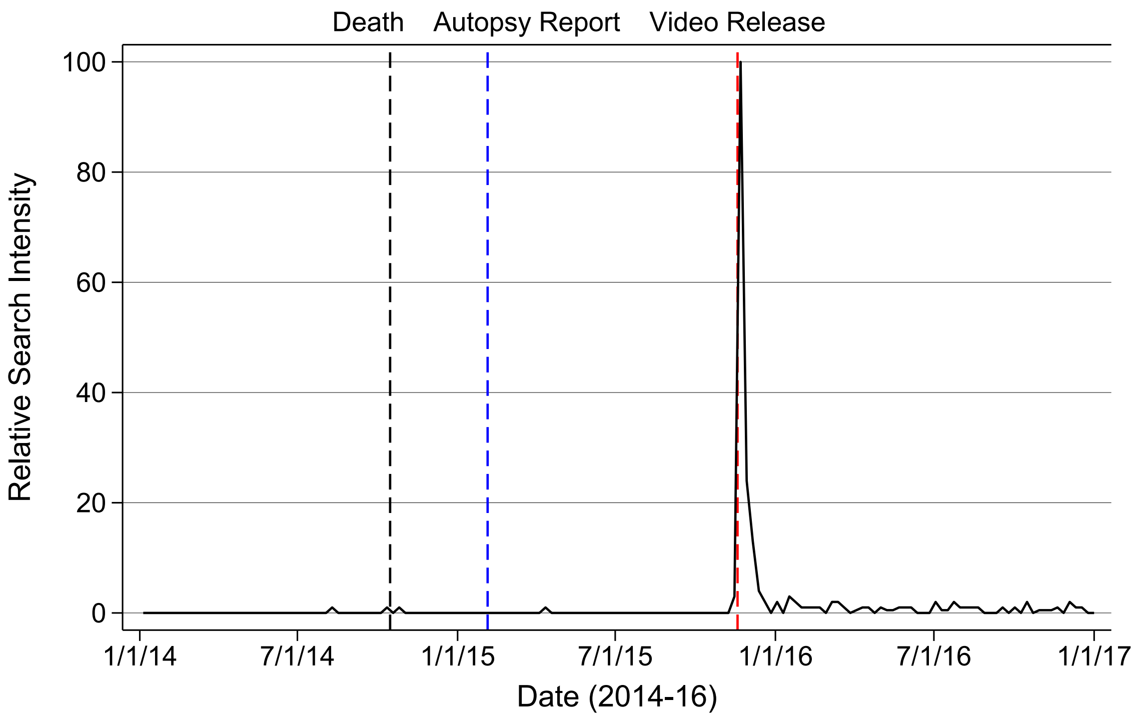
(d) Log(Theft)



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 use the IW estimator from Sun and Abraham (2020) and represent the change in crime after the highest-profile police killing per police department. 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.

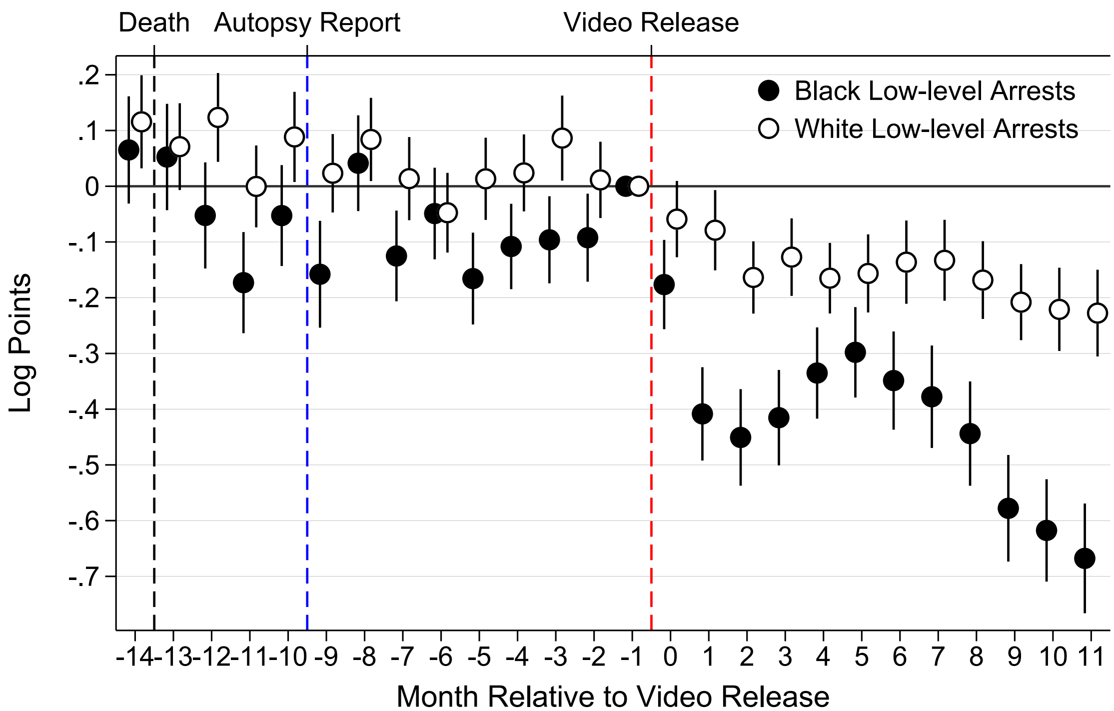
B.4 Case Study of Laquan McDonald and the Chicago Police Department

Figure B.18: Google Searches for Laquan McDonald in Chicago (2014-2016)



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).

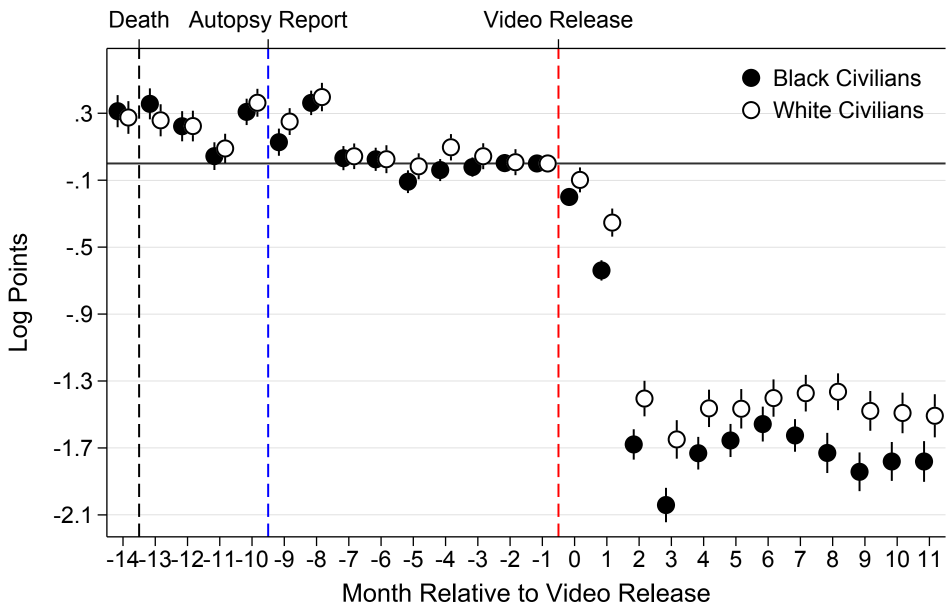
Figure B.19: Laquan McDonald's Death on Low-level Arrests by Race of Suspect



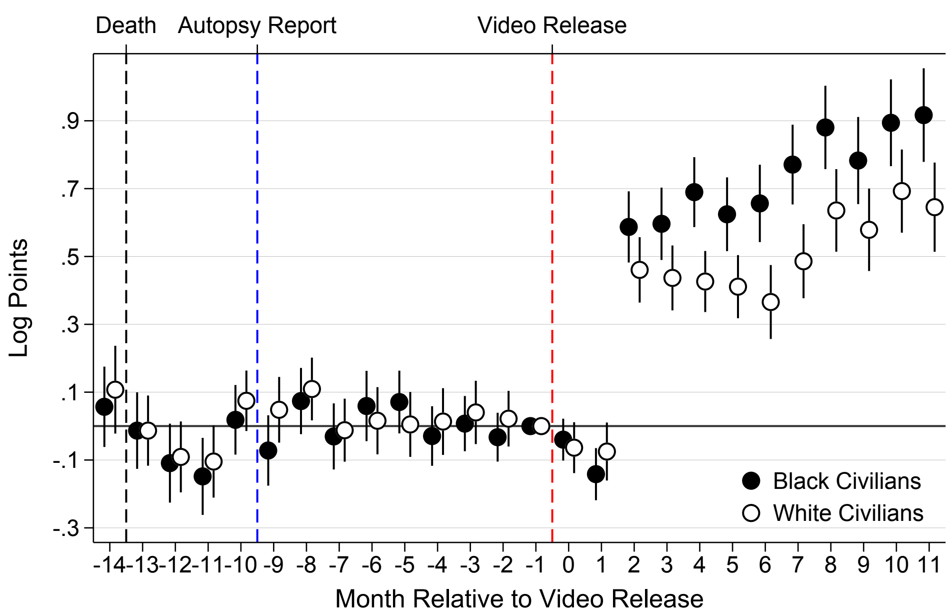
The dotted black line (left) depicts the month of death for Laquan McDonald (October 2014). The dotted blue line (middle) denotes when an article with the associated autopsy report was released (February 2015). The dotted red line (right) highlights when the video of his death was released to the public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

Figure B.20: Laquan McDonald's Death on Stops by Race of Civilian

(a) Log(Pedestrian Stops) by Race of Civilian

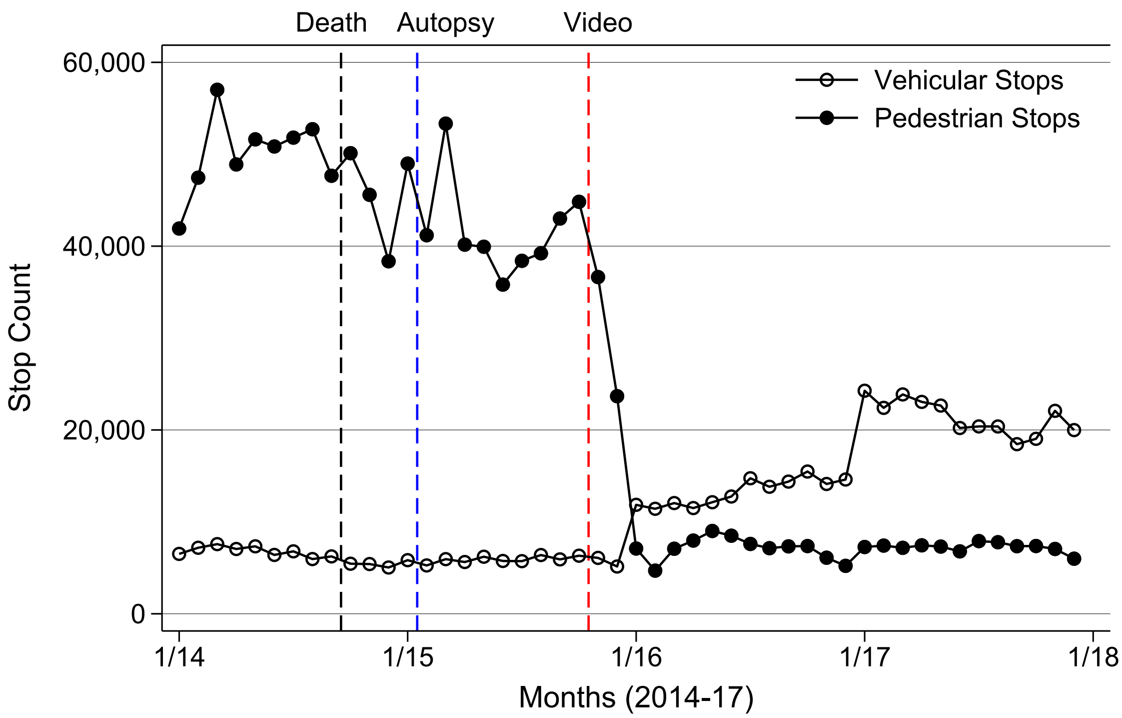


(b) Log(Vehicular Stops) by Race of Civilian



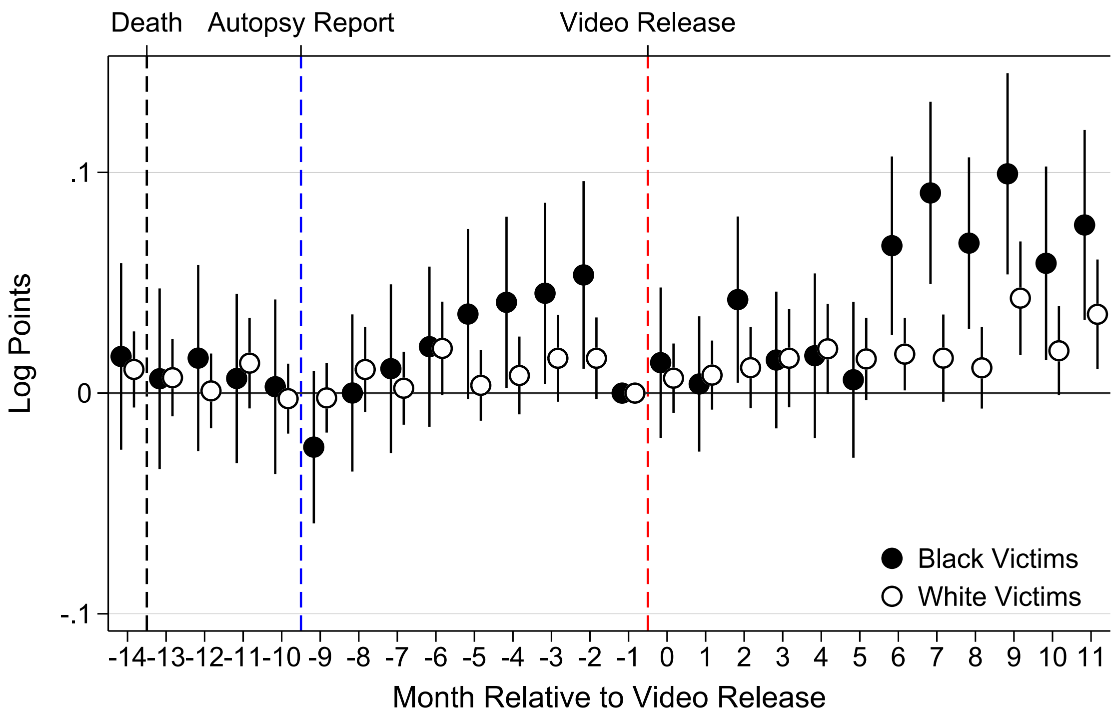
Circles display monthly event time coefficients from a regression of log pedestrian or vehicular stops at the beat-month-race level in Chicago from 2014–2017. Lines represent the 95% confidence interval using standard errors clustered at the beat level. I control for beat-level fixed events and linear trends, which are interacted with race indicators.

Figure B.21: Chicago Police Department: Substitution from Pedestrian to Vehicular Stops



This figure depicts the unadjusted counts of Black and White civilians stopped as pedestrians or in a vehicle in Chicago from 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).

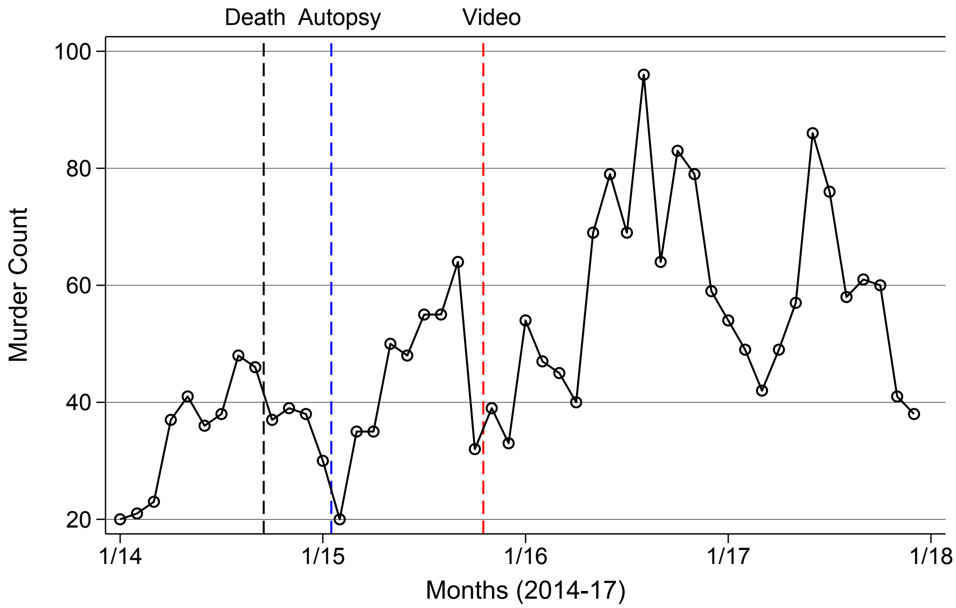
Figure B.22: Laquan McDonald's Death on Murders by Race of Victim



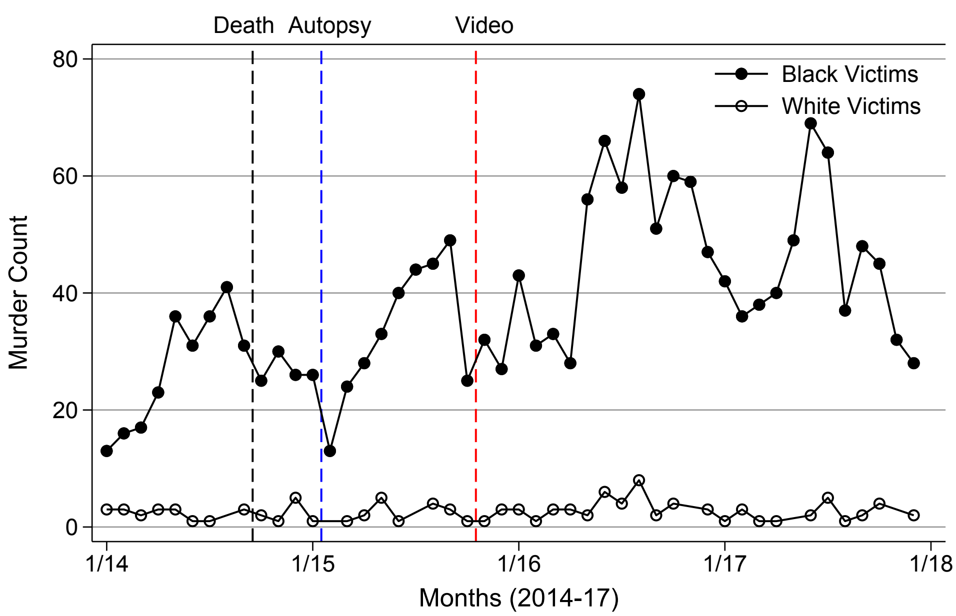
The dotted black line (left) depicts the month of death for Laquan McDonald (October 2014). The dotted blue line (middle) denotes when an article with the associated autopsy report was released (February 2015). The dotted red line (right) highlights when the video of his death was released to the public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

Figure B.23: Murders in Chicago (2014–2017)

(a) Murder



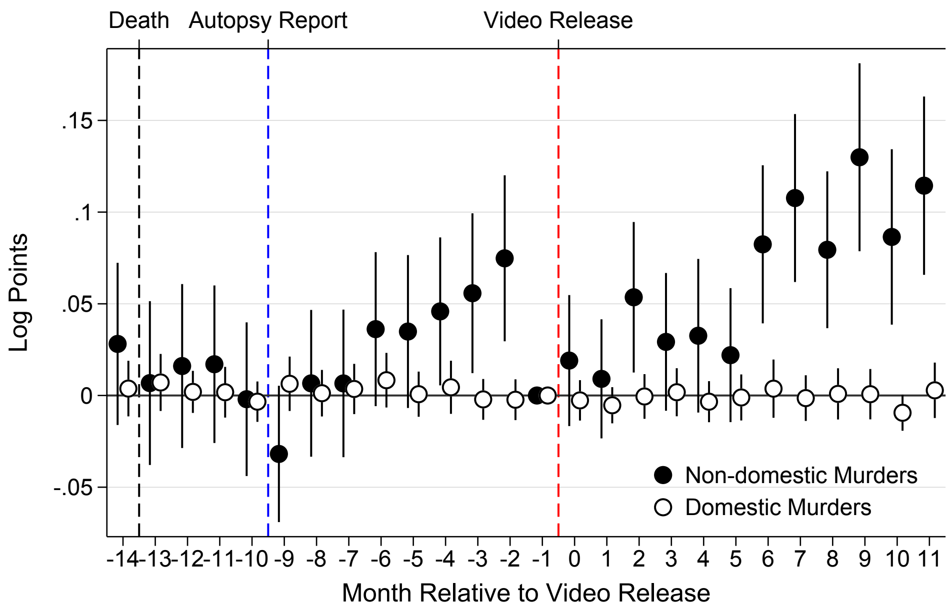
(b) Murder by Race of Victim



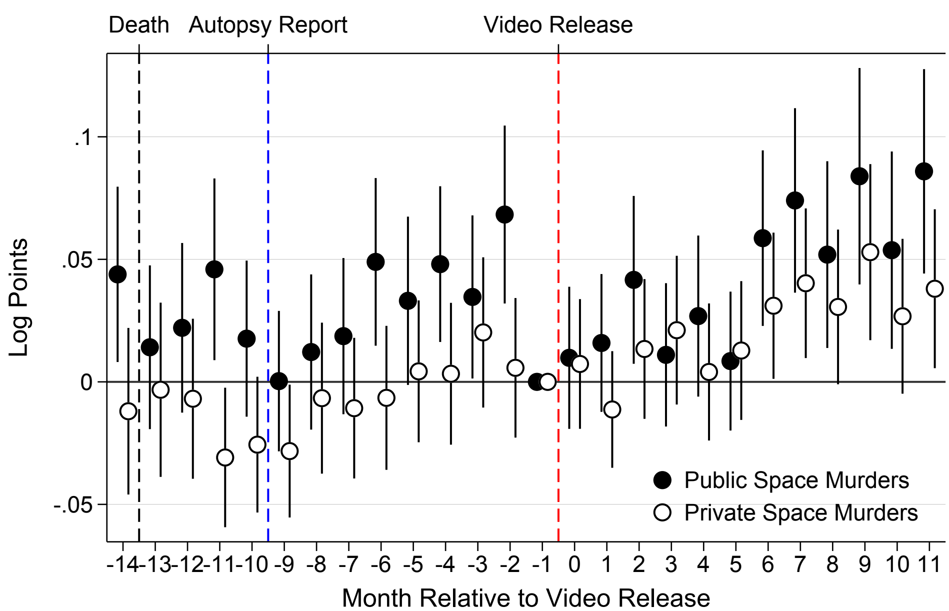
The first figure depicts the unadjusted counts of murders in Chicago from 2014–2017. The second figure separates out the increase by race of victim. To be consistent with the rest of the analysis that contains race data, I focus on deaths where the victim is a Black or White civilian. The dotted black 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).

Figure B.24: Effect Heterogeneity of Laquan McDonald's Death on Murders

(a) Log(Murders) by Victim-Offender Relationship



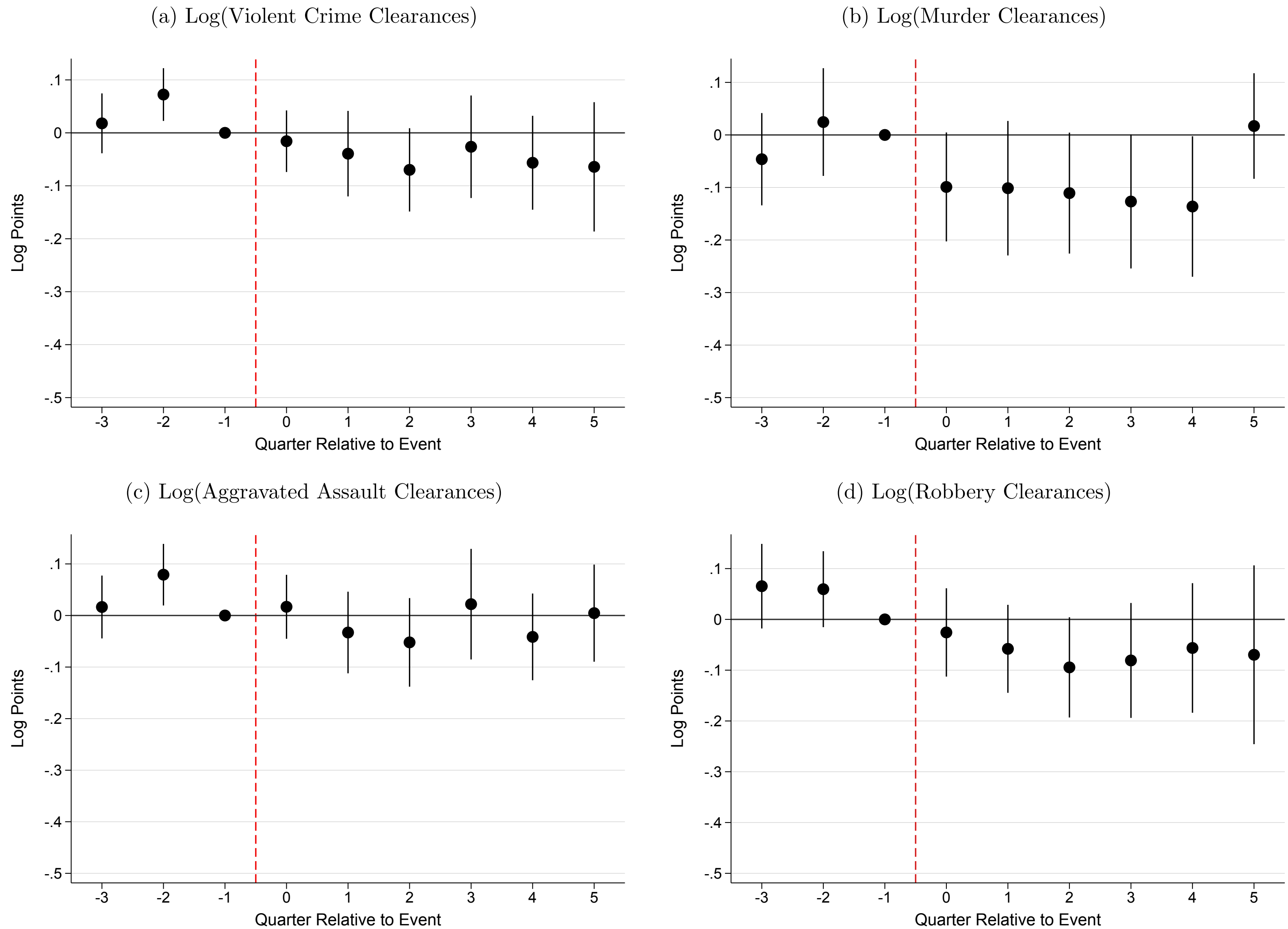
(b) Log(Murders) by Location



In part (a), circles display monthly event time coefficients from a regression of log murders at the beat-month-relationship level in Chicago from 2014–2017, while part(b) involves a regression of log murders at the beat-month-location level. Lines represent the 95% confidence interval using standard errors clustered at the beat level. I control for beat-level fixed events and linear trends, which are interacted with relationship and location indicators respectively.

B.5 Clearance Rate Analysis: Effect of High-Profile Police Killing on Arrests over Reported Crime

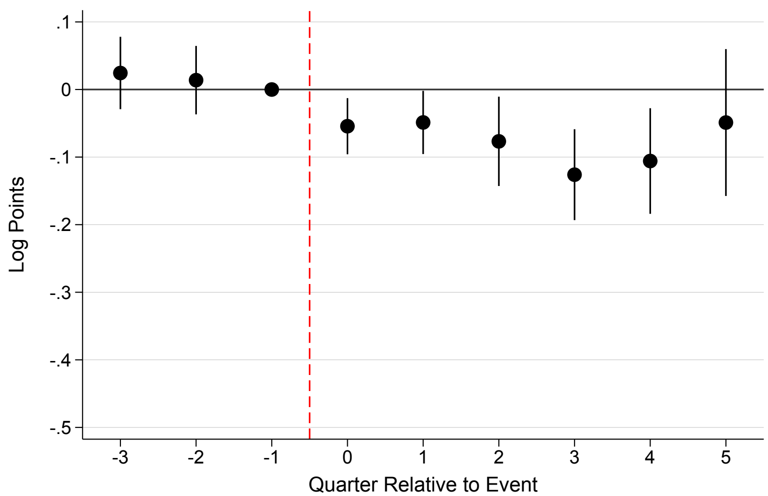
Figure B.25: Effect of High-Profile Police Killing on Violent Crime Clearance



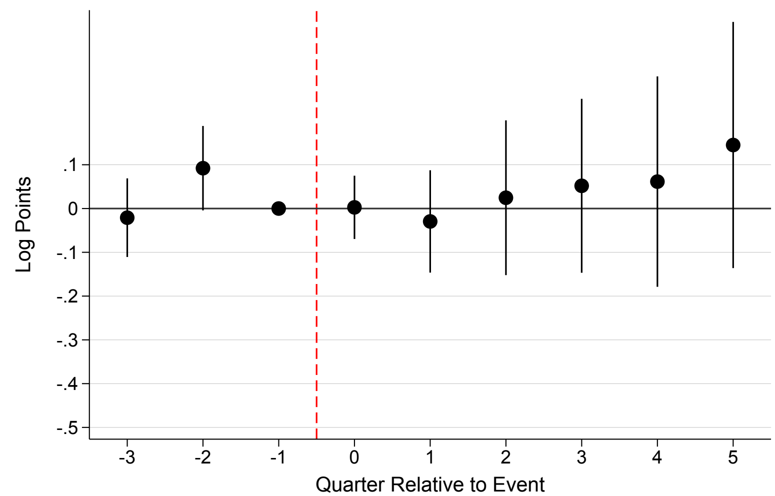
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines are the 95% confidence interval using standard 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; 52 treated; 2,687 control agencies]

Figure B.26: Effect of High-Profile Police Killing on Property Crime Clearance

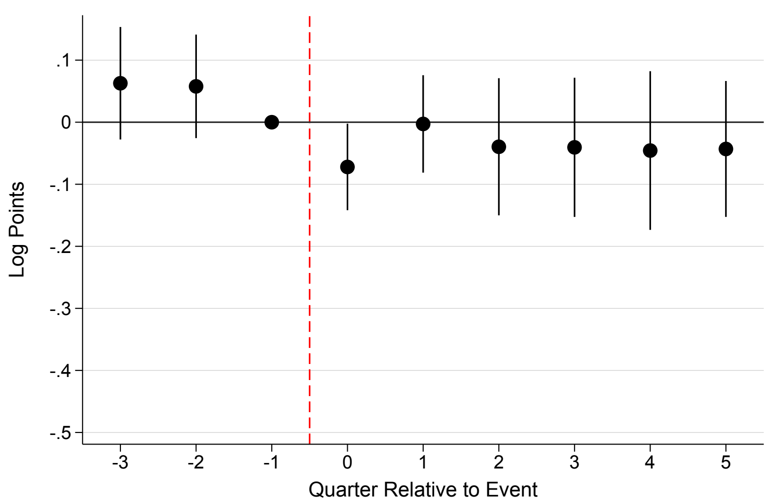
(a) Log(Property Crime Clearances)



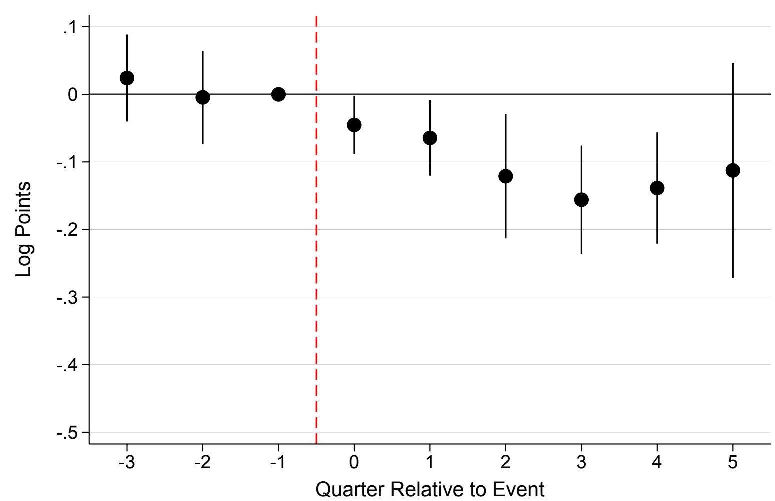
(b) Log(Motor Vehicle Theft Clearances)



(c) Log(Burglary Clearances)



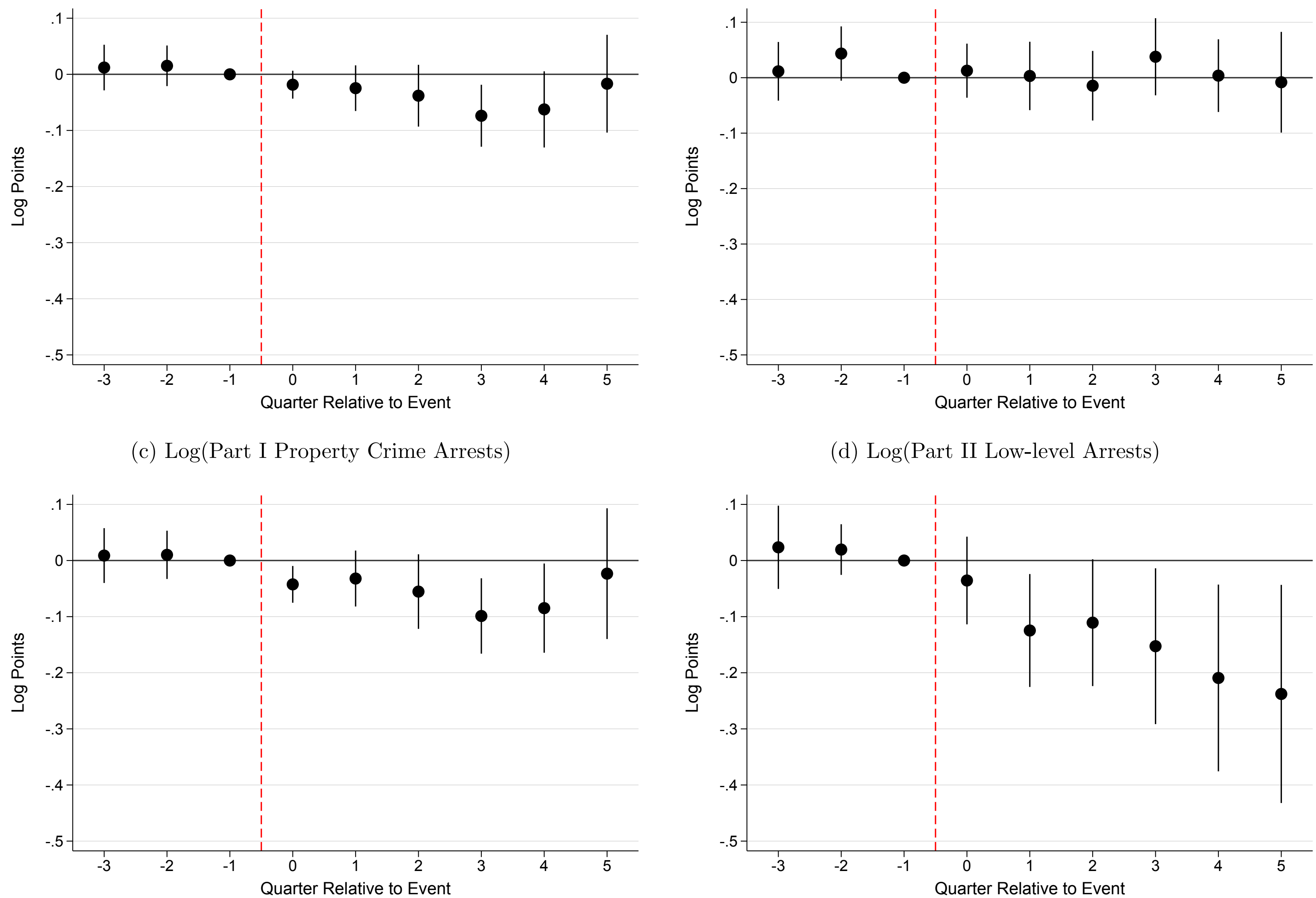
(d) Log(Theft Clearances)



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; 52 treated; 2,687 control agencies]

B.6 Arrest Analysis with Crime Controls: Effect of High-Profile Police Killing on Arrests

Figure B.27: Effect of High-Profile Police Killing on Arrests with Crime Controls



(a) Log(All Part I Arrests)

(b) Log(Part I Violent Crime Arrests)

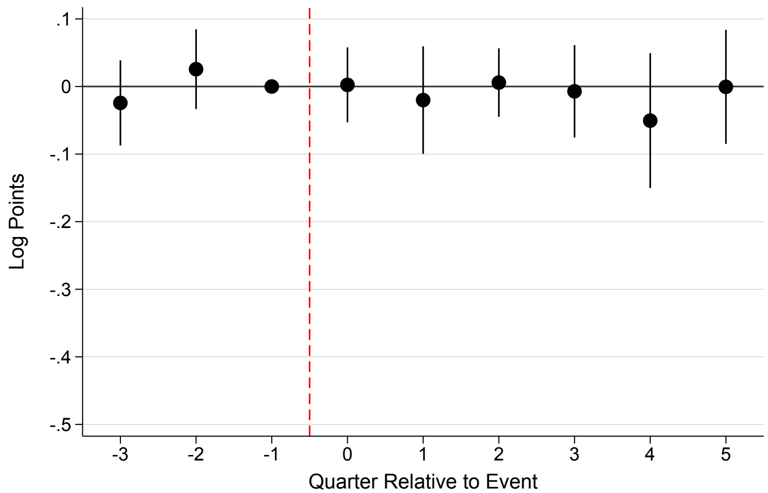
(c) Log(Part I Property Crime Arrests)

(d) Log(Part II Low-level Arrests)

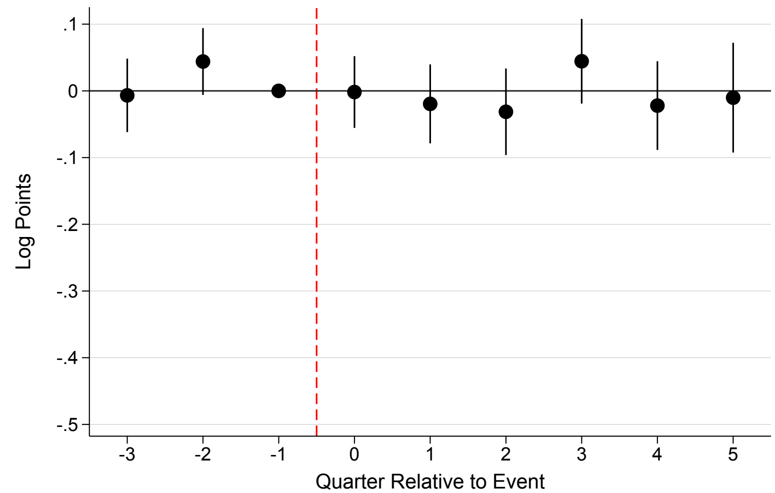
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines are the 95% confidence interval using standard 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; 52 treated; 2,687 control agencies]

Figure B.28: Effect of High-Profile Police Killing on Violent and Property Crime Arrests with Crime Controls

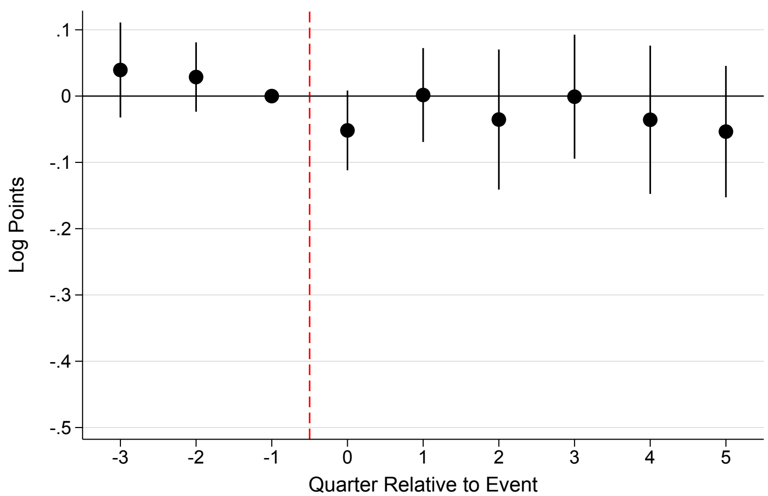
(a) Log(Murder Arrests)



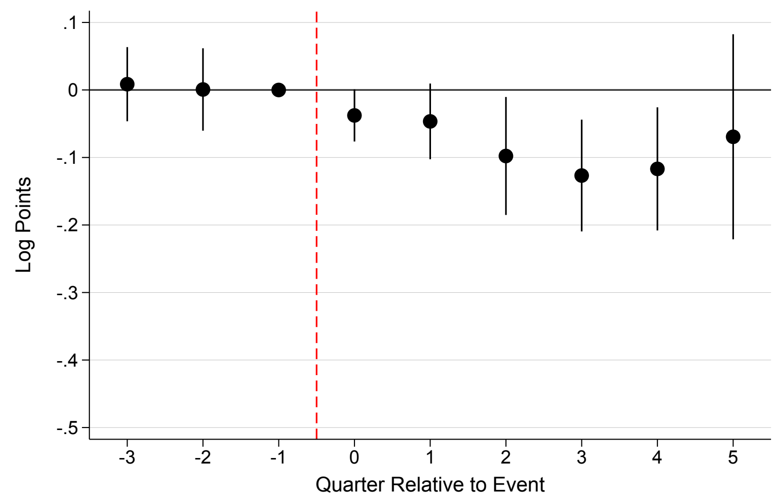
(b) Log(Aggravated Assault Arrests)



(c) Log(Burglary Arrests)



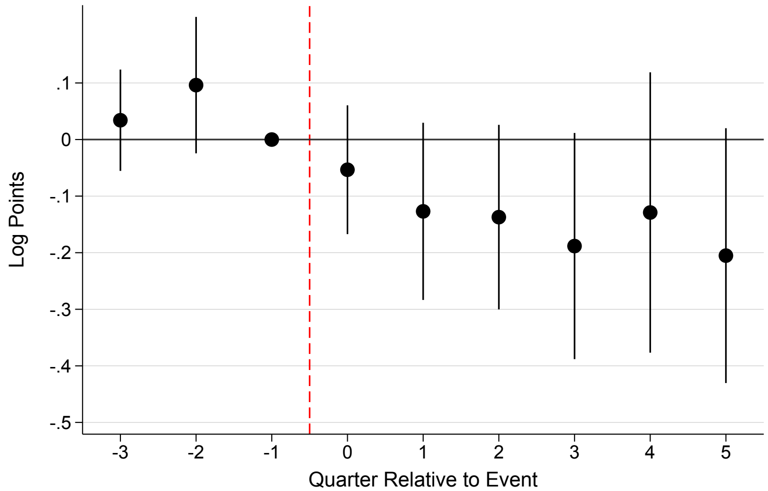
(d) Log(Theft Arrests)



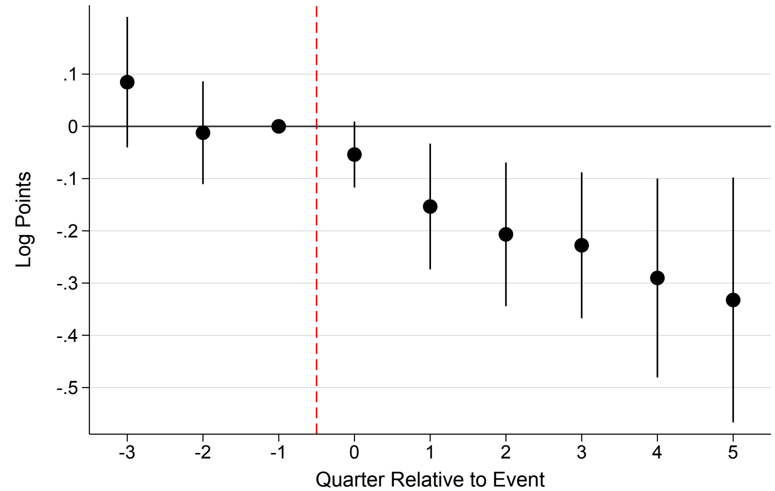
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, 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; 52 treated; 2,687 control agencies]

Figure B.29: Effect of High-Profile Police Killing on Low-level Arrests with Crime Controls

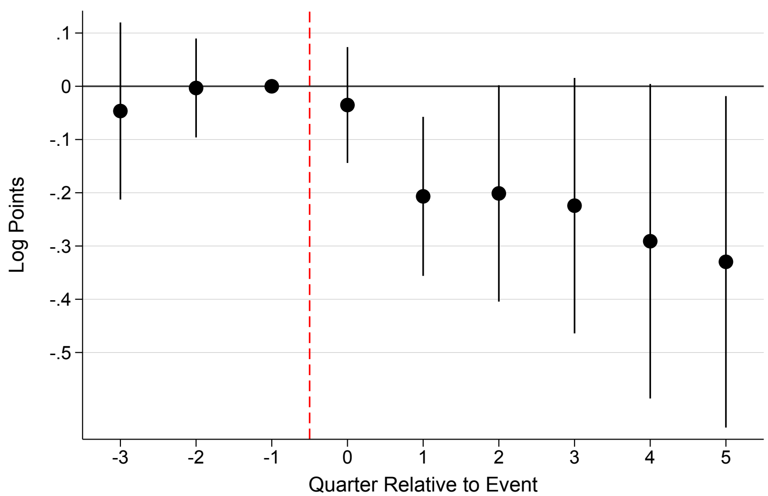
(a) Log(Disorderly Conduct Arrests)



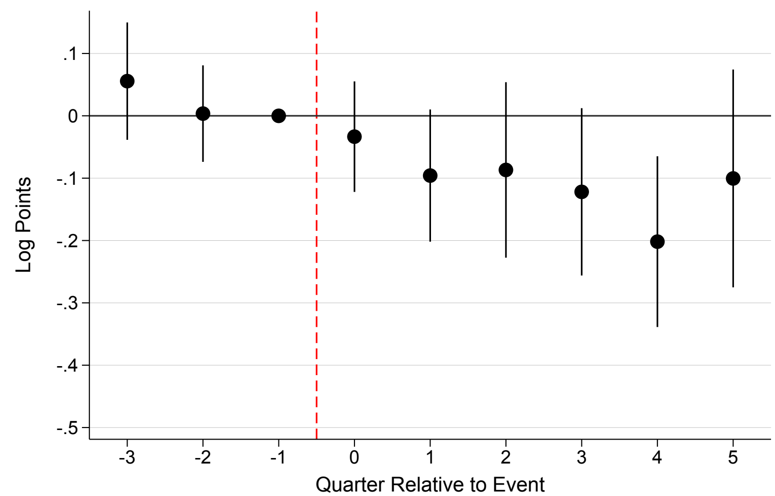
(b) Log(Marijuana Possession Arrests)



(c) Log(Liquor Law Violation Arrests)



(d) Log(Marijuana Sale Arrests)

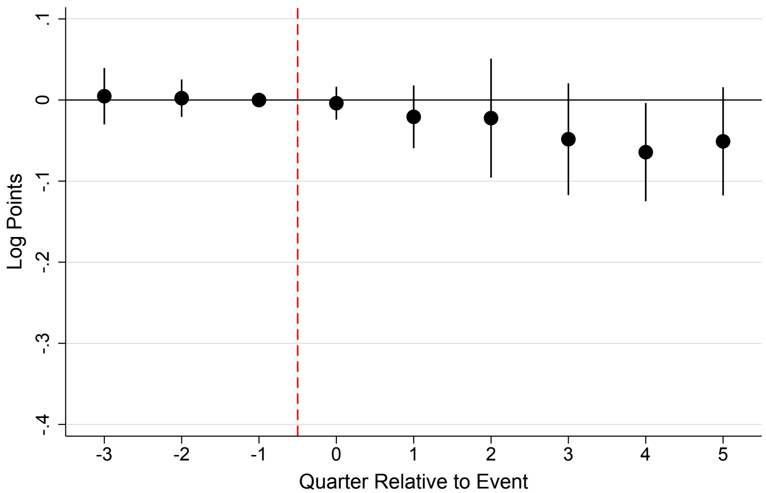


Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard errors clustered at the department level. Each regression controls for population, and violent and property crime. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [599,088 observations; 52 treated; 2,687 control agencies]

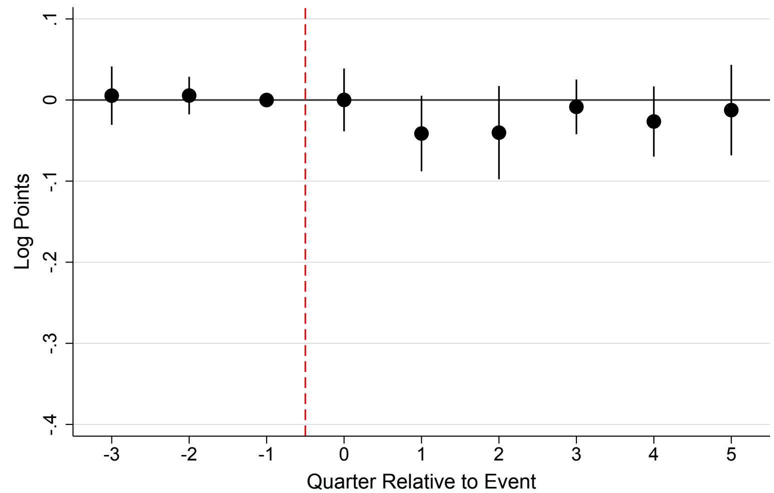
B.7 Results with Population Weighting and No Population Control

Figure B.30: Population Weighting: Effect of High-Profile Police Killing on Arrests

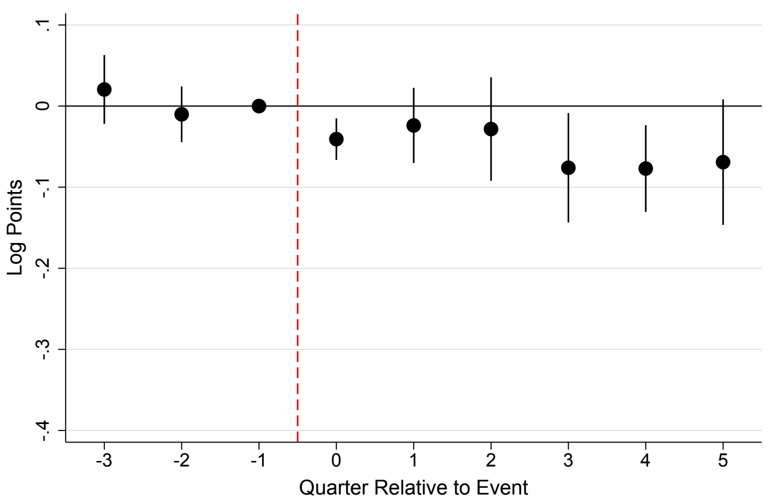
(a) Log(All Part I Arrests)



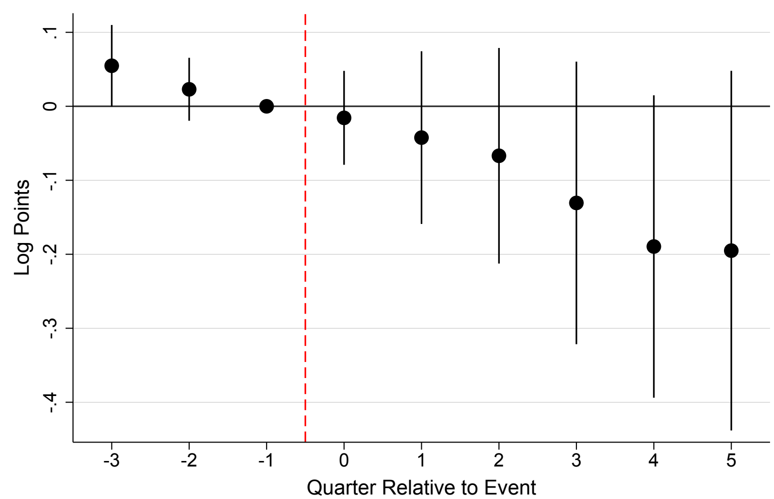
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)

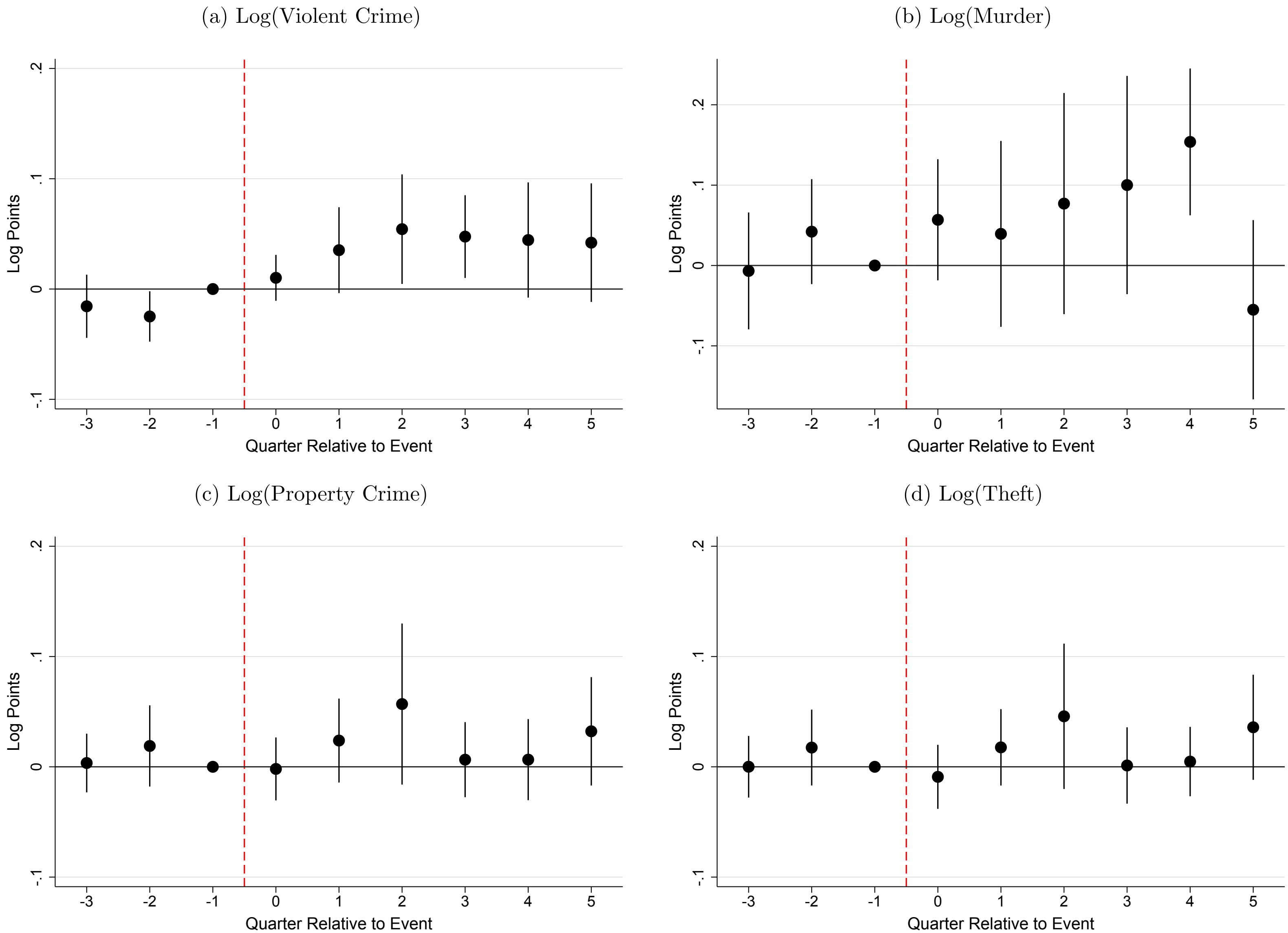


(d) Log(Part II Low-level Arrests)



Using a sample of treated 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 weights the estimates by municipality population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

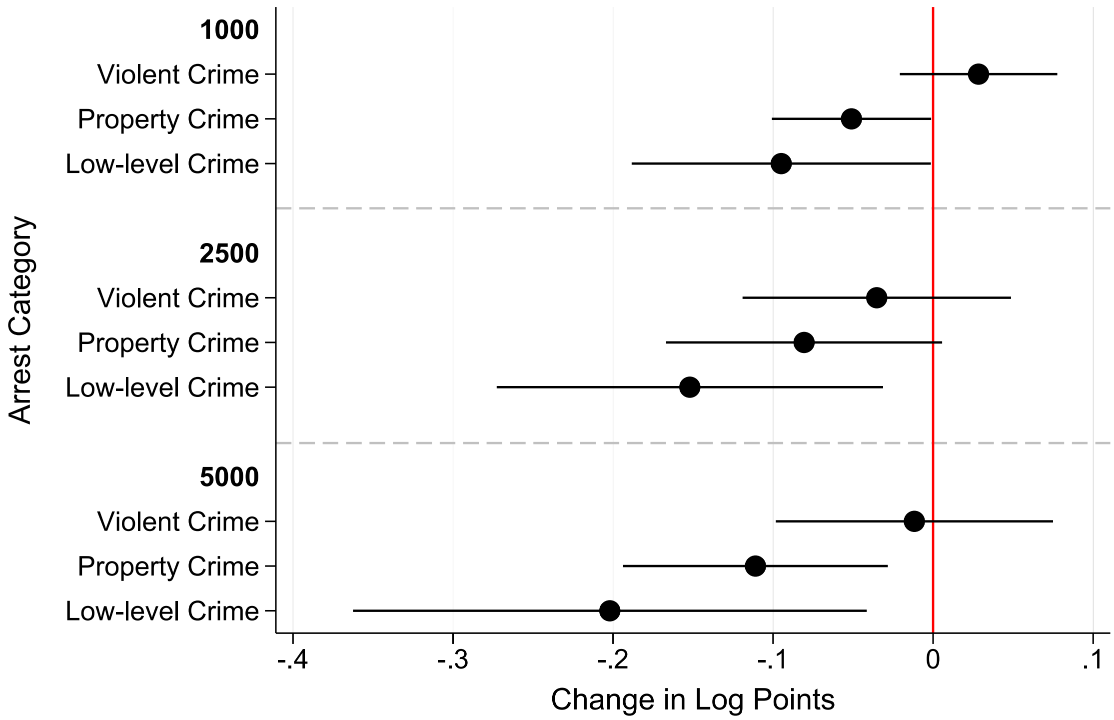
Figure B.31: Population Weighting: Effect of High-Profile Police Killing on Violent and Property Crime



Using a sample of treated 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 weights the estimates by municipality population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [598,788 observations; 52 treated; 2,687 control agencies]

B.8 Effect Robustness and Heterogeneity by Media Article Threshold

Figure B.32: Difference-in-differences (DD) Coefficient Estimates by Arrest Category and Media Article Threshold



Using different thresholds for what fatalities are considered high profile and included in the analysis, the circles display DD coefficients from separate regressions—in descending order of the social cost of crime. There is a significant difference between the average pre-treatment period of control and treated departments for property crime arrests using the 1000 article threshold, but no others. I use a sample of city police departments with fewer than 9 outliers and a population greater 10,000. 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. The arrest DD tables by threshold begin on the next page. The number of officer-involved fatalities used in analysis by threshold: [1,000: 72; 2,500: 32; 5,000: 15]

Table B.12: Effect of the Highest-Profile Police Killing on Log(Violent Crime Arrests) by Media Coverage Threshold

	1000	2500	5000
	(1)	(2)	(3)
Panel A: Violent Crime Arrests			
Treat*Post	0.029	-0.035	-0.012
	(0.026)	(0.043)	(0.044)
Panel B: Murder Arrests			
Treat*Post	0.017	0.031	0.035
	(0.019)	(0.032)	(0.055)
Panel C: Aggravated Assault Arrests			
Treat*Post	0.001	-0.066	-0.003
	(0.028)	(0.046)	(0.044)
Panel D: Robbery Arrests			
Treat*Post	0.027	0.009	-0.031
	(0.024)	(0.032)	(0.043)
Observations	741,532	741,532	741,532
Number of Agencies	2,739	2,739	2,739

Coefficients are from double difference (DD) regressions, using different thresholds for what fatalities are considered high profile [1,000: 72 fatalities; 2,500: 32; 5,000: 15]. There are no significant differences between the average pre-treatment trends of control and treated departments in any these regressions. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.13: Effect of the Highest-Profile Police Killing on Log(Property Crime Arrests) by Media Coverage Threshold

	1000	2500	5000
	(1)	(2)	(3)
Panel A: Property Crime Arrests‡			
Treat*Post	-0.051**	-0.081*	-0.111***
	(0.025)	(0.044)	(0.042)
Panel B: Motor Vehicle Theft Arrests‡			
Treat*Post	-0.010	-0.057	-0.004
	(0.040)	(0.077)	(0.067)
Panel C: Burglary Arrests			
Treat*Post	-0.051	-0.003	0.081
	(0.031)	(0.059)	(0.075)
Panel D: Theft Arrests			
Treat*Post	-0.074**	-0.154**	-0.251***
	(0.034)	(0.063)	(0.084)
Observations	741,532	741,532	741,532
Number of Agencies	2,739	2,739	2,739

‡There is a significant difference between the average pre-treatment trends of control and treated departments for property crime and motor vehicle theft arrests using the 1000 article threshold, but no others.

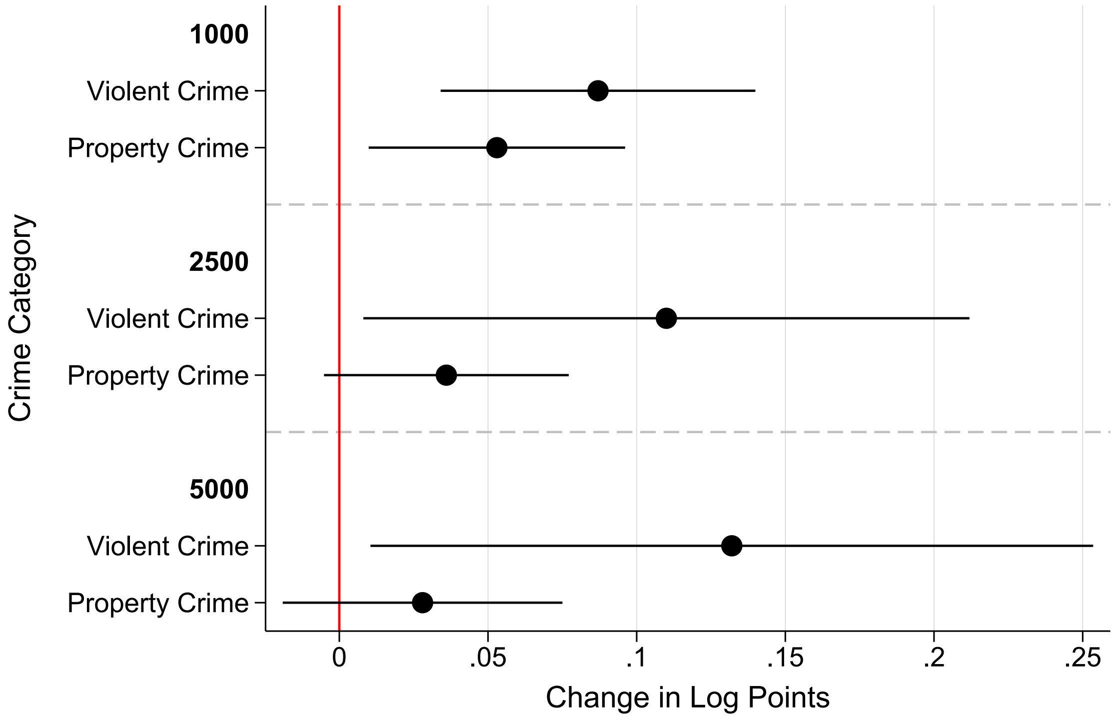
Coefficients are from double difference (DD) regressions, using different thresholds for what fatalities are considered high profile [1,000: 72 fatalities; 2,500: 32; 5,000: 15]. Standard errors, clustered at the agency level, in parentheses. Uses only city police departments with less than 9 outliers and a population greater 10,000. Each regression controls for population. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.14: Effect of the Highest-Profile Police Killing on Log(Low-level Arrests) by Media Coverage Threshold

	1000	2500	5000
	(1)	(2)	(3)
Panel A: Low-level Arrests			
Treat*Post	-0.095**	-0.152**	-0.202**
	(0.048)	(0.062)	(0.082)
Panel B: Marijuana Sale Arrests			
Treat*Post	-0.133***	-0.084	-0.065
	(0.042)	(0.069)	(0.053)
Panel C: Disorderly Conduct Arrests			
Treat*Post	-0.140**	-0.239*	-0.193*
	(0.066)	(0.123)	(0.101)
Panel D: Marijuana Possession Arrests			
Treat*Post	-0.141**	-0.168*	-0.296**
	(0.055)	(0.099)	(0.137)
Observations	741,708	741,708	741,708
Number of Agencies	2,739	2,739	2,739

Coefficients are from double difference (DD) regressions, using different thresholds for what fatalities are considered high profile [1,000: 72 fatalities; 2,500: 32; 5,000: 15]. There are no significant differences between the average pre-treatment trends of control and treated departments in any these regressions. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.33: Difference-in-differences (DD) Coefficient Estimates by Crime Category and Media Article Threshold



Using different thresholds for what fatalities are considered high profile and included in the analysis, the circles display DD coefficients from separate regressions—in descending order of the social cost of crime. There are no significant differences between the average pre-treatment period of control and treated departments. I use a sample of city police departments with fewer than 9 outliers and a population greater 10,000. 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. The crime DD tables by threshold begin on the next page. The number of officer-involved fatalities used in analysis by threshold: [1,000: 72; 2,500: 32; 5,000: 15]

Table B.15: Effect of the Highest-Profile Police Killing on Log(Violent Crime) by Media Coverage Threshold

	1000	2500	5000
	(1)	(2)	(3)
Panel A: Violent Crime			
Treat*Post	0.087***	0.110**	0.132**
	(0.027)	(0.052)	(0.062)
Panel B: Murder			
Treat*Post	0.144***	0.192***	0.267***
	(0.036)	(0.038)	(0.052)
Panel C: Aggravated Assault			
Treat*Post	0.041*	0.053	0.114**
	(0.024)	(0.045)	(0.051)
Panel D: Robbery			
Treat*Post	0.125***	0.147***	0.132*
	(0.030)	(0.050)	(0.068)
Observations	741,532	741,532	741,532
Number of Agencies	2,739	2,739	2,739

Coefficients are from double difference (DD) regressions, using different thresholds for what fatalities are considered high profile [1,000: 72 fatalities; 2,500: 32; 5,000: 15]. There are no significant differences between the average pre-treatment trends of control and treated departments in any these regressions. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B.16: Effect of the Highest-Profile Police Killing on Log(Property Crime) by Media Coverage Threshold

	1000	2500	5000
	(1)	(2)	(3)
Panel A: Property Crime			
Treat*Post	0.053**	0.036*	0.028
	(0.022)	(0.021)	(0.024)
Panel B: Motor Vehicle Theft			
Treat*Post	0.079***	0.063*	0.026
	(0.029)	(0.034)	(0.044)
Panel C: Burglary			
Treat*Post	0.039	0.086**	0.118**
	(0.027)	(0.036)	(0.047)
Panel D: Theft‡			
Treat*Post	0.051**	0.002	-0.014
	(0.026)	(0.026)	(0.038)
Observations	741,532	741,532	741,532
Number of Agencies	2,739	2,739	2,739

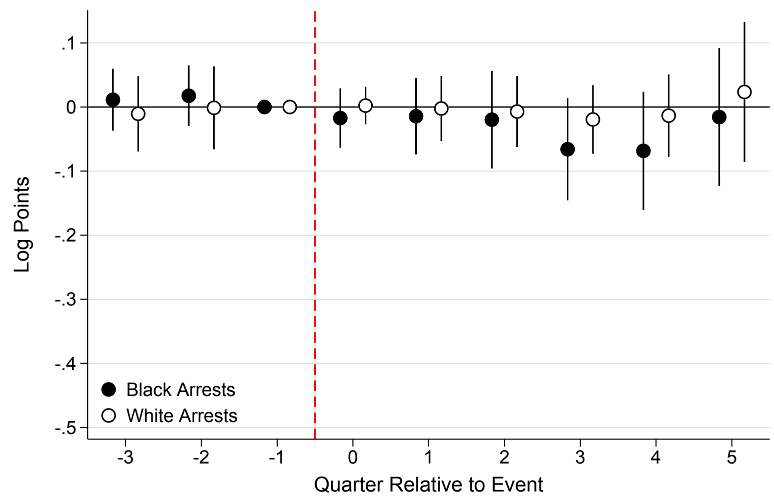
‡There is a significant difference between the average pre-treatment trends of control and treated departments for theft using the 5000 article threshold, but no others.

Coefficients are from double difference (DD) regressions, using different thresholds for what fatalities are considered high profile [1,000: 72 fatalities; 2,500: 32; 5,000: 15]. 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

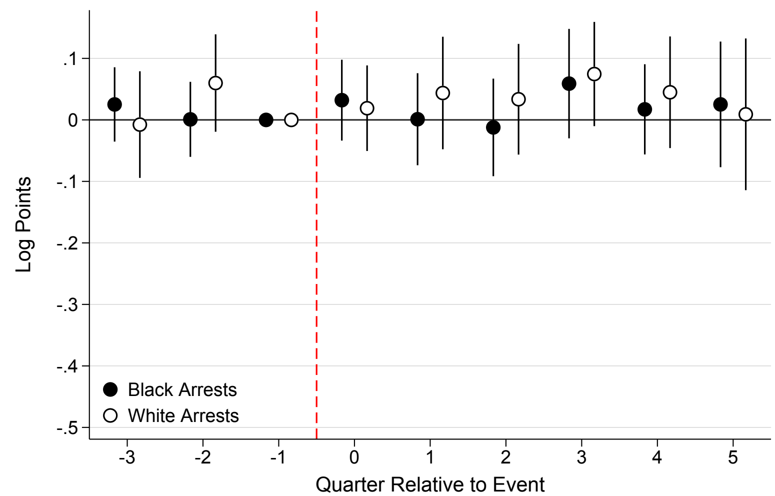
B.9 Effect Heterogeneity by Race of Suspect

Figure B.34: Effect of High-Profile Police Killing on Race-Specific Arrest Categories

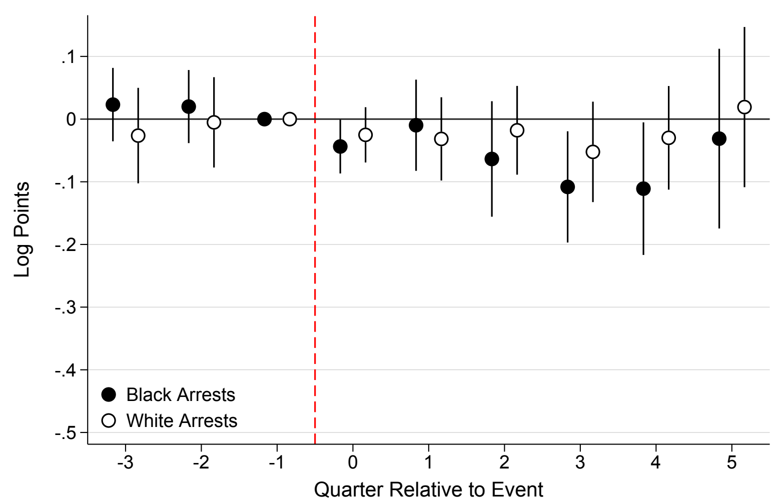
(a) Log(All Part I Arrests)



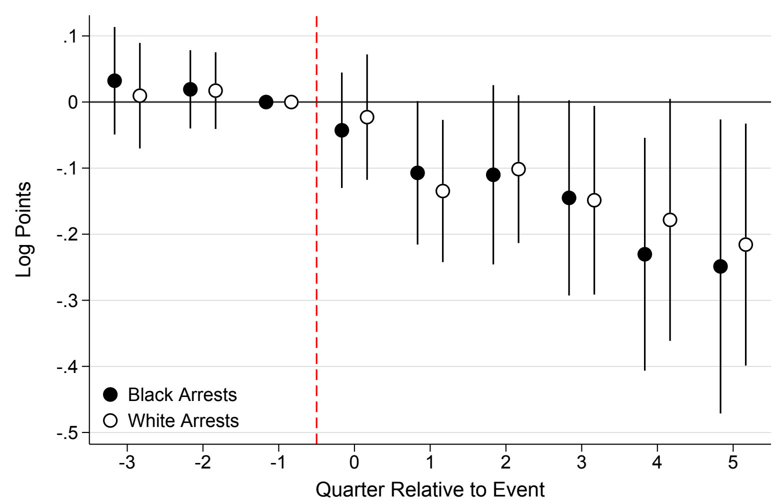
(b) Log(Part I Violent Crime Arrests)



(c) Log(Part I Property Crime Arrests)



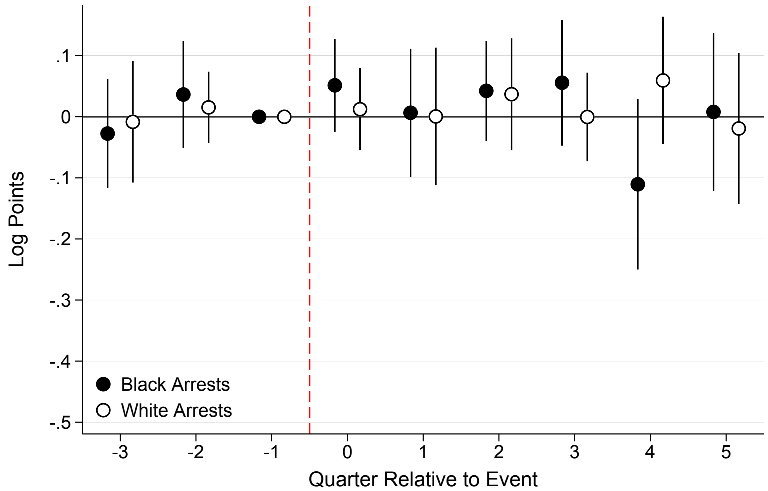
(d) Log(Part II Low-level Arrests)



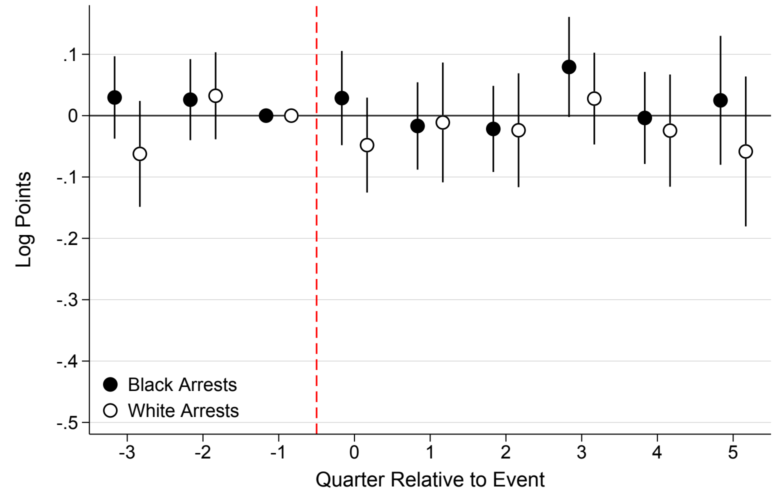
Using a sample of city police departments with over 10,000 people and fewer than 9 outliers from 2005–2016, the event time coefficients (circles) right of the red dotted line are the change in arrests after a high-profile, officer-involved fatality. Lines represent the 95% confidence interval using standard 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 B.35: Effect of High-Profile Police Killing on Race-Specific Violent and Property Arrests

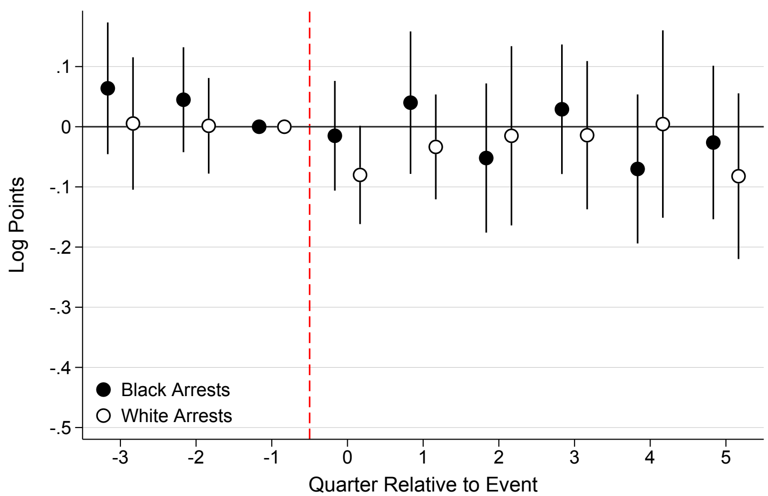
(a) Log(Murder Arrests)



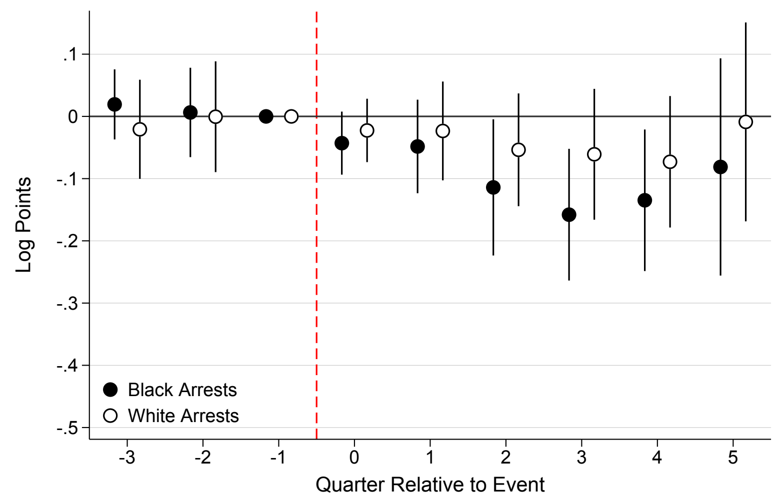
(b) Log(Aggravated Assault Arrests)



(c) Log(Burglary Arrests)

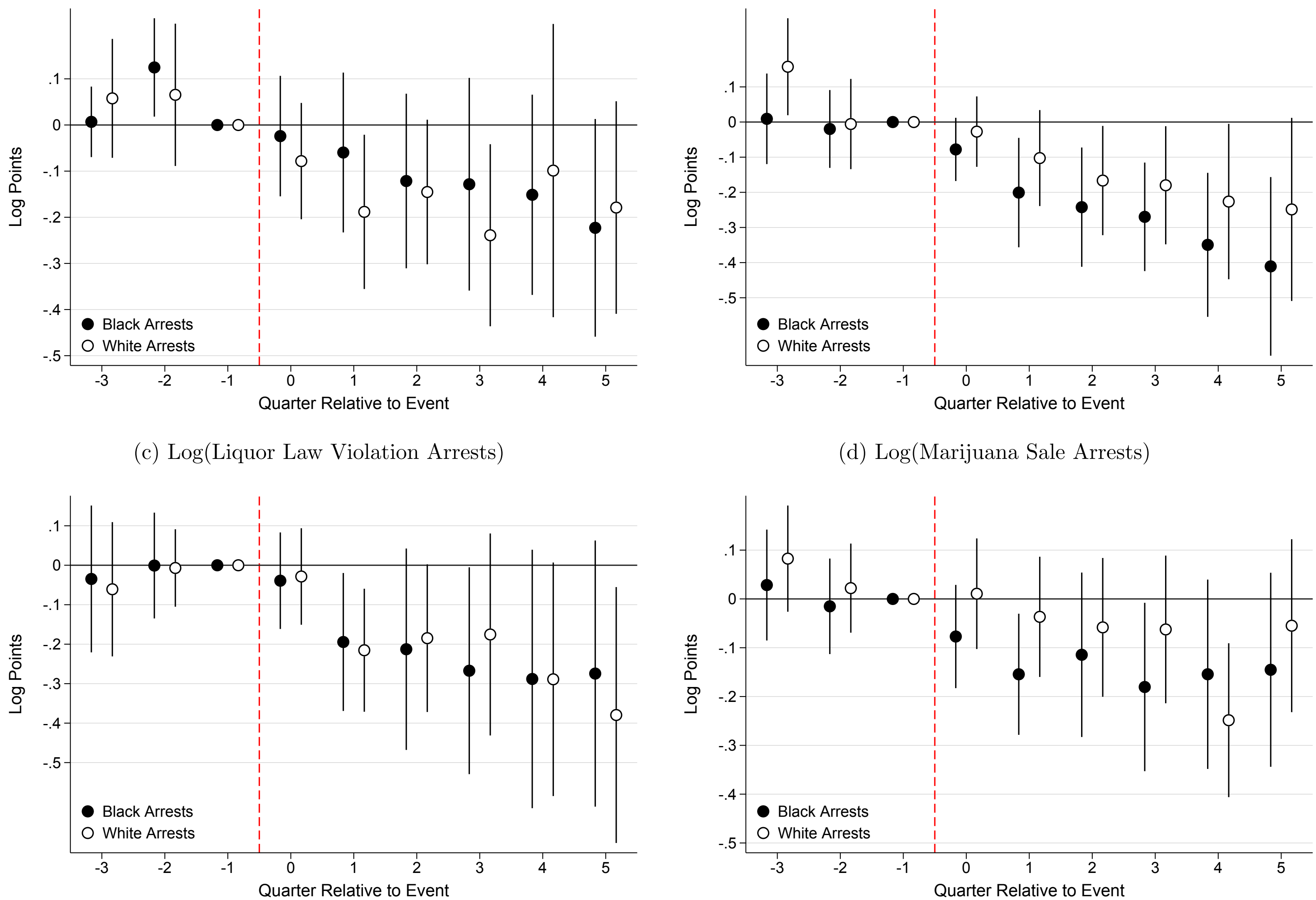


(d) Log(Theft Arrests)



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. 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 B.36: Effect of High-Profile Police Killing on Race-Specific Low-level Arrests



(a) Log(Disorderly Conduct Arrests)

(b) Log(Marijuana Possession Arrests)

(c) Log(Liquor Law Violation Arrests)

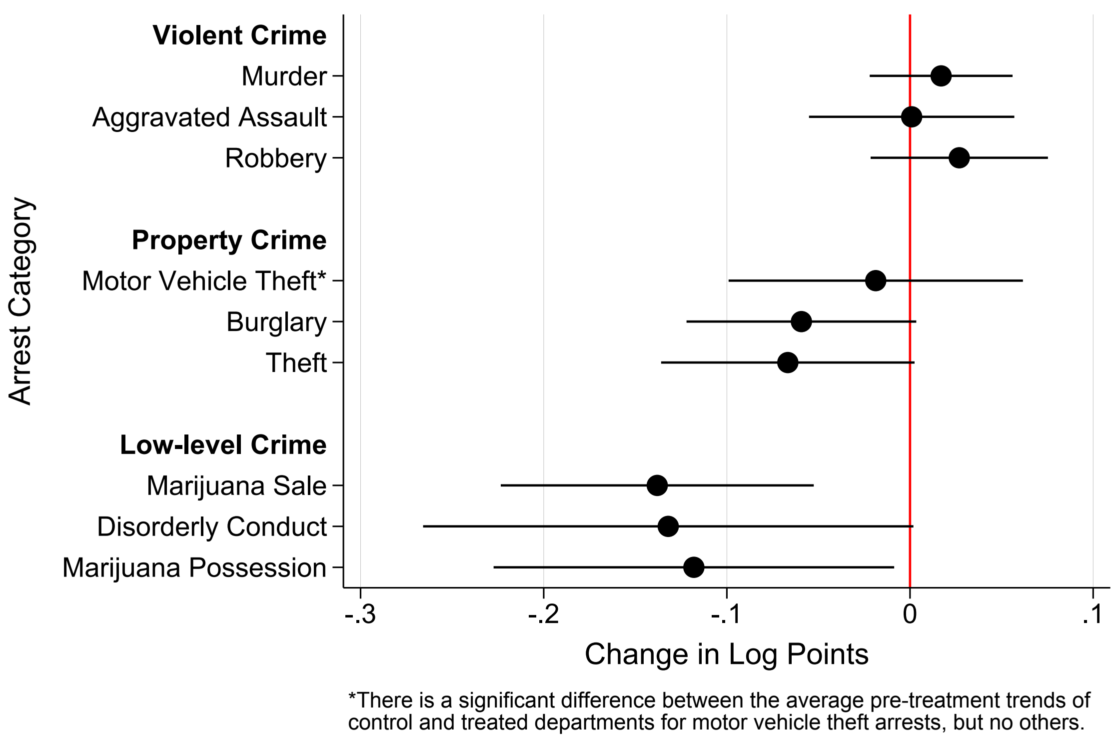
(d) Log(Marijuana Sale Arrests)

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. 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]

B.10 Effect Robustness to Excluding Agencies with Pattern-or-Practice (PoP) Investigations

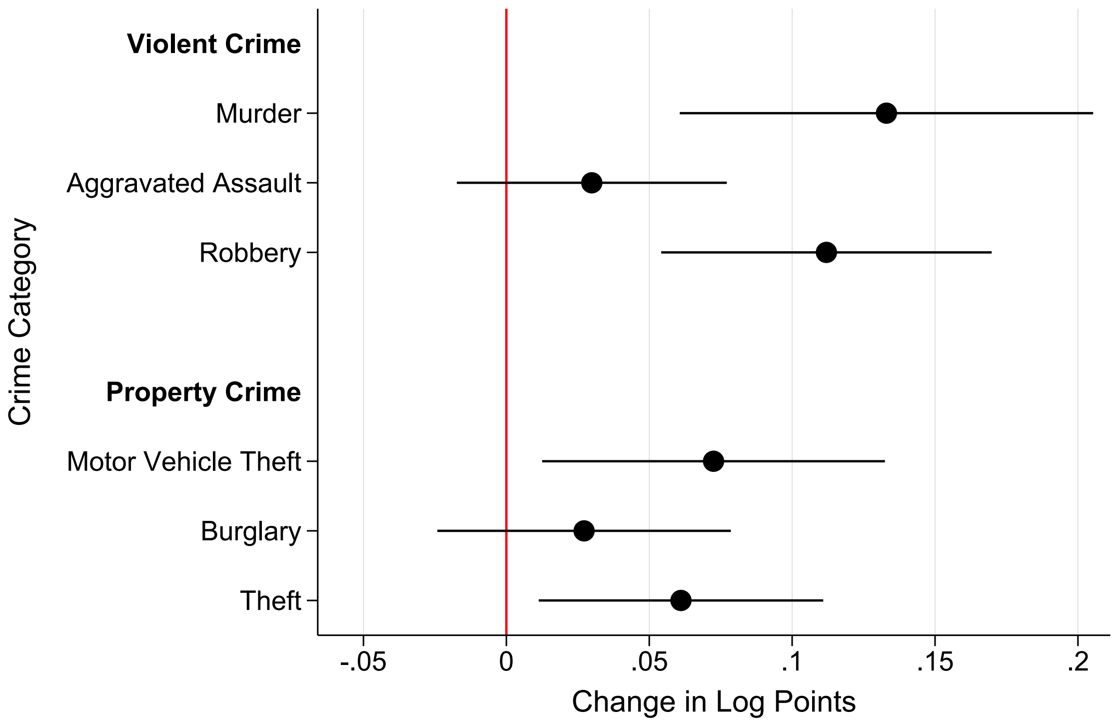
B.10.1 Excluding agencies whose PoP investigation is plausibly instigated by a high-profile police killing

Figure B.37: Difference-in-differences (DD) Coefficient Estimates by Arrest Category



This analysis sample excludes agencies whose PoP investigation is plausibly instigated by a high-profile police killing. 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. 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. [740,318 observations; 50 treated; 2,687 control agencies]

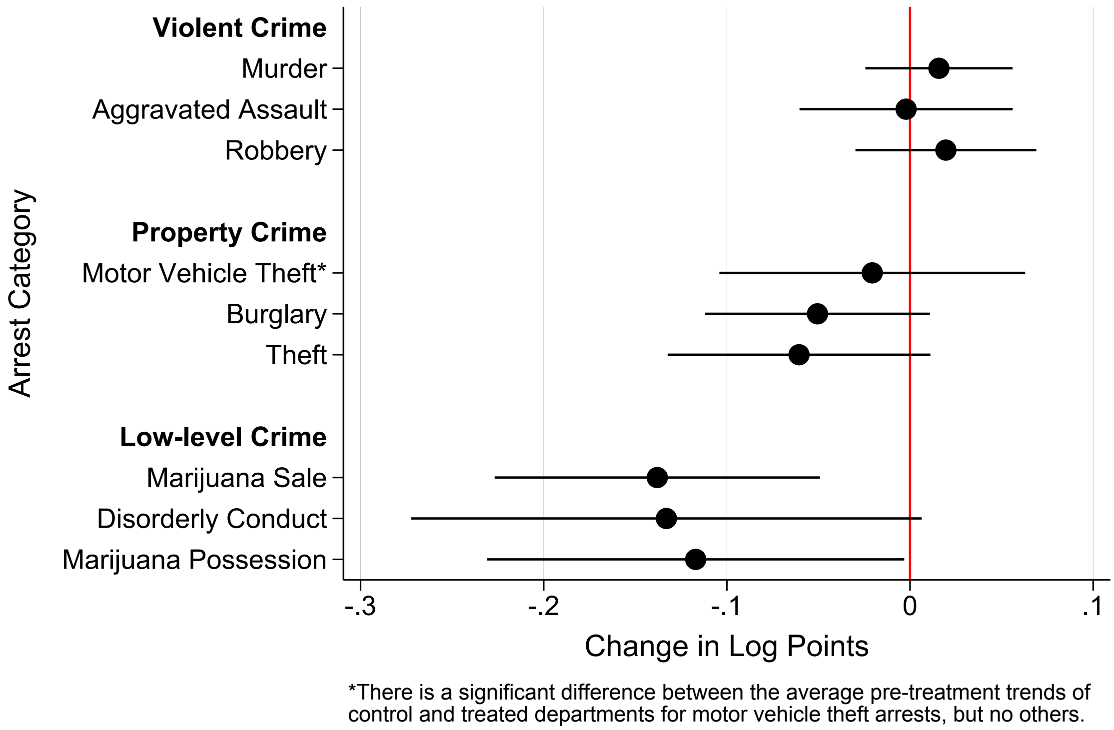
Figure B.38: Crime Analysis: Difference-in-differences (DD) Coefficients by Crime Category



This analysis sample excludes agencies whose PoP investigation is plausibly instigated by a high-profile police killing. 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 average 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [740,318 observations; 50 treated; 2,687 control agencies]

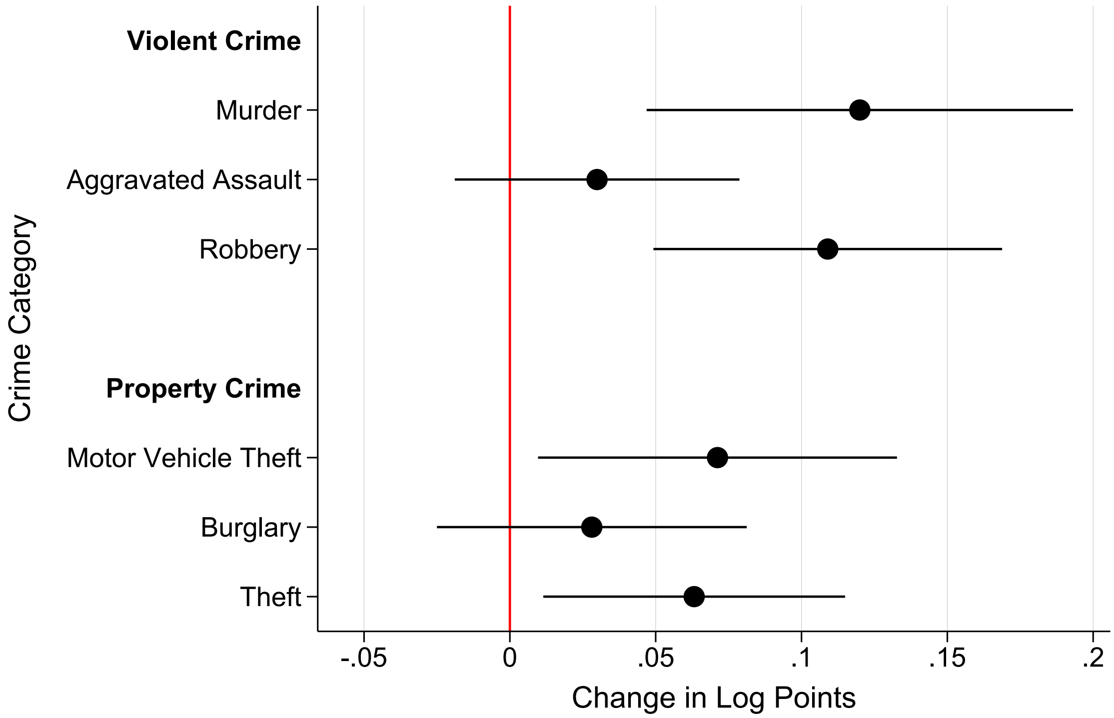
B.10.2 Excluding agencies who experience a PoP investigation from 2004–2016

Figure B.39: Difference-in-differences (DD) Coefficient Estimates by Arrest Category



This analysis sample excludes any agencies that experience a PoP investigation from 2004–2016. 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. 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. [737,348 observations; 47 treated; 2,679 control agencies]

Figure B.40: Crime Analysis: Difference-in-differences (DD) Coefficients by Crime Category



This analysis sample excludes any agencies that experience a PoP investigation from 2004–2016. 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 average 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. All regressions have month-of-sample and department fixed effects, as well as linear county trends. [737,348 observations; 47 treated; 2,679 control agencies]