THE UNIVERSITY OF CHICAGO

ESSAYS IN PERSONNEL ECONOMICS OF POLICING

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
TAEHO KIM

CHICAGO, ILLINOIS
JUNE 2021

Copyright © 2021 by Taeho Kim

All Rights Reserved

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This work would not have existed without the help of many people. I would like to thank my family and friends in the US and Korea for their endless love and support throughout my academic journey in the US.

My graduate education is made possible through some perseverance and resilience, which I owe most to my parents who have inspired those virtues in me from early on in my life. I am also thankful that my brother Taehyung was in Chicago in the early stage of my dissertation, during which I received crucial feedback and tremendous encouragement. I am also very lucky to have met my wife Aileen during my graduate studies who have given me enormous moral support and have been by my side through the ups and downs of the dissertation stage.

I am indebted to my thesis committee Steve Levitt, Michael Greenstone, Heather Sarsons, and Alessandra Voena. They have been central for my growth at the University of Chicago. I appreciate their being extremely generous with their time and resources and patient in helping me grow intellectually. I greatly benefited from their advice, critiques, and encouragement on this project. I would also like to thank my mentors from college – Doug Staiger, Bruce Sacerdote, and Taryn Dinkelman for instilling my love for research in applied microeconomics.

Many thanks to all professors at the University of Chicago for furthering my passion for economics, giving me world-class education in economics, and providing me with helpful feedback. I am thankful to have made friends with talented peers during my graduate program – I have learned just as much through interactions with them as in the classroom. The Economics Department and the Social Sciences Division at the University of Chicago have provided me with generous funding.

I am grateful to many others who have given me the necessary tools to write the dissertation. William McCarty familiarized me with the body camera literature, and officers at the Chicago Police Department and Schaumburg Police Department, IL, provided me with institutional background. All errors are my own.

ABSTRACT

Controversial police use-of-force incidents have eroded police legitimacy, and body-worn cameras (BWCs) have received extensive attention as a key reform. In the first chapter, I study the causal effects of BWCs on the use of force and law enforcement outcomes. Previous studies that randomized BWC deployment at the officer level within a single agency faced empirical challenges as (1) the control group officers are also indirectly affected by BWCs due to interactions with the treatment group officers (spillover), (2) there may be fundamental differences between agencies that agree to be researched and agencies that do not (site-selection bias), and (3) researchers could not directly examine agency-wide variables such as crime rates and public opinion. I overcome these limitations by conducting the first nationwide study of BWCs across more than 1,000 agencies in the US. I find that BWCs lead to substantial decreases in the use of force, both against whites and minorities. Nationwide, they reduce police-involved homicides by 58%. In contrast to previous studies on police accountability, I find no evidence of an association between police oversight through BWCs and reduction in policing efforts. By examining social media usage from Twitter, I find that BWC adoption has improved public opinion toward the police. These findings imply that BWCs can be an important tool for improving police accountability without sacrificing policing capabilities.

In the second chapter, I include my work from another related line of my research that aims to understand of bureaucrats' incentives and their decision-making. Police agencies, like other public bureaucracies, have rigid promotion structure. Combining the rigid structure based on exam or tenure with one based on merit may incentivize better police behavior or retain officers. In the Chicago Police Department, eligibility requirement in service length expedited the first promotion opportunity to officers with enough tenure by seven years relative to officers with slightly less tenure who are otherwise similar. I use this discontinuity in promotion chance to examine the effects of promotion opportunities on performance and career decisions. Overall, I find that the promotion opportunity led to less police misconducts, while I do not find evidence of changes in arrest performance or retention. I also find that it encouraged officers to invest in their career capital by joining specialized units, where promotion occurs more frequently. The results suggest a more flexible promotion policy based on merit by an evaluation board could be used to complement rigid promotion structures to improve bureaucratic behavior.

CHAPTER 1

FACILITATING POLICE REFORM: BODY CAMERAS, USE OF FORCE, AND LAW ENFORCEMENT OUTCOMES

1.1 Introduction

Recent high-profile and controversial police use-of-force incidents have spurred protests across the nation and calls for police reform. In ensuing debates, officer body-worn cameras (BWCs) have been examined as a key to police reform by providing video documentation of police encounters with community members. BWCs have the potential to significantly affect policing by strengthening accountability and promoting community relations.

However, the question of whether BWCs are a viable solution to the police legitimacy crisis is far from settled. BWCs may lead to reduction in policing efforts as officers become more afraid of making errors and become less proactive in crime control activities. Moreover, BWCs may not lead to reduction in the use of force if there is no margin of improvement in current police practices or police unions are so powerful that officers who are found to have used excessive force through BWCs do not face real consequences.

In this paper, I study the causal effects of BWCs on the use of force, enforcement outcomes, and public opinion toward the police in the first nationwide study of over 1,000 local police agencies in the US. I examine the staggered adoption of BWCs in a quasi-experimental difference-in-difference (DID) approach. My empirical method relies on the idiosyncratic timing of BWC adoption that are attributable to administrative hurdles in the adoption process. Anecdotal evidence suggests that as agencies rushed to adopt BWCs starting in 2014, they faced bureaucratic delays in implementation that stemmed from uncertainties about funding and BWC policies. Consistent with this, I find that while there are fundamental differences between agencies that do and do not adopt BWCs, among adopters of BWCs, characteristics of police agencies are not related to when they adopted BWCs. While selection into BWC adoption may be endogenous, a key assumption of this approach is that the timing of BWC adoption is uncorrelated with other determinants of changes in the use of force.

To leverage the variations in adoption timing, I gather data on agency-level use of force as well as data on BWC adoption. My data on BWC adoption status come from the Body Worn Camera Supplement to the Law Enforcement Management and Administrative Statistics (LEMAS) Survey, a novel survey by the Bureau of Justice Statistics that asks police chiefs in the US if they have adopted BWCs, and if so, in which year and what month. At the national level, I gather data on police-involved homicides from 2013 to 2019. I aggregate the incident-level observations at the agency-month level and merge them with time-varying adoption status.

In my national analysis on police-involved homicides, I find that BWCs contribute to substantial reductions (58%) in these incidents in an immediate trend break after BWC adoption. When I further disaggregate these results according to the races of the victims, I find that this decrease can be attributed to both white and minority populations. Previous studies have largely neglected disaggregated analysis of the use of force by race. However, this finding is especially relevant given the recent controversy surrounding police use-of-force incidents, in which the victims tend to be minorities.

I use several strategies to rule out alternative explanations for the reduction of the use of force. First, I use high-frequency monthly data to isolate BWC adoption from other concurrent policies that may have been adopted together. As it typically takes about 18 months to implement BWCs, the monthly data allows me to observe the time trends of police-involved homicides near the month of adoption at a granular level. Other confounding reforms such as training and changes in use-of-force policies are unlikely to generate the trend break in the use of force immediately following adoption because they do not involve bureaucratic processes that are necessary for BWC implementation and thus would not have implementation schedules that precisely overlap with that of BWCs. Second, I more directly test whether BWC adoption coincided with other interventions, I gather data on purchase orders of police training. Additionally, I analyze the possibility that agencies with high logistical capacities can quickly implement BWC programs as an immediate response to a scandal. I run robustness checks excluding agencies that indicated they did not face obstacles to implementation and find that the results are not driven by such agencies.

To shed light on possible explanations for the reduction in the use-of-force incidents, I consider three factors which I describe in a model of police behavior: disengagement by police, improved police tactics, and improved citizen behavior. I find that an explanation based on improved police tactics is most consistent with my empirical results. To check whether disengagement has driven the reduction in the use of force, I test whether BWCs have led to decreases in arrests or changes in crime rates; I do not find evidence for either. This result is surprising given that previous studies on police oversight have found that public scrutiny on the police can lead to severe police disengagement (Prendergast, 2001, Shi, 2009, and Devi and Fryer Jr., 2020). Improved citizen behavior is also unlikely to drive the reduction in the use of force as I do not find that BWCs have led to measurable changes in assaults on the police. On the other hand, an explanation based on improved police tactics is most consistent with my empirical results. Finally, I investigate whether BWCs have reduced the use of force by improving police tactics using the LEMAS survey of police chiefs which directly asks about the ways in which BWCs have been useful. I find that the majority of police chiefs who have adopted BWCs agree that BWCs help improve the professionalism of the officers and BWCs help facilitate officer training. In accordance with their answers, I find that agencies that agreed that BWCs pose these benefits experience greater reductions in the use of force than those that do not.

I also explore if BWCs have also led to changes in the use of force less severe than deadly force, by supplementing my national analysis with an analysis of lower-level use of force in New Jersey. I gather administrative data on all use of force in New Jersey from 2012 to 2016.

I find that in New Jersey, implementation of BWCs has decreased the total use of force by 20% and subject injuries by 42% following the implementation of BWCs.

Public perceptions are an important measure of police performance. Distrust in the police can interfere with proper crime control and lead to rises in crime levels. However, it is difficult to measure changes in public opinion using conventional data sources such as surveys because they do not capture high-frequency variations over wide geographic regions. I overcome this issue by gathering police-related tweets on Twitter. Using these data sources, I find that BWCs lead to improved sentiments on Twitter. After BWC adoption, total tweets about police drop by 21% and I gather suggestive evidence that sentiments also improve.

In the final section of my analysis, I interpret the reductions in the use of force in the context of social welfare. Using a monetized value of averted police-involved homicides and my estimates, I find that BWC programs lead to an increase in social welfare.

My national study on BWCs overcomes limitations in previous studies that have evaluated BWCs in single-agency settings by randomizing the treatment of equipping BWCs at the officer level in the same agency (Ariel, 2016; Braga et al., 2018; Jennings et al., 2015; Yokum et al., 2019). While this research has revealed a wealth of information about various aspects of BWCs, the single-agency research design presents several shortcomings. For instance, this study design does not adequately account for spillover effects on control group officers who are also indirectly affected by working closely with treatment group officers. Moreover, these studies only examine agencies willing to cooperate with researchers, and as such, it may be difficult to generalize from them due to site-selection bias. Also, my cross-agency analysis allows me to study agency-wide performance measures such as crime rates and public attitudes that single-agency studies cannot adequately examine.

In addition to contributing to current research on BWCs, which I explore in Section 1.2.3, I make several contributions to existing literature. First, this paper contributes to the literature on the effects of police inputs on crime outcomes. Most of this literature focuses on measuring the impacts of hiring more police officers (Levitt, 1997; McCrary, 2007; Miller and Segal, 2019; Mello, 2019). Recent papers, however, evaluated innovations in policing, such as the use of computerization (Garicano and Heaton, 2010) and DNA databases (Doleac, 2017). By studying an input that has been introduced primarily to decrease the use of force and "side-effects" from policing, this paper differs from previous research on inputs applied to improve general crime rates and clearance rates.

Additionally, this paper contributes to the recent literature that has analyzed on police use of force as well as civilian complaints. Fryer Jr. (2018) explores racial differences in the police use of force, while Ba (2017) and Rivera and Ba (2019) study the interactions between civilian oversight and the use of force and complaints. Annan-Phan and Ba (2019) examine the effects of the patrol environment on deadly force. Rozema and Schanzenbach (2019) study intervention methods that can predict which officers will display the most civilian allegations of misconduct. This paper differs from these papers by evaluating an accountability tool that could help reduce the use of force and complaints.

More broadly, this paper contributes to a large body of literature on agency issues in the public sector. It is more difficult to incentivize government bureaucrats than private sector employees because it is hard to find an objective measure of performance and the set of contracts governments can offer their employees is limited (Finan et al., 2017). Among different types of government agencies, enforcement agencies face particularly severe challenges because of multitasking problems (Khan et al., 2016, Prendergast, 2001, Shi, 2009, Ba, 2017, Rivera and Ba, 2019, Devi and Fryer Jr., 2020). This paper provides one of the first evidence that suggests an accountability tool could bring benefits in enforcement agencies.

Finally, this paper relates to the literature that examines technology adoption, work organization, and performance (e.g., Bresnahan et al., 2002; Hubbard, 2003; Athey and Stern, 2002; Acemoglu et al., 2007; Aral et al., 2007). Although organizations have invested substantially in monitoring employees, employee-monitoring technologies have been under-studied in the literature. Such technologies have the potential to curb misconducts while invasion of privacy can also have negative consequences for productivity or retention (Bernstein, 2012). Pierce et al. (2015) finds that a theft-monitoring software significantly reduced theft and improved productivity in restaurants. I extend this literature by analyzing a video recording technology that can more comprehensively capture employee behavior but is more intrusive. My agency-level analysis allows me to study multiple dimensions of performance in addition to undesirable behavior.

The remainder of the paper proceeds as follows. In Section 2.2, I provide institutional details about BWCs and present a stylized model of police behavior and BWCs. I also describe previous studies on BWCs. Section 1.3 describes my data on BWC adoption status, the use of force, and other performance measures. Section 1.4 compares the characteristics of agencies that adopt BWCs and those that do not, in addition to laying out an empirical strategy for overcoming obstacles to the comparison of these two groups of agencies. In Section 2.4, I present my main results on the use of force and enforcement outcomes. I also explore possible mechanisms underlying reductions in the use of force. Section 1.6 analyzes changes in public attitudes toward police after BWC adoption. In Section 1.7, I perform a social welfare calculation of BWC programs and a budgetary analysis of BWC programs from the perspective of police agencies. I conclude in Section 2.5.

1.2 Background and Expected Effects of Body-Worn Cameras

1.2.1 Adoption of BWCs and Policies

Widespread public support and buy-ins from law enforcement executives and officers have materialized into policy changes across the nation. In 2014, the Obama administration proposed a subsidy of $263 million for the purchase of 50,000 BWCs by local law enforcement agencies. Backed by an infusion of these federal funds, and reinforced by grants from state and local governments, BWCs are now widely used in the US. Specifically, as of 2016, they have been fully deployed by 60% of local police departments and 49% of sheriffs' offices in the US (Hyland, 2018). Currently, debates are active nationwide about whether to equip all US law enforcement officers with BWCs.

BWCs are not a new technology, although they have evolved over time through the incorporation of recent smart technological advances. BWCs have been tested in the United Kingdom as early as 2005 (Harris, 2010). However, they did not gain widespread adoption in the US until a string of high-profile and controversial officer-involved killings in New York, Missouri, Illinois, and Ohio spurred public demand in the second half of 2014 (Maskaly et al., 2017). Figure 1.1 depicts the evolution of the coverage of BWC programs across the US. After a modest rise in BWC implementation in the US to around 15% of all agencies, BWC adoption escalated in 2014 at a steeper rate, entering 67% of all agencies by June 2016. This trend captures agencies' haste to outfit officers in order to avoid controversies. Before BWCs became the subject of public attention, agencies that adopted the technology had generally deployed the devices to specialized positions such as traffic details or officers covering major events such as protests and sporting events (Police Executive Research Forum, 2017).

From the perspective of police departments that were looking to obtain BWCs, BWCs are more than just a device. First, they are expensive. BWC programming involves rigorous logistics, vast data management systems, public records requests processing systems in addition to the cameras themselves. For example, Federal Ways, CA, that has 134 officers has allocated for its BWC program $1.1 million for the initial year and $450K per year, which represents 1.4% of the total annual budget and 3.8% of the budget for Field Operations. A survey of US police executives, which serves as one of my primary data sources, shows that the costs of BWC implementation are a real concern. 86% of executives that did not adopt BWCs said that costs were a primary reason; this was followed by the burden of public records request, which were mentioned by 70% of executives. Among agencies that adopted

BWCs, the biggest obstacle they faced in implementation was the fact that costs were higher than anticipated, as reported in Table 1.2. Because of the high expenses, agencies looking to purchase BWCs resorted to funding opportunities provided by local, state or federal governments. Second, the implementation of BWCs requires careful crafting of policies and negotiations as BWCs carry sensitive information. Officers are sensitive to materials that is recorded, stored, and released, because BWCs can intrude on their privacy and negatively affect their employment status.

As a recording device that is more mobile than previous technologies such as car dashboard cameras, the BWC has the potential to reshape police-citizen interactions. BWCs can be worn on the front of officers' uniforms or clipped to headgear. Different departments have varying policies on what events to record. In my data sample, the majority of police departments that adopted BWCs require officers to turn on BWCs during routine calls for service, traffic stops, officer-initiated citizen contacts, firearms deployments, and executions of arrests and search warrants. A common "fail-safe" feature of BWCs allows them to always be on and save the 30 seconds of footage prior to the officer activating the record button.

Once the footage is recorded, it is usually stored with security protections to address concerns of privacy and evidence integrity. Most agencies record and track internal access to video files. The video files are kept in storage for varying periods of time, with the modal response in my data being one month to a year. The videos may be kept for longer periods of time if, for example, they are associated with use of force, a citizen complaints, or legal proceedings.

The process of implementing BWCs is lengthy and contains many administrative hurdles. This paper's identification strategy relies on the bureaucratic process that has prevented agencies from maintaining full control over the timing of eventual adoption. The implementation process begins when the police chief consults with the local council and recommends capital purchase. These two entities make decisions based on the positions of the public and officers. The process takes about 18 months.* The agency must study devices and requests proposals from vendors. Because the BWC program is expensive, the agency prepares a budget according to its availability and writes grants which often contain stipulations that the award can only be used for cameras that have yet to be purchased. The agency must also develop policies that satisfy the requests of the council as well as police unions, which may prolong the process. Finally, the agency must train officers and install necessary hardwares and hire personnel to process public records requests.

In the initial years of BWCs, as agencies navigated these bureaucratic processes, anecdotal evidence abounds that agencies faced long and arbitrary delays. For example, in Nashville, TN, BWC was named a top priority in 2016 by city and police leadership, but in 2019, months before roll-out, the mayor delayed the program citing simmering budget crisis and concerns over massive downstream costs. Agencies also awaited official policies from county and state governments regarding BWC use and data release policies before implementing their programs. These bureaucratic factors will serve as the basis for my empirical design that relies on variation in BWC adoption timing.

1.2.2 Model of Police Behavior and BWCs

How do BWCs change the use of force? First, as officers recognize that their actions may be subject to closer scrutiny and that bad policing tactics are more likely to be detected and punished, they consequently attempt to reduce actions that lead to bad behavior. Thus, the total amount of policing efforts may decline, as bad policing may be proportional to the total policing efforts. This phenomenon is commonly known as de-policing and has been widely documented (Prendergast, 2001, Shi, 2009, Devi and Fryer Jr., 2020). Second, BWCs may lead officers to develop more prudent police tactics by investing in their skills. In addition,

∗. A recent Chicago Tribune article (09/11/2020) quotes the Chief of Naperville, IL, police who made the assessment based on other police departments' timelines (https://www.chicagotribune.com/suburbs/naperville-sun/ct-nvs-naperville-police-body-cameras-st-0913-20200911-bbticmklbbadzhcxjludf2qfoi-story.html)

BWCs may further facilitate training by providing objective views into policing in action. Anecdotal evidence suggests that agencies use footage from BWCs to provide scenario-based training, to evaluate the performance of new officers in the field, and to identify areas in which further training may be needed (Police Executive Research Forum, 2014). Third, citizens may be more likely to cooperate with police if they hold more positive beliefs about the police who are equipped with BWCs or are aware that they are more likely to be prosecuted for resisting arrest with video evidence.

To illustrate these concepts more clearly, I have constructed a model of police decision-making based on the principal-agent models of policing in Prendergast (2001) and Shi (2009). The objectives of the city (the principal) are to minimize crime while also minimizing errors and expenditure. Police officers and their field supervisors (agents) decide upon the amount of training and policing efforts facing the ex-post oversight. In the first stage, the agent makes the decision of improving her force mitigation tactics such as tactical positioning, verbal de-escalation techniques, and implicit bias controls. Through the principal-agent model, bad policing naturally arises as a fixed proportion of policing efforts, and investment in policing tactics reduces chances that bad policing will occur. These investments are costly as they require close examination of current practices and efforts to detect and correct any suboptimal behavior. BWCs may affect policing in the model in three different ways; they may increase the rate of investigation of bad policing, reduce the cost of skill development, or decrease hostility of citizens toward the police.

In the first stage of the model, the agent's problem can be written as

$$U_1 = \max_k w + \delta U_2(k) - ck, \tag{1.1}$$

where she maximizes the first period utility U_1 by choosing $k \in \{0, 1\}$. If she chooses to invest, she pays the cost c. She receives a fixed wage w from the principal and cares about discounted future U_2, which depends on her investment decision. In the second stage, she decides the effort level in making arrests taking into account the expected penalty from errors in arrests, rewards from arrests, and dis-utility from the effort:

$$U_2 = \max_e \{w - N(e) \times \rho_h \times q(k) \times \rho_i \times \Delta + f(N(e)) - \frac{1}{2}e^2\}, \tag{1.2}$$

where N is the number of arrests made by the officer (which increases in effort e), ρ_h is the citizen hostility parameter drawn i.i.d. from a distribution between 0 and 1 (where a higher outcome implies more investigative cases), $q(k)$ is the expected error rate which is a decreasing function of skill investment from the first period, ρ_i is the rate of investigation and finding that the officer is guilty of wrongdoing, Δ is the punishment for being found guilty. $f(N(e))$ represents the reward the principal gives to the agent for each arrest, and the final term parametrizes the cost of arrest efforts.

The amount of the use of force F is increasing in arrest efforts e, decreasing in force mitigation techniques k, and increasing in citizen hostility:

$$F = \frac{1}{A^k} e^{\alpha} \rho_h^{1-\alpha}, \qquad A > 1. \tag{1.3}$$

.

In the model, the primitive variables that are directly affected by BWCs are ρ_i, c, and ρ_h; BWC adoption increases ρ_i, and ρ_h, and decreases c. Through these changes, BWCs may change the use of force in the following three channels:

1. BWCs decrease arrest efforts as greater ρ_i implies higher costs from police errors. Lower arrest efforts lead to lower levels of use of force in Equation 1.3.

2. BWCs increase investment in policing skills as greater ρ_i incentives skill development to reduce errors and cost of investment c decreases. The agent employs a cutoff strategy when choosing to invest, in which lower costs of skill investment and higher rates of investigation lead to decision to invest in skill development. Better skills directly lead

to less use of force[†]

3. As BWCs increase police legitimacy, lower ρ_h directly reduces the use of force. The model mechanically generates this relationship in Equation 1.3.

The three channels have a few implications about law enforcement outcomes, which I will use to guide my empirical findings. The first channel of lower policing efforts imply that BWCs would also lead to lower arrests. The second channel implies increase in arrest rates, as lower error rates from higher skills incentives greater policing efforts. The third channel of lower citizen hostility implies lower assaults against officers. I use these predictions to guide my analysis when I examine mechanisms behind changes in the use of force.

Not modeled here are potentially beneficial effects of BWCs on investigation and prosecution, which may lead to reduction in cost of arrest efforts. However, since these aspects are not primary factors that determine the use of force, I do not consider them in my model.

1.2.3 Previous Studies of BWCs

The gradually increasing body of academic research of BWCs, all in the field of criminology, has lagged behind the rapid growth of interest in and adoption of the technology and has sought to uncover whether BWCs fulfill promises of greater police accountability, police efficacy, improved relations with the public (Lum et al., 2019). A significant number of those studies have used a deterrence framework to examine the impacts of BWCs on officer behavior, including use of force (Ariel, 2016; Braga et al., 2018; Jennings et al., 2015; Yokum

†. Change in the use of force from skill development can be thought of as a combination of the direct effect and the indirect effect through increased policing efforts (lower error rates lead officers to increase efforts). In a model with continuous skill development k, the change in the use of force can be expressed as the total derivative $\frac{dF}{dk}$, which equals

$$\frac{\partial F}{\partial e}\frac{\partial e}{\partial k} + \frac{\partial F}{\partial k}. \tag{1.4}$$

With the added assumption that the direct effects of mitigation skills dominates the indirect effects through increased arrest efforts, use of force declines.

et al., 2019), as well as citizen complaints about officer behavior or conduct (Hedberg et al., 2017; Peterson et al., 2018), and arrests (Ariel, 2016; Ready and Young, 2015).

Table 1.1 lists randomized controlled trials that examine the effects of BWCs on use of force listed in a recent literature review (Lum et al., 2019). The estimates of effect range from -53% to +50%, with many finding null results. Some papers that find large reductions (e.g. Ariel et al. (2015)) also find large reductions for the control group relative to the pre-intervention period, indicating that there could be large spillover effects. Overall the results are mixed.

Others studies of BWCs investigate the impacts of the technology on various aspects such as citizen behavior, generally through examination of physical responses to police actions – for example, assaults on officers (Ariel, 2016; Ariel et al., 2018) or deterrence of criminal or antisocial behaviors due to the presence of cameras (Ellis et al., 2015; Police and Crime Standards Directorate, 2007; ODS Consulting, 2011). Here, again, the results are mixed. Regarding assaults on officers, a number of studies have actually found the presence of BWCs increases violent incidents, while others find that they have a null effect (Headley et al., 2017; White et al., 2017). Taken together, then, this group of studies finds little evidence to suggest that BWCs have a civilizing effect on citizen behavior.

This growing body of research have added considerably to the overall knowledge of the effects of BWCs on various aspects of policing. Nonetheless, the previous studies have been limited by contamination issues in single-agency settings and an inability to consider agency-wide effects.

Single-agency studies are constrained in their ability to effectively deal with spillover effects, as they randomize BWCs to officers and shifts. Officers with BWCs may work, or even partner, with officers without BWCs, and in studies where BWCs are randomized by shifts, the same officers may work with or without BWCs depending on the day. Thus, researchers significantly lose their ability to assign treatment status in an unbiased way, and consequently, their ability to uncover the true effect of BWCs. Furthermore, officers in the control group can learn from officers with BWCs through peer effects and adjust their behavior accordingly. By using agencies as the unit of observations in this study, I am able to overcome this contamination issue.

Relatedly, the agency-level framework in this paper is able to capture agency-wide effects from BWCs. Officers, potential victims, potential offenders, and other civilians are likely to adjust their behaviors over time as they become accustomed to BWCs. The actions and presence of BWCs-equipped officers have implications beyond their own activities and eventually affect all the officers in the agency. Previous studies are likely to have neglected the agency-wide effects that affect both the control and treatment groups. Additionally, directly assessing BWC effects on crime rates has not been possible for previous single-agency studies. It is important to measure their effects on crime, as this allows us to understand whether there have been trade-offs from reduction in use of force.

Lastly, single-agency studies suffer from selection bias, as the researchers can only perform research on agencies that are willing to cooperate and share data (Allcott, 2015). These agencies are less likely engage in problematic practices and culture, which may lead to findings of a null effect. On the other hand, studies of these agencies may display bias toward a greater effect if BWCs truly can fix problems with more cooperative officers.

1.3 Data

1.3.1 Data on BWC Adoption

This study uses the differential adoption timing of law enforcement agencies to quantify the effects of BWCs. I gather agencies' adoption decisions from the Law Enforcement Management and Administrative Statistics (LEMAS) survey which has been administered by the Bureau of Justice Statistics (BJS) every three years since 1987. Garicano and Heaton (2010) uses the LEMAS survey to construct a panel dataset of police departments and examines the effects of information technology on law enforcement productivity. Similarly, I employ the Body-Worn Camera Supplement (LEMAS-BWCS) which was administered for the first time in 2016. As it was released recently in 2019, to the best of my knowledge, this data has not yet been used in academic research. This survey was distributed to all agencies with over 100 officers as well as a nationally representative sample of smaller law enforcement agencies in the US. This data includes about a quarter of the total agencies extant in 2016. The LEMAS-BWCS contains responses from heads of agencies on topics that range from the current status of BWC use and reasons for adoption to obstacles they faced in BWC implementation. Most importantly for the purposes of this research, it contains agencies' answers as to when (year and month) they adopted BWCs.

The decision to adopt BWCs is not binary. Law enforcement agencies may choose to adopt this technology on varying scales. Additionally, adoption may begin with a small-scale pilot and grow into a larger deployment later. One limitation presented by the dataset is that I cannot use it to observe adoption initial adoption scale, and only the scale at the time of the survey (June 2016). Agencies that indicate they adopted BWCs include those that deployed the technology to only a small part of the force. Agencies that adopted BWCs at low rates may not experience meaningful impacts of their implementation. To address this possibility, I examine the effects of BWC implementation at agencies that adopt the technology at rates above a certain threshold. I do not claim to know what the ideal threshold is. Additionally, this concern must be weighted against the need for a large enough sample. For the nationwide study where I have a large enough sample, I run estimations with thresholds of 30%, 40%, and 50% in terms of the the proportion of officers wearing cameras in 2016 as well as estimates using all adopters. The results are qualitatively similar. For conciseness I report my findings using agencies that adopt BWCs at a rate of over 40%, and present the others in table form.

For my analysis in New Jersey, I expand the LEMAS data by submitting FOIA requests to all New Jersey agencies, as the LEMAS data capture adoption status for less than 30 NJ adopters. The FOIA requests provide me with an adoption status of 60% for all NJ agencies, of which 100 are adopters.

1.3.2 National Data on Police-involved Homicides

There are currently two government-funded datasets on police-involved homicides, neither of which are reliable: the voluntary justifiable homicides collected in the Supplemental Homicide Reports (SHR) under the UCR system of the FBI and the Department of Justice's arrest-related-death (ARD) dataset. Researchers have long pointed out that these datasets contain substantial omissions in these data (Klinger, 2012; Schwartz and Jahn, 2020; Edwards et al., 2019). Currently only 750 of approximately 17,985 agencies voluntarily submit justifiable homicides (as defined by agencies) to the SHR, and the Bureau of Justice Statistics reports that the ARD misses between 30-50% of deaths (Banks et al., 2015; Finch et al., 2019). The BJS claims that data from the ARD does not meet BJS quality standards and suspended data collection in 2014. The ARD does not overlap with the time period studied in this paper. Although the SHR covers the time period, I do not use this data, as it suffers from substantial omission and is likely to under-report cases that are less justifiable in the eyes of the public.

In the wake of controversial use-of-force incidents, journalists and independent researchers have worked to fill these gaps in data needs. They collected media mentions and crowd-sourced cases and cross-examined them using social media, obituaries, criminal records databases, and police reports. I use Mapping Police Violence (MPV), which is compiled from the three largest and most comprehensive crowd-sourced databases and is currently, to my knowledge, the most comprehensive accounting of police-involved homicides. MPV began in 2013 and contains incident-level information on cases in which civilians die as a result of being intentionally harmed from police officers. The data fields include agency responsible for the death, circumstances surrounding the death, and demographics of the victim. I aggregate the data at the agency-month level and merge with the other data.

As a crowd-sourced dataset, MPV likely underreports the true extent of police-involved homicides. However, it is the best data that we have on this important issue, and according to at least one measure, this data seems reliable; in 2015, the BJS estimated around 1,200 police-involved homicides occurred, and the MPV identified 1,106 cases during this period (Banks et al., 2016).

One of MPV's data sources, Fatal Encounters (FE), extends even further back to 2000. However, the quality of this data source deteriorates as one moves back away from 2013, as the data on incidents before 2013 was gathered retroactively. The administrator of the MPV has suggested that data before 2013 is not reliable. Data gathering and cross-examination of cases is difficult for historical events, and crowdsourced inputs are most likely to be reliable for current events. Additionally, the fervent national attention on use of force that began after 2013 did not exist before 2013, and the media was less likely to report police-involved homicides. In Appendix Figure 1.A.1, FE shows a marked ascent up until 2013. This increase is consistent with the data problems I discussed above. For these reasons, my analysis on police-involved homicides begins in 2013, when more reliable data becomes available.

I describe the time series of my data on police-involved homicides in Figure 1.2. Police-involved homicides (population-weighted mean) decline from the period between 2013 and 2019 for the agencies that are matched to the LEMAS data.[‡] On the right panel, I separately examine the differential trends of adopters and non-adopters. We first see that the adopters and non-adopters are different at the baseline. In the data, the adopters display roughly twice as many cases of police-involved homicides as the non-adopters in 2013. The gap diminishes

‡. The national trend including agencies that did not match with the LEMAS data is quite flat and stable. An up-to-date plot can be accessed at https://mappingpoliceviolence.org/nationaltrends. The discrepancy exists because the LEMAS survey oversample larger agencies with more than 100 officers

as the adopters experience a decline. This figure uncovers the interesting relationship between BWC adoption and how police-involved homicides have evolved which I will investigate in a more systematic way to examine if BWC adoption is indeed responsible for the decline in police-involved homicides.

1.3.3 Use of Force Data in New Jersey

For my analysis of use of force in New Jersey, I rely on the Force Report collected and maintained by NJ Advance Media. NJ Advance Media, the leading media company in New Jersey, has collected and digitized all use-of-force forms from 2012 to 2016 from all local law enforcement agencies in New Jersey.[§] The database contains incident-level information on levels of force, date, agency and officer names, subject demographics, and injuries.

As the only existing set of data on the use of force that contains the full geographical coverage of a state, this data provides a breakthrough in research on the use of force, which previously could only be obtained from a handful of agencies. When the data was released at the end of 2018, the New Jersey Attorney General praised the work as “nothing short of incredible” and “something they should be doing.”[¶] The Force Report is less likely to be affected by site-selection bias which is introduced when researchers attempt to acquire sensitive data about subjects such as use of force from organizations.

On the other hand, the data still presents a couple of limitations. The data is constructed from reported that are created by officers and hence are thus likely to fail to capture the full extent of use of force. Any use of force beyond physical force (e.g. compliance hold, use of hands, and fists) must be reported in New Jersey as in most agencies in the US. When citizen complaints initiate investigations or lawsuits against officers, these use-of-force reports can

§. In 2000 NJ Attorney General required all agencies to report use of force incidents. However, this tracking and reporting system did not work properly. NJ Advance Media performed a 16-month investigation and gathered 72,677 use of force reports through FOIA

¶. https://www.nj.com/news/2018/11/njcom-probe-of-police-force-nothing-short-of-incredible-njs-top-cop-says-now-hes-promising-major-reform.html provide critical evidence. During the sample period of study, officers were incentivized to under-report use of force, as society heavily scrutinized use-of-force incidents. For low-level use of force, without trails of evidence, officers have incentive to neglect reporting to avoid scrutiny. For more severe uses of force that leave behind evidence, officers are less likely to neglect to report. Not only do reporting requirements become more stringent as severity increases, but victim injuries also make it hard to hide use-of-force incidents. Both the supervisors and officers have incentives to clearly document cases in the event that citizen complaints arise. BWCs tend to decrease reporting bias. They allow less room to hide use of force and increase the likelihood of citizen complaints about excessive force, as they provide citizens with objective evidence. Thus, findings that BWCs reduce use of force are likely to be understating their true effects. Along with use of force I also examine subject injuries to alleviate this concern.

Focusing on New Jersey gives us a window into overall police use of force, which is more frequent than rare police-involved homicides, though this finding comes at the cost of reduced generalizability. New Jersey has relatively been spared from controversial use-of-force incidents and is an affluent state (9th richest in the US). Its law enforcement agencies may thus be better maintained than the national average and have better systems of accountability compared to those of other states.

1.3.4 Data on Other Law Enforcement Outcomes

Data on incidents and crime control activities come from the Uniform Crime Reporting (UCR) database maintained by the Department of Justice. This database contains agency-wide information on monthly index crimes, arrests, and officer safety incidents. It also contains information on the number of officers and citizens the agency serves for each year. Previous literature using this data notes the importance of cleaning this data, as agencies have wide discretion as to their methods of reporting. I clean the data in ways similar to those used by other researchers. The specific procedures used are described in the detail in the Appendix.

I merge BWC surveys with the monthly data of UCR so that my final sample only includes local police departments that were recruited for and participated in the LEMAS-BWCS survey. In total, there are 2,380 departments, with 1,001 adopting agencies in the main sample. For my study on the use of force in New Jersey, I also merge in monthly aggregates of the use of force in New Jersey police departments.

The merged data runs from January, 2013, when the MPV starts, to June, 2016, when LEMAS was collected. This time period allows me to focus on agencies that adopted beginning in 2014, when agencies began to adopt BWCs on a large scale due to public interest.

I bring in a few other data for my analysis on public sentiment. I introduce these data later in the relevant section.

1.4 Empirical Strategy

My primary empirical approach exploits staggered adoption of BWCs by law enforcement agencies using time variations within the sample of adopters. The availability of data on the exact adoption dates of BWCs and high-frequency monthly data on police activities further helps me finely tease out the effects of treatment effects of BWC adoption from alternative explanations.

The main goal of this paper is to convincingly attribute changes in policing variables of interest to the adoption of BWCs by different agencies. BWC adoption is not random and is related to variables potentially linked to the outcomes of interest. Table 1.3 compares the characteristics of departments by adoption status. The adopting agencies are those that adopted BWCs in 2014 or later. The data on these agencies come from the five-year estimates of Census places and county subdivisions in the 2013 American Community Survey. The adopting agencies were more likely to be located in cities with higher populations, higher minority populations, and lower incomes. Mirroring the differences in population sizes, the adopting agencies were also much larger.

Although there are significant differences between adopting and non-adopting agencies, I argue that the adopting adopters demonstrate relatively limited control over the exact timing of BWC adoption. As a result of high-profile use-of-force incidents on national media and ensuing debates, law enforcement agencies rushed to adopt BWCs. At the same time, as shown in Table 1.2, agencies faced logistical difficulties in implementing BWC programs. Especially in early years of BWCs around 2014, agencies that wanted to adopt BWCs experienced long and arbitrary bureaucratic delay in deployment, leading to arguably exogenous variations in the timing of adoption among adopters for reasons unrelated to time-varying conditions such as police tactics and crime control as discussed in Section 1.2.1.

A comparison between early and late adopters supports this claim. In my empirical analysis, I define early adopters as those that adopted BWCs between 2014 and June of 2015 to provide one year of post-event window in my event study; late adopters are those that adopted BWCs after June of 2015. The differences between these two groups are smaller than the differences between non-adopters and adopters. I do not reject that the two groups of adopters are different.

Table 1.4 also supports this claim. It examines the differences in a regression framework and tests whether the timing of adoption can be predicted using pre-determined covariates of the agencies. In the first column, the decision to adopt is predicted by multiple covariates including race, income, and poverty. However, in the second column, all of the variables fail to predict the specific timing of adoption. For agencies that adopt on a nontrivial scale – for example those adopting at rates of over 30% and 50% – we see that all key variables are also statistically insignificant.

Motivated by these empirical findings, I follow the approach of Deshpande and Li (2019) and compare the evolution of outcome variables of early adopters to those of later adopters.

For the primary national analysis, the data ranges from January 2013, when the MPV data starts, to June 2016, when the LEMAS survey was collected. To conduct a balanced event study one year before and one year after adoption, I analyze agencies adopting between January 2014 and June 2015. I compare these agencies with later-adopting control agencies that adopt BWCs after one year or later. The later-adopting agencies are assigned the adoption date of the respective treatment agencies and compared across the event times of the treatment agencies.

Specifically I construct my sample as follows. I create separate datasets for every month between January 2014 and June 2015. In each dataset, agencies that experience the current adoption are labeled the treated agencies while agencies that adopt after over a year are labeled the control agencies. Event months are specified as relative to the month of current adoption timing. This procedure allows me to study the period one year before and one year after the adoption of BWCs. To estimate the effects of BWC adoption in regression form, I estimate the following regressions for agency j and month t:

$$\begin{aligned} Outcome_{jt} = \gamma Treated_j + \sum_{\tau=-11}^{11} D_{jt}^{\tau} + \sum_{\tau=-11}^{11} \beta_{\tau}(Treated_j \times D_{jt}^{\tau}) + \phi_j \\ + \delta_t + \phi_j \times f(time)_t + \epsilon_{jt}, \quad (1.5) \end{aligned}$$

where D_{jt}^{τ} is a dummy variable that indicates whether the agency adopts BWCs in τ months. $Treated_j$ is an indicator variable equal to 1 if agency j is a treated agency that adopts BWCs early. ϕ_j and δ_t denote agency and calendar-month fixed effects. The main coefficients of interest β_{τ} estimate the divergence in outcome variables net of changes in control agencies after adjusting for covariates and secular trends. $\phi_j \times f(time)_t$ are agency-specific time trends using a linear function of time. I weight all my regressions by the population count the agency serves.

To summarize the results, I pool the estimates in the pre- and post- event in the following fashion:

$$Outcome_{jt} = \gamma Treated_j + \sum_{\tau=-11}^{11} D_{jt}^{\tau} + \sum_{\tau=-11}^{11} \beta(Treated_j \times Post_{jt}) + \phi_j$$
$$+ \delta_t + \phi_j \times f(time)_t + \epsilon_{jt}, \quad (1.6)$$

where I replace the event time variables with a single indicator variable $Post_{jt}$ that takes a value of 1 if month t is after adoption.

The main outcome variables are police-involved homicides, overall use of force, and injuries, as well as other crime outcomes such as arrests and index crime rates. In addition to the overall use of force, I further examine the use of force separately by three different types of force based on New Jersey Attorney General's Use of Force Policy: physical, mechanical, and deadly force. Physical force includes various methods of hand-to-hand confrontation such as wrestling a resisting subject to the ground, wrist locks or arm locks, and striking with the hands or feet. Mechanical force involves the use of some kind of device or substance, other than a firearm, such as a baton, canine physical contact, chemical spray, or more enhanced methods of conducted energy devices. Deadly force denotes the use of firearms intended to cause death or serious bodily harm.

Relative to a more traditional difference-in-differences (DID) design, this stacked difference-in-differences strategy uses a control group to eliminate event time trends that do not appear in calendar time. When agencies make the decision to adopt BWCs, they may account for trends in the use of force and policing activities as a criterion. Event time effects help to control for such consideration which would not be suitably controlled for in an event study design that only has calendar-time effects. Additionally, this stacked difference-in-difference design circumvents problems with DID models that are recently uncovered (e.g. De Chaisemartin and D'Haultfoeuille, 2018; Goodman-Bacon, 2019, Sun and Abraham, 2020, Callaway and Sant'Anna, 2020) by avoiding already-treated units in the control group.

The main assumption of my empirical approach is that after accounting for the covariates, the agencies that adopt BWCs at one point are comparable to agencies that adopt at different points of time. As a test of the identifying assumption, I will be careful to observe any pre-trends preceding the adoption of BWCs, or $\tau < 0$.

1.5 Results on the Use of Force and Enforcement

1.5.1 Effects on Police-Involved Homicides

Figure 1.3 shows the results of estimating Equation 1.5. I present results for different adoption rate cutoffs: 0%, 30%, 40%, and 50%. I plot the event time estimates β_{τ}. For the agencies with all adoption rates, the estimates for the pre-event period are slightly unstable, however, after the adoption, police-involved homicides drop. The patterns are clearer for higher cutoffs. Police-involved homicides are reasonably flat and close to zero. After adoption, they decline and persist for the event study window.

Table 1.5 presents summary estimates in which I restrict the event estimates after the event to one *Post* coefficient. All the estimates are negative at different adoption thresholds and obtain statistical significances of 5% for all the thresholds examined. The estimates indicates that BWC adoption leads to a decrease in police-involved homicides of about 0.02 per 100,000 citizens. To put this in perspective, the pre-adoption mean of police-involved homicides is 0.035. This implies a drop of 58%.

In Appendix Figure 1.A.2, I also present results using an event study framework that includes non-adopters for robustness. Prior to adoption, there exists an upward trending pre-trend. However, the evolution takes on a similar shape to Figure 1.3 in that it displays a break in trend after adoption. Appendix Figure 1.A.3 presents results without agency-specific trends. We see that before BWC adoption, police-involved homicides trend slightly upward trend, but they drop immediately after adoption.

How have BWCs affected different types of agencies? I explore heterogeneity of the effects in Table 1.6. Agencies that are in more urban and racially diverse areas have experienced larger declines in police-involved homicides. I also check examine whether agencies that have more requirements of turning on BWCs had different changes. As officers have less latitude over the decision of turning BWCs on and off, they may experience more pressure to change their behavior or may be more likely to detect possible blind spots in their police tactics. I find that agencies that have more than median level of requirements have experienced greater declines. Finally, the last column shows that agencies that have adopted at higher rates of adoption have experienced larger declines in police-involved homicides.

1.5.2 Alternative Explanations and Robustness

Testing other explanations

In the visual evidence, the sharp trend break immediately after BWC adoption helps me rule out alternative explanations based on other potential reforms such as training and changes to use-of-force policies. For other such reforms to explain the visual patterns we observe, their implementation schedules would have to overlap with that of BWCs for all adopting agencies. This is unlikely because implementing BWCs takes 18 months on average and has unique bureaucratic and logistical processes not involved in other reforms.

Nonetheless, I use couple approaches to check if the reduction in the use of force is driven by other concurrent reforms. The LEMAS data contains a question asking police chiefs what were primary motivation behind BWC adoption. I remove agencies that answered that they have adopted BWCs for "reform-driven" reasons including for reducing the use of force or enhancing police accountability. For remaining agencies, who have adopted BWCs for other reasons such as making cases more prosecutable, I run my main specification. Although this exercises removes more than half of the sample, Table 1.12 shows that even those agencies also experienced large drops in police-involved homicides.

To more directly disentangle BWC adoption from other interventions, I obtain historical data on purchase orders of government purchases from a commercial vendor of government business intelligence.[‖] Around half of my sample (455) agencies are observed in this data. With this data, I should be able to indirectly test the hypothesis that BWC adoption happened concurrently with other major purchases that are related to police accountability. As police training is most likely to be combined with purchases of BWCs, I gather all purchase orders that contain the word "police" and "training." If police department had bundled BWCs into other projects, we would see prior to BWC adoption an increase in purchase orders of both these projects and BWCs.

The purchase orders data cover the sample period between January 2013 and June 2016. For 452 agencies, I identify 3,107 purchase orders that are related to police training and BWCs, which translates to an average of 7 per agency. In my data, I observe that 34% of the sample are observed to have submitted purchase orders for BWCs. Other agencies are not observed likely because they received BWCs from grants or other agencies at the county-level or state-level, aside from data errors.

I observe the expected patterns of purchase orders of BWCs before adoption. In Figure 1.5, we see a sudden jump in purchase orders in the month the adoption. Purchase orders of BWCs return to their previous level. On the right hand side of Figure 1.5, I observe patterns of purchase orders that are related to police training. The trends are relatively well-behaved before BWC adoption, and I do not see spikes concurrent with purchase orders for BWCs. This suggests that BWC adoption were not part of broader reform packages, which strengthens my claim that the timing of BWC adoption was more or less random and isolated from other changes.

Another alternative explanation I consider is that a use-of-force incident can create in-

‖. I obtained data from GovSpend, which has collected purchase orders for a large set of US agencies in various sectors. A previous draft of this paper has included an analysis using requests for bids. However, this approach yielded very few matches agencies use formal bids for only small cases of purchases tense social pressures to adopt BWCs and reduce the use of force in the same month. Although adoption in such a short time frame is unlikely, there may be extraordinary events that forces police chiefs to shorten the implementation times. To address this issue, I exclude agencies that experience both BWC adoption and Black Lives Matter protests in three three months leading up to BWC adoption; Table 1.12 demonstrates that the results are largely similar. Later in my analysis of public opinion, I also show that there is no spike of negative public opinion near BWC adoption.

Additionally, one may be concerned that some agencies may have the logistical capabilities to implement BWC programs quickly as a response to a use-of-force incident. To address this issue, I remove agencies that did not indicate logistical obstacles during BWC implementation in LEMAS. The results do not change much.

Other specifications

Furthermore, I check my main result by switching my main data from MPV to FE. The results, as presented in Table 1.12 are similar. An alternative way to construct the outcome variable is to look at the use of force per number of officers instead of per number of citizens. From the pre-adoption mean of 21 occurrences per 100,000 officers, police-involved homicides drop by 11.87, which yields a magnitude in the same ballpark as my initial findings. Additionally, because the outcome variable is rare and has many zeros, I run my main analysis using the Poisson pseudo-maximum likelihood (PPML) estimator. As noted in Santos Silva and Tenreyro (2006), PPML is attractive when dealing with cases with many zeros, such as that of police-involved homicides. As a pseudo-maximum likelihood estimator, PPML does not require the data to follow a Poisson distribution. One drawback of PPML in my methodology is that PPML displays convergence issues when the model has fixed effects for units with all 0 values Santos Silva and Tenreyro (2010). Dealing with this issue significantly reduces the sample size. In Table 1.12, the first column indicates that log of police-involved homicides decreases by 0.627, which translates to a 47% decrease from the mean. In subsequent cutoffs, the estimates are all negative in a similar range, although they lose statistical significance.

Finally, recent literature in the method of difference-in-differences (DID) has pointed out potential problems in DID with staggered treatment timings (Goodman-Bacon, 2019; Callaway and Sant'Anna, 2020; Borusyak and Jaravel, 2018; Sun and Abraham, 2020; De Chaisemartin and D'Haultfoeuille, 2018). The problem arises because the DID estimator is a weighted average of all 2-by-2 difference-in-differences in the data and introduces problematic comparisons already-treated units are used as control group. The main empirical analysis used in this paper circumvents this issue by reconstructing datasets. As a robustness check, I follow another corrective method used in the literature following Callaway and Sant'Anna (2020). This procedure involves estimating DID for each adoption month cohort with control groups that have not adopted BWCs, and taking a weighted average of the estimates across the cohorts with weights equal to each cohort's share. I present the results in Figure 1.A.4. The visual evidence looks similar to the main DID results.

1.5.3 Impact on Police-involved Homicides by Race

One question this study hopes to answer is whether BWCs affect the fatality rates of minority citizens versus white citizens differently. The protests against police brutality have been spurred by controversial uses of force against minority subjects. In Figure 1.4, I break down citizen deaths by racial groups. For the sake of conciseness, I present visual evidence for agencies that adopt BWCs at rates of over 40% and summarize the results of cutoffs in table form. I run the difference-in-difference regression for non-Hispanic whites and minorities. Minorities include blacks and Hispanics. The estimates are more imprecise than those of overall police-involved homicides. The pre-adoption means are 0.043 for whites, 0.041 for minorities, 0.035 for Hispanics, and 0.07 for black individuals (the low number for minorities seems to stem from misclassification of Hispanics in the data). For the white group, police-involved homicides drop by around 0.02 after adoption. In a more clear visual evidence, minorities experience a drop in homicides of about 0.037 after adoption. For black individuals, there is a similar drop after adoption. A drop of 0.054 for blacks represents a 77% decrease from the pre-adoption mean, although this does not obtain statistical significance. I do not separately examine Hispanics because this ethnic group does not seem to be well-coded in MPV. I summarize these results in Table 1.7.

1.5.4 Did Agencies Sacrifice Policing Capabilities?

The evidence presented thus far strongly indicates that BWCs placed downward pressure on the use of force. Insofar as BWCs reduced unnecessary or excessive use of force and encouraged subjects to comply more, this suggests that BWCs can be successfully used in agencies' efforts to avoid controversial uses of force, reduce liability, and restore public faith in policing. However, this study also intends to explore the question of whether this improvement comes at the expense of policing capabilities. For example, as discussed in Section 1.2.2, perhaps officers reduced policing efforts in order to reduce the burden of being monitored.

A direct measure of policing efforts is arrests per capita. Patterns seem to indicate that police did not typically reduce enforcement following the adoption of BWCs. Arrests actually seem to increase after adoption, as seen in the clear break in Figure 1.6. In Table 1.9, the increase of 0.043 represents a 6% increase from the pre-adoption mean, which is statistically significant at 10%.

Next, I test the effects of BWCs on crime. Figure 1.6 shows trends surrounding BWC adoption. The trends show more fluctuations than in previous figures, but there is no marked increase in crime rates. Table 1.9 summarizes the results for crime. Regarding homicide rates, which are less likely to suffer from reporting bias, I find no significant difference after adoption.

These results indicate that officers did not reduce crime control behaviors in a meaningful way (in fact, the point estimate suggest the opposite), but they did reduce the use of force on civilians.

1.5.5 Mechanisms Behind the Reduction in the Use of Force

I consider two other mechanisms from Section 1.2.2 in turn. First, did officers change their policing tactics to mitigate the use of force? In the model, this can arise because officers face higher costs of errors from oversight and they face lower costs of skill development. One strategy I use to test the validity of this mechanism is through the LEMAS survey of police chiefs. LEMAS asks police chiefs who have adopted BWCs about the ways in which BWCs have been useful. 61% of the chiefs agreed that BWCs improved the professionalism of the officers, 39% agreed that BWCs helped identify instances of officer misconduct, and 72% agreed that BWCs were useful for supervising officers.

To examine if improved police tactics have played a role in reducing the use of force, I examine whether agencies that claimed they improved upon these aspects experienced larger reductions than those that did not. In Table 1.10, I present the results. In the first column, I split the early adopters into agencies that agree that BWCs improve professionalism of officers and those that do not. The coefficients are the divergence from the control group after adoption for each category. Although the coefficients are not statistically different, the main effects are driven by those that agreed that BWCs have helped change officer behavior. There is a similar pattern for agencies that agreed that BWCs helped identify officer misconduct. On the other hand, there is a smaller difference for officer supervision, indicating that decentralized changes in behavior, rather than formal training, may have played a role in reducing the use of force. I also present visual evidence for these differences in Appendix Figure 1.A.6, which show the contrasts.

Another possible explanation for the observed reductions in the use of force is citizen behavior. Uses of force occur during interactions between officers and citizens. If officers use force to have effective arrests, they are likely to use less force if they expect citizens to be less hostile and cooperate more. On the other hand, citizens are likely to be less hostile if they know they are being recorded and any hostility or resistance could be more easily prosecuted. I do not have direct information about the circumstances surrounding the use of force. Instead, I use an indirect method of examining assaults on police. If the results are driven by the fact that citizens become less hostile, we would see a drop in assaults on police as this is mainly driven by subjects. Figure 1.6 demonstrates that although the estimates are noisy, we do not see a drop in felonies against officers like ones we see in police-involved homicides. This evidence is consistent with the hypothesis that officers drive reductions in the use of force.

1.5.6 Long-Term Effects on Police-involved Homicides

The DID comparison between early and late adopters is attractive because it provides a close comparison between similar agencies and it effectively addresses issues with staggered DID discussed in the recent literature (e.g., Abraham and Sun, 2018; Goodman-Bacon, 2019). However, its empirical design limits us to a short post-event period. Long-run effects may be smaller or larger in magnitude than short-run effects. If officers get used to BWCs over time and revert to their old habits, then long-run effects would be attenuated. On the other hand, if officers have to learn how to integrate BWCs to their policing or supervisors implement further training in response to BWC footages, then we might see greater long-run effects.

To examine the long-term impacts of BWCs, I deviate from the main empirical design and use a conventional DID that includes all adopters and non-adopters. The caveat here is that agencies that did not adopt BWCs by the time of the LEMAS survey (June 2016) might have adopted later, leading to an underestimation of BWC effects. To mitigate this concern, I exclude agencies that replied that they are very likely to adopt BWCs within one year of the survey. In addition to 502 agencies that adopted BWCs, the event study contains 1,379 non-adopters. The large group of non-adopters reduces the weight on the problematic DID comparison that use already-treated units as controls. Figure **??** shows the results on an event study that examines trends one year before and three and half years after adoption. In the first panel, I use the same set of adopters as in the short-run analysis as well as non-adopters. As expected, the model does not capture the pre-trends well, indicating differences between adopters and non-adopters. However, we see that police-involved homicides drop from the pre-trends. The drop in police-involved homicides that occurred immediately after adoption can also be seen up to three and half years after adoption. In the second panel, the sample includes agencies that adopted BWCs between June 2015 and June 2016, which I have excluded in my short-run analysis of adopters and no longer serve as my control group in the long-run analysis.

1.5.7 Effects on Low Levels of Force in New Jersey

We have examined how BWCs lead to changes in police-involved homicides; how about lower levels of force such as the use of force involving fists and feet? Lower levels of force account for the majority of uses of force and play a significant role in how society perceives the police. To examine how BWCs changed the use of force, I examine the BWC experience in New Jersey. The New Jersey sample allows me to analyze use-of-force incidents, which occurred more frequently than police-involved homicides. The event study consists of 96 BWC-adopting agencies, of which 21 are in the early adopting treatment group. Because of the small sample size that makes the monthly estimates noisy, I present quarterly estimates instead of monthly estimates (See Appendix Figure 1.A.5 for the monthly event study).

For my study of NJ agencies, in which the outcome variables begin in 2012 and end in 2016, I can extend the pre-period to 2 years, giving me 8 quarters pre-period and 4 quarters post-period. Additionally, because I have adoption dates for NJ agencies that adopted BWCs after 2016, I set the treatment agencies as those that adopted between January 2014 and December 2015.

In Figure 1.8, I present the results for all use of force. The pre-trends for all use of force are relatively stable before adoption, indicating that the treatment group agencies behaved in a similar way as the control later-adopting agencies. After the adoption of BWCs, the use of force declines sharply. The decline corresponds to 20% of the pre-adoption mean of 3.61. Another measure of the use of force is subject injury, which is less likely to suffer from measurement bias. Subject injury also exhibits a declining trend after adoption, with a decrease of 42%. Table 1.8 summarizes the results by restricting the estimates in the post period into one coefficient. When I break down all use of force by levels, the reductions in the use of force are mostly driven by physical force, such as the use of hands, fist and feet (Appendix Figure 1.A.5).

This analysis indicates that BWCs played a role in also reducing incidents of lower-level uses of force in New Jersey. With a caveat that the analysis in New Jersey may represent unique experiences of the state, we have evidence that body cameras have the potential to reduce broader uses of force in addition to police-involved homicides.

1.6 How do BWCs Affect Public Sentiment toward the Police?

In this section, I examine whether the implementation of BWCs changed public attitudes about the police. In light of the recent social upheaval and the erosion of public trust in the police, understanding whether BWCs can ameliorate public attitudes about the police is important.

It is difficult to measure changes in public opinion due to BWCs using conventional data sources such as surveys, because they do not capture high-frequency variation over wide geographical regions. Consequently, I rely on two methods. First, I measure sentiment regarding the police as posted on Twitter. Twitter's focus on current events allows me to capture sentiment that concurrently arises from BWC adoption. Additionally, the need to be concise on Twitter renders its postings amenable to text-based sentiment analysis.

For the sample period of my analysis (January 2013 to June 2016) I obtained Twitter postings that feature the words "police" or "cop" as well as the name of a city included in my sample. The main challenge of associating a tweet with an agency is that tweets in general are not geo-coded. Hence, I reduced my sample of adopting agencies from 1,001 to 594 by limiting it to agencies with uncommon names.** Even after this restriction, searches for some agencies did not generate enough Tweets. Thus, I restricted my sample to 312 agencies that have tweets from every month of the sample period and a monthly average of at least 30 Tweets on average. The sample includes about 2.4 million tweets attributed to these agencies. As expected, these agencies are large and located in urban areas.

I construct the following two measures to study changes in sentiment on social media: First is the total amount of police-related tweets:

$$TotalTweets_{jt} = \frac{\#Tweet_{jt}}{\#\overline{Tweet_j}}, \tag{1.7}$$

where $\#Tweet_{jt}$ is the total number of Tweets for agency j in month t, $\#\overline{Tweet_j}$ is the number of agency-specific average monthly tweets. This measure is inspired by the conjecture that police services are more conspicuous when they deviate from the norm which in the US generally is effective.

The second measure is average sentiment captured in tweets. I apply a widely-used lexicon-based natural language processing (TextBlob) to assign sentiments to these tweets. TextBlob assigns a sentiment score from the range -1 (negative) to 1 (positive). The score is obtained by taking a weighted average of all words in the text while taking into account

**. To be in this reduced sample, the agency needs to be in a city that is comprised of at least 80% of the combined population from cities that have the same names word ordering, negation, and intensifiers. A drawback of using sentiment as a measure is that because police-related Tweets inherently contain words associated with negative sentiment, such as "murder" and "kill," this measurement error may attenuate the estimated effects.

I use the main empirical framework of Equation 1.6 to measure Twitter sentiments. In Figure 1.9, I plot the difference-in-difference estimates. One may observe that after BWC adoption, total tweets decrease, indicating less complaints about the police. The drop in 1.3 translates to a 21% drop from the pre-event mean. I summarize the difference-in-difference estimates in Table 1.11, which shows statistical significance across all thresholds. Figure 1.9 demonstrates that sentiments, as expected, turns positive after BWC adoption, although the estimates are noisy. When I estimate the summarized form in Table 1.11, the increase in sentiment is shown to be large. An increase of 0.005 indicates an improvement of 50% in sentiment. However, the coefficients for sentiment are not statistically significant, reflecting much noise in the measure.

Overall, BWCs lead to improvement in public perceptions of the police. A limitation of this empirical exercise is that I am not able to tease out answers as to whether the change in citizen perceptions of BWCs led to decreases in the use of force or is a byproduct of a reduced use of force induced by BWCs. I have provided sufficient evidence, however, to show that BWCs lead to improved public attitudes toward the police, and to the extent that they independently induce improved citizen-police interaction in the short and long term, the improved public attitudes have the potential to further decrease the use of force.

1.7 How do BWCs Affect Social Welfare and Police Budget?

In this section, I perform a social welfare analysis of BWC adoption. A reasonable cost estimate is $1,100 per camera per year, based on a calculation of the cost of implementing a BWC program in the Las Vegas Metropolitan Police Department (Braga et al., 2017). This calculation accounts for amortized equipments, the IT infrastructure, training, and the labor costs of responding to freedom of information requests. For the average agency that adopts BWCs in my data, there are 184 officers for 78,000 citizens, so a complete BWC program costs $202,400 per year.

In terms of savings, my estimates indicate that BWCs can save 0.02 per 100,000 capita lives in each month. Using a value of statistical life of $7 million, for a population of 78,000, the yearly savings of life translate to $1,310,400. This far outweighs the cost of a BWC program.

Omitted from these calculations are aspects that are no less important but more difficult to quantify: a reduction in civilian injuries and their implications for the well-being of civilians, improvement in trust in the police, and more effective prosecutorial evidence.

Although improvement in social welfare is itself a valid reason to adopt BWCs, I also perform a cost-benefit analysis of implementing BWCs from the perspective of law enforcement agencies. I compare the costs of BWC programs with the potential gains from reduced lawsuits and settlements related to excessive policing. Though data on lawsuits and settlements by police departments are not readily available, I use large US cities as a guide. From 2010 to 2014, the 10 largest US cities each spent $204 million annually on misconduct cases.[††] I assume that lawsuits and settlements decrease by the same percentage (41%) as the decreases in police-involved homicides calculated in this study. This may be an underestimate if BWCs decrease excessive use of force more than justified use of force. A 41% reduction in yearly settlement payouts is equivalent to savings of $83 million. On the other side of the equation, based on a calculation of the cost of implementing a BWC program in the Las Vegas Metropolitan Police Department, a reasonable cost estimate is $1,100 per camera per year (Braga et al., 2017). This calculation accounts for amortized equipments, the IT infrastructure, training, and the labor costs of responding to freedom of information requests. For an average police force of 8,415 (characteristic of the largest largest depart-

††. https://www.wsj.com/articles/cost-of-police-misconduct-cases-soars-in-big-u-s-cities-1437013834

ments) a full deployment of BWCs would cost \$10 million per year at most. Given that in the study sample the reduction occurred at the much lower deployment rate of 68%, I conclude that BWCs can also serve as a powerful cost-saving tool for police departments.

1.8 Conclusion

Since the early 2010s, BWCs have played a prominent role in efforts to improve police accountability and transparency. Currently policy debates concern whether or not to continue to expand BWC programs. Previous research on BWCs has not provided conclusive answers due to the limitations inherent to research settings.

This cross-agency study is the first to examine the effects of BWCs in a setting that deals credibly with intra-agency spillovers and idiosyncratic agency characteristics. An important implication of this study is that investments in BWCs can deliver large returns evident in improved police-civilian interactions. This study also delivers the first evidence on the effects of BWCs on agency-wide law enforcement outcomes such as crime control activities and public opinion toward the police.

The finding that agencies adopting BWCs have largely preserved policing capabilities contrasts with findings from previous literature on police accountability. Several previous studies that have examined the effects of enhancing police accountability through natural experiments such as use-of-force scandals and federal investigations have found that increase in police oversight can lead to reduction in policing efforts. Why didn't BWC adoption hamper policing capabilities in a similar way? One reason may be that BWCs did not pose serious threats to officers. In my survey data of police chiefs, over 90% of agencies that have adopted BWCs stated lack of officer support was not an obstacle to BWC implementation, and over 95% of agencies that did not adopt BWCs stated the same reason was not an obstacle. If officers believe BWCs prevent false accusations of police misconduct, they may increase policing efforts in certain types of citizen interactions. Relatedly, officers may enjoy strong enough support from the agency leadership and unions that oversight through BWCs does not lead to extreme disengagement from officers.

Prendergast (2020) makes this point where he shows that when officer engagement is high enough – for example through intrinsic motivation or victim oversight – and officer oversight is low enough, more oversight through monitoring may reduce excessive force without decreasing the level of engagement. Related empirical evidence comes from Rivera and Ba (2019) who find that self-monitoring in Chicago Police Department led by a memo from the police union reduced civilian complaints without reducing policing efforts, while a scandal increased both civilian complaints and crime rates. Further improving our understanding of how we can enhance police accountability and performance, and identifying what factors trigger reduction in policing efforts, are much needed areas of future research.

In conclusion, the results from this study show that BWCs can help police departments meet the societal demands placed on them by recent high-profile excessive policing cases and ensuing protests.

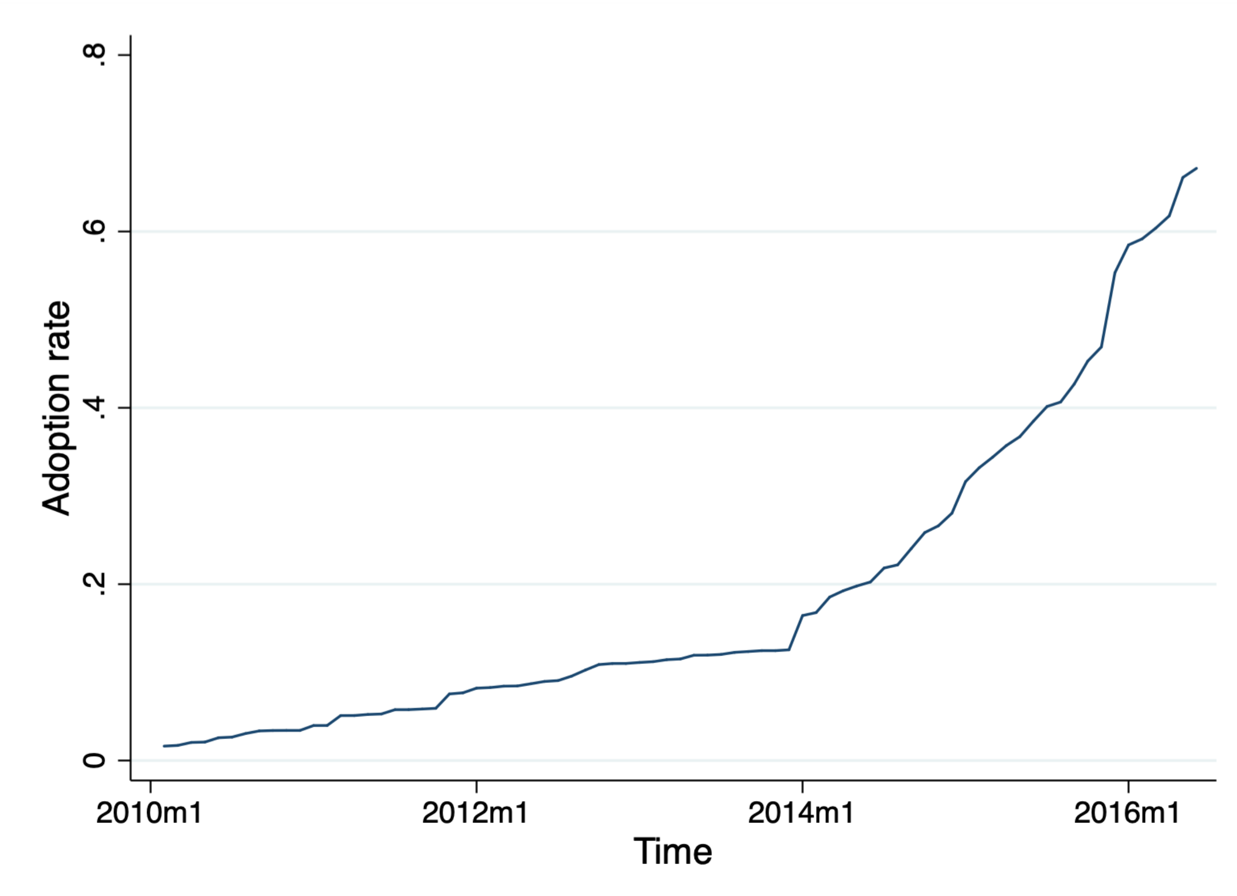


Figure plots the nationwide adoption trends of BWCs by local agencies. The solid trend line is the population weighted proportion of agencies that adopted BWCs by each point in time.

Figure 1.1: National Trend of BWC Adoption

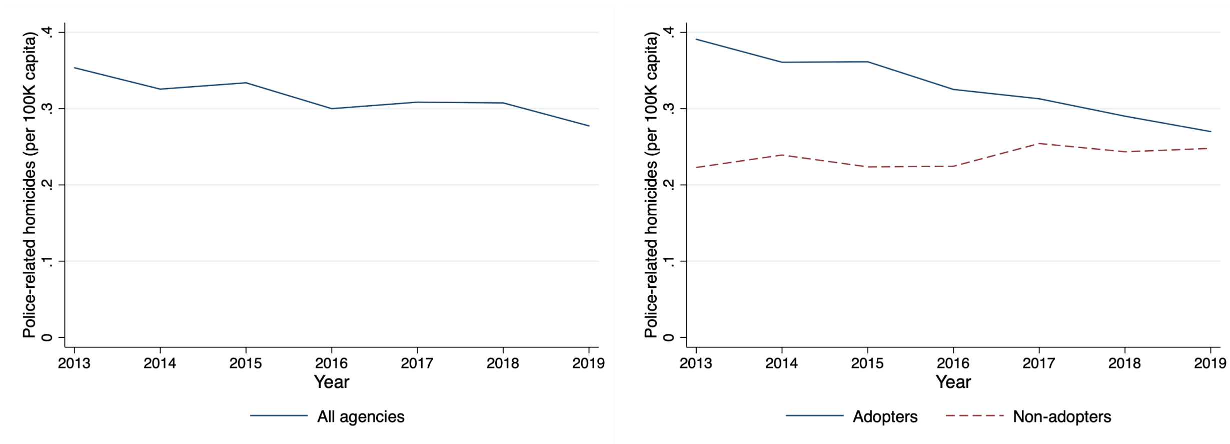
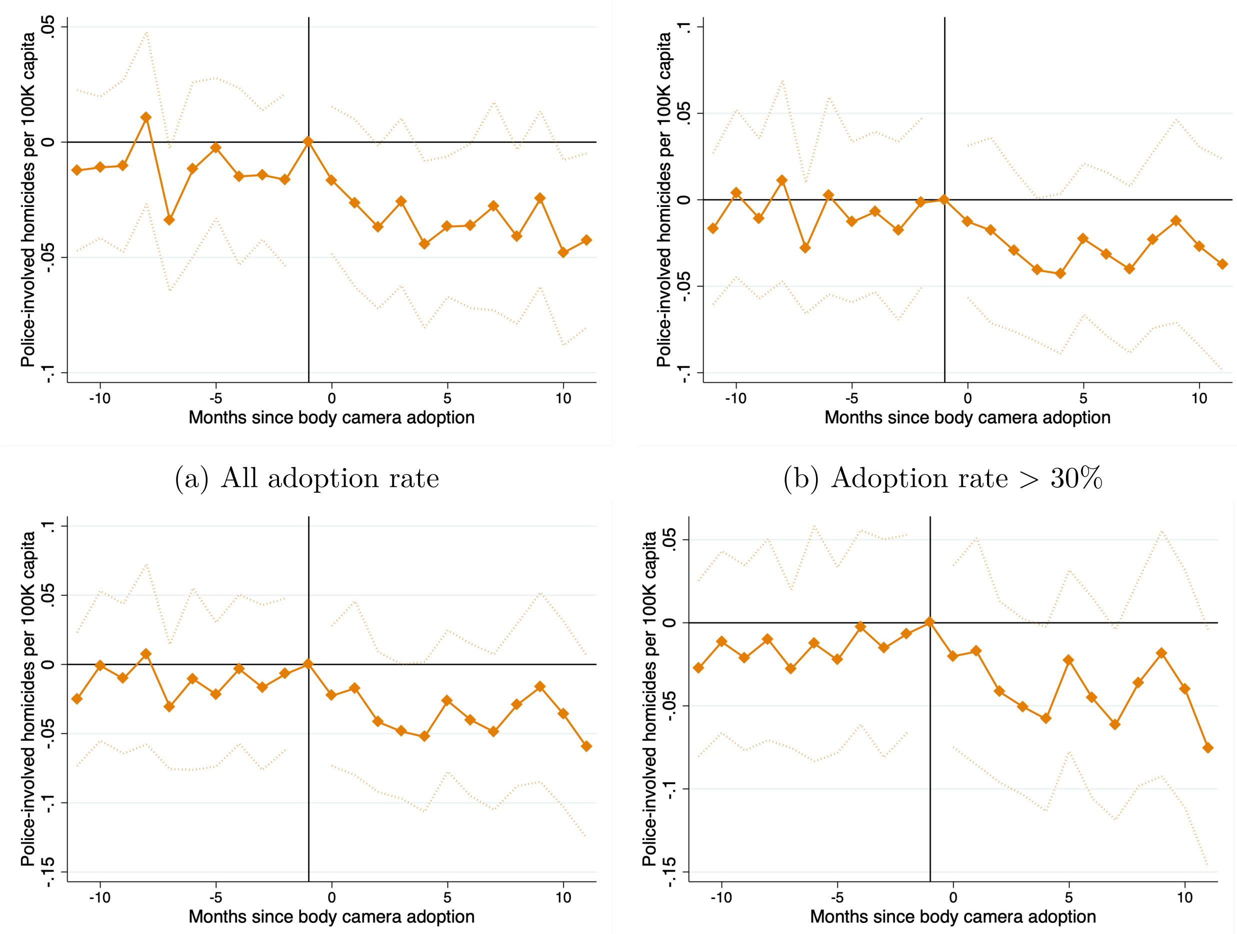


Figure describes the time series of the main data (MPVD) on police-involved homicides for agencies that are matched with the LEMAS data. I link this data with LEMAS data on BWC adoption status. The left panel shows trends for all agencies matched with the adoption data. I split the time series of police-involved homicides into adopters and non-adopters. Adopters are agencies that I observe to have adopted BWCs at any point in time in the sample.

Figure 1.2: Description of Data on Police-involved Homicides by Adoption Status



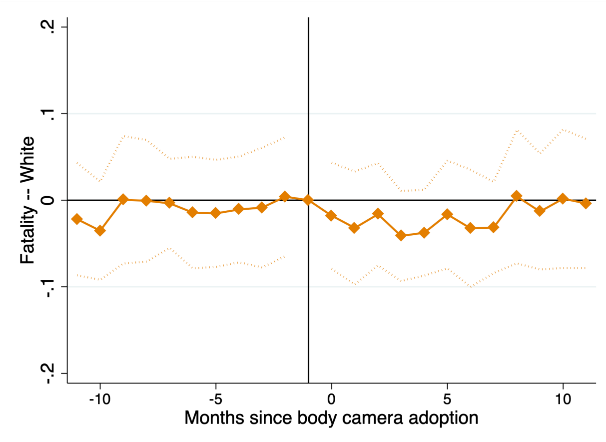
(a) All adoption rate

(b) Adoption rate > 30%

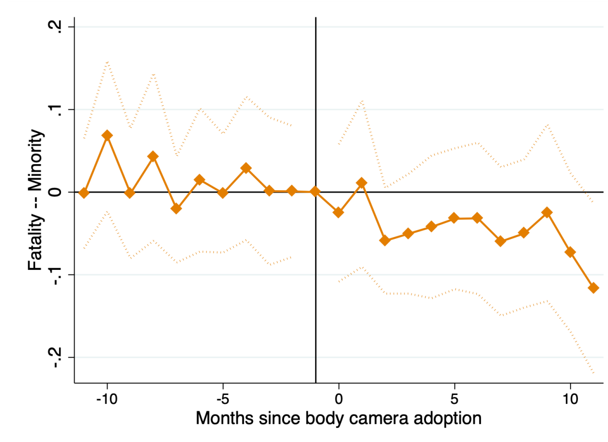
(c) Adoption rate > 40%

(d) Adoption rate > 50%

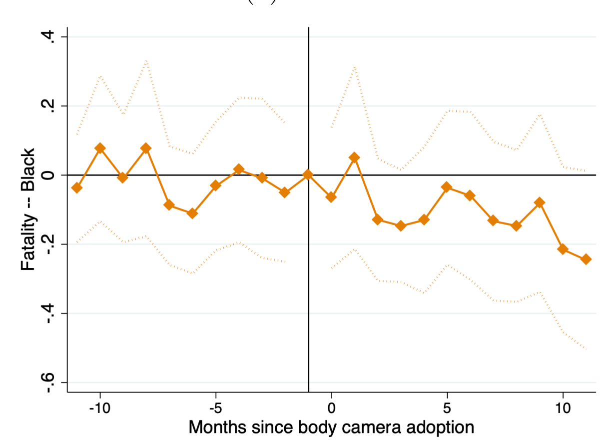
Figure 1.3: DID Effects of BWCs on Police-involved Homicides for Adopters
Figure plots the DID estimates (with 95% confidence intervals) of homicides by police around BWC adoption. The sample includes adopters only; early adopters (adopting between 01/2014-06/2015) are matched with any later adopters adopting 1 year or later. Dependent variable is the number of police-caused homicides relative to the population the agency serves. There are 1,001 agencies in the sample with half being early adopters. The first panel includes adopting agencies in the sample; the second panel includes adopters with adoption rate greater than 30%; the others have 40% and 50% as the thresholds. Coefficients are normalized to t = -1. All regressions include time FE, agency-specific time trends, and agency FE and are clustered at the agency level. The regressions are weighted by population the agency serves.



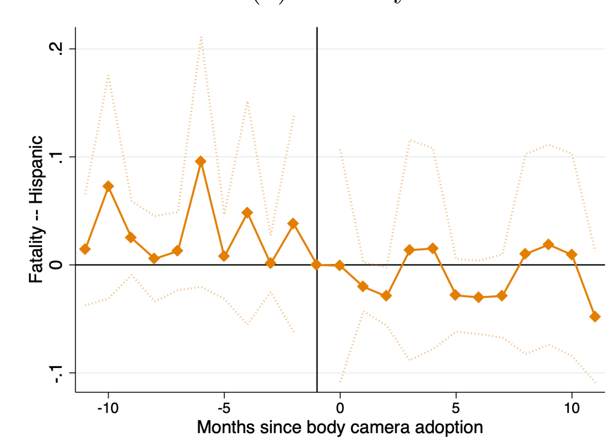
(a) White



(b) Minority



(c) Black



(d) Hispanic

Figure 1.4: DID Effects of BWCs by Race

Figure plots the DID estimates (with 95% confidence intervals) of homicides by police around BWC adoption for different race of the subjects. The plots represent the estimates for agencies that adopt by 40% or more. The regressions are weighted by population of the racial group the agency serves.

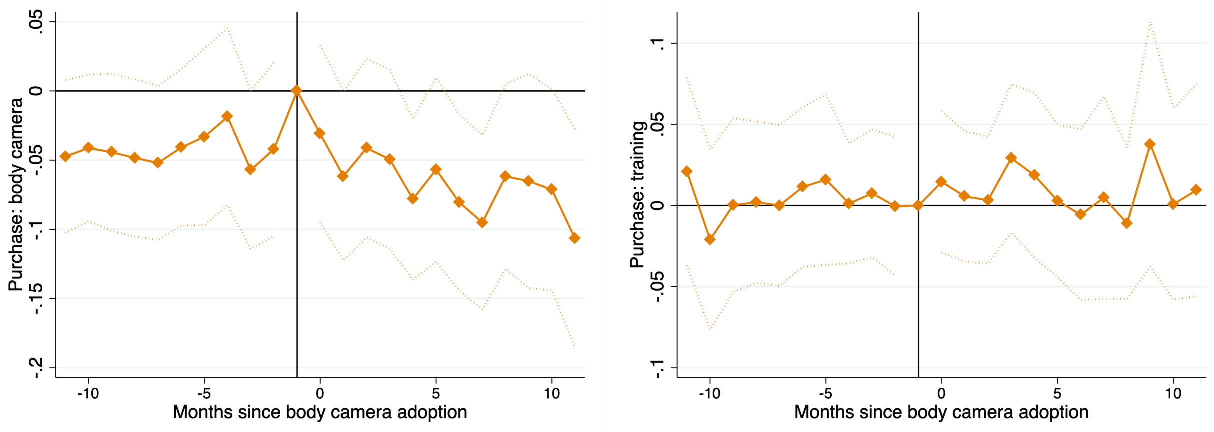


Figure plots the difference-in-difference estimates for purchase orders around the timing of adoption. The sample includes adopters only. All regressions include time FE, agency-specific time trends, and agency FE and cluster at the agency level.

Figure 1.5: Purchase orders near BWC Adoption

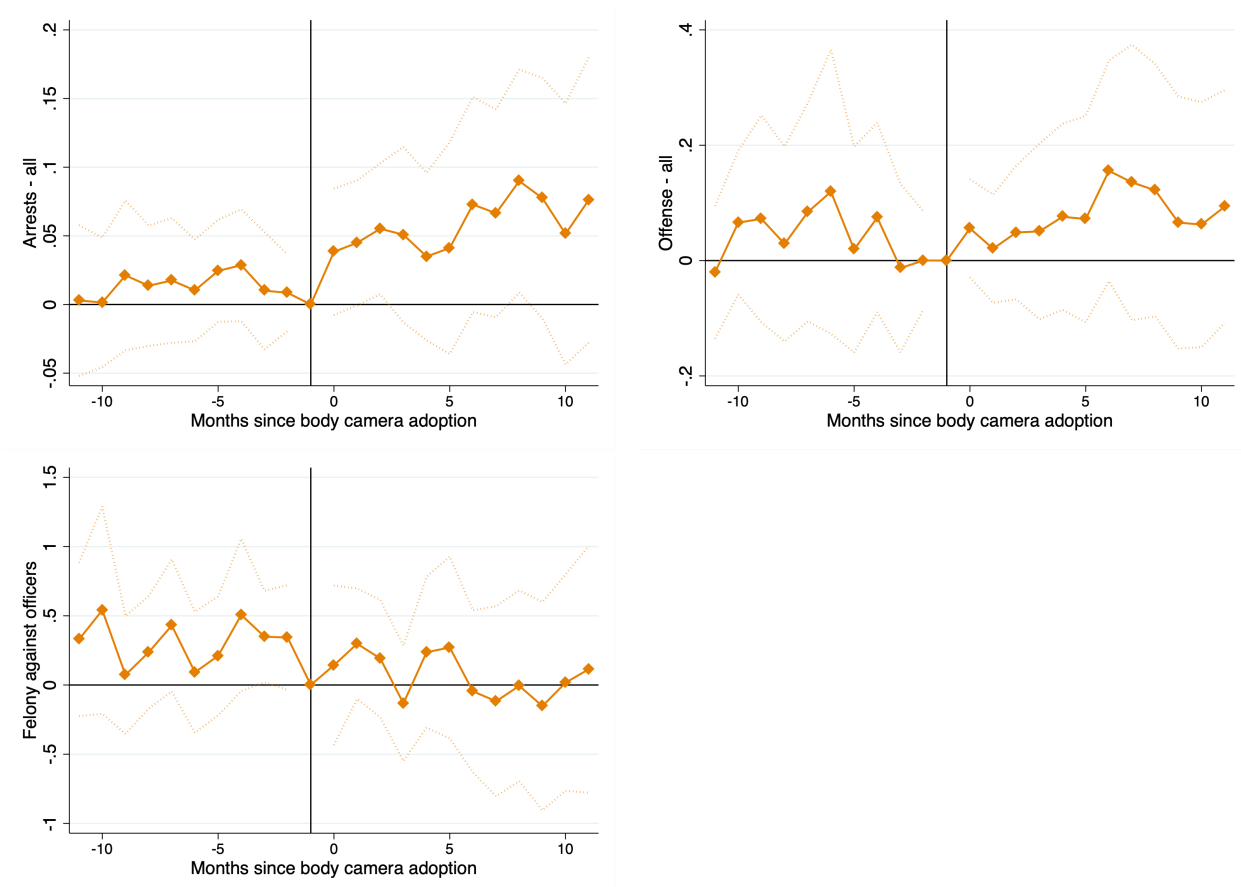
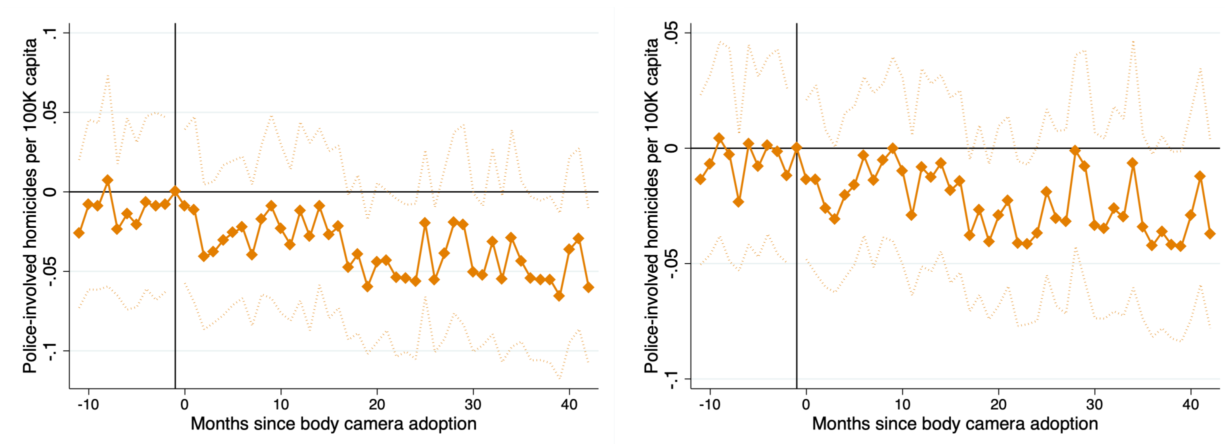


Figure plots the DID estimates of arrest rates, index crime rates, and felonies against the police around BWC adoption. The results are presented for agencies that adopt BWCs by more than the rate of 40%. See Figure 1.3 for details.

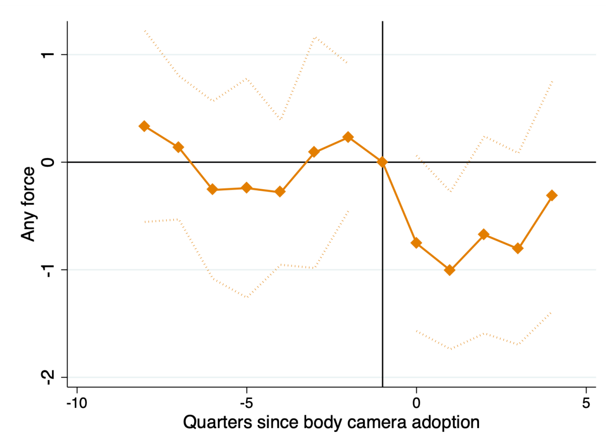
Figure 1.6: Effects of BWCs on Other Police Performance



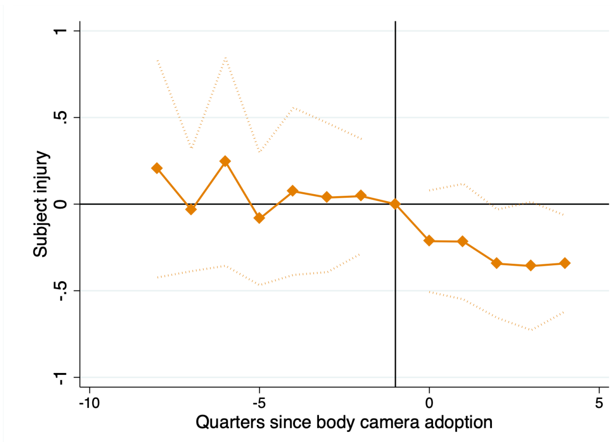
(a) Main adopter sample

(b) All adopter sample

Figure 1.7: Long-run Effects of BWCs on Police-involved Homicides
Figure plots the DID estimates (with 95% confidence intervals) of homicides by police around BWC adoption. The sample includes adopters and non-adopters. In the first panel, I use the sample of adopters I've used in the short-run analyses (adopting between 01/2014-06/2015. In the second panel, since I am no longer constrained to have an adopting control group, I include all adopters (including those adopting between 07/2015-06/2016) as well as non-adopters. Dependent variable is the number of police-caused homicides relative to the population the agency serves (per 100K capita). Adopters are defined as any that have adopted at rates of 40% or more. Coefficients are normalized to t = -1. All regressions include time FE, agency-specific time trends, and agency FE and are clustered at the agency level. The regressions are weighted by population the agency serves.



(a) Use of force



(b) Subject injury

Figure 1.8: Effects of BWCs on Use of Force in New Jersey
Figure plots the DID estimates of use of force around BWC adoption. On the New Jersey data from 2012 to 2016, I run event study comparing early adopting agencies (between 2014 and 2015) with agencies that adopt 4 quarters later or more. All event studies control for agency FE, time FE, and controls for agency-specific time trends. Standard errors are clustered on agency.

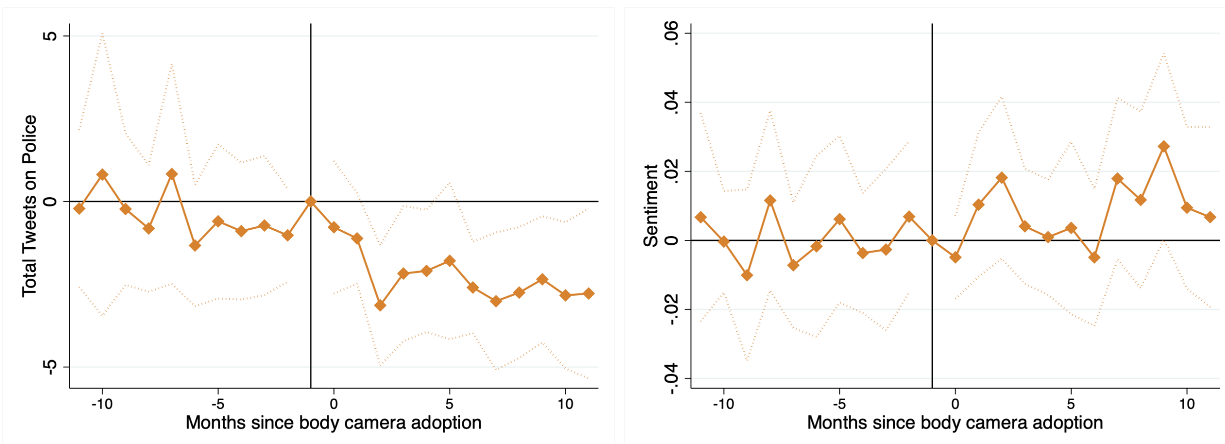


Figure plots the difference-in-difference estimates for public opinion as captured in Twitter. In the first panel the dependent variable is total tweets index (tweets in a month adjusted for average tweets over sample) and in the second panel it is sentiment. The sample includes adopters only that have tweet sample throughout the sample period. All regressions include time FE, agency-specific time trends, and agency FE and cluster at the agency level.

Figure 1.9: Effects of BWCs on Public Opinion in Twitter

Table 1.1: Effects of Body Cameras on the Use of Force in Previous RCT Studies

Paper	Location	Effects	Pre-post (control)
Ariel et al. (2015)	Rialto, CA	-53%	-50%
Jennings et al. (2015)	Orlando, FL	-26%	-38%
Ariel et al. (2016)	10 experiments	Overall null; -38% to +50%	
Henstock and Ariel (2017)	UK	-19%	
Yokum et al (2017)	DC	Null	Null
Peterson et al (2018)	Milwaukee, WI	Null	
Braga et al (2018)	Las Vegas, NV	-46%	Null

Table 1.2: Survey: Obstacles in Using BWCs

Greatest obstacles encountered	%
Costs greater than anticipated	13.5%
Storage Procedures	11.1%
Privacy	7.7%
Technical obstacles	6.6%
Public request burden	5.5%
Security	2.9%
Liability	2.3%
Officer support	1.2%
No benefits	0.4%
Public support	0%
Others	2.7%
Missing	46.1%
Total	1,788

Notes: This table shows survey answers from all local agencies that adopted BWCs as to what was the single greatest obstacle they faced. The data source is the LEMAS-BWCS.

Table 1.3: Comparison of Adopters and Non-adopters

	(1)	(2)	(3)	(4)	(5)	(6)
	Non-adopters	Adopters	Early-adopters	Late-adopters	None vs. adopt	Early vs. late
Population (10K)	4.06	7.81	7.01	8.61	-3.75**	-1.59
	(15.85)	(34.27)	(25.96)	(40.97)	(-3.22)	(-0.73)
White	82.95	76.97	77.22	76.72	5.98***	0.50
	(18.17)	(20.38)	(20.65)	(20.12)	(7.39)	(0.39)
Black	8.24	13.61	13.81	13.41	-5.37***	0.40
	(15.05)	(18.73)	(19.24)	(18.21)	(-7.49)	(0.34)
Hispanic	9.96	12.61	11.56	13.68	-2.65***	-2.12
	(15.33)	(17.55)	(16.99)	(18.06)	(-3.84)	(-1.91)
Male	48.76	48.73	48.67	48.78	0.03	-0.10
	(3.24)	(3.86)	(4.12)	(3.59)	(0.21)	(-0.43)
Income	56693.72	48973.07	48030.65	49921.16	7720.65***	-1890.51
	(25017.66)	(21988.03)	(23050.68)	(20844.64)	(7.98)	(-1.36)
College	27.62	24.17	23.75	24.59	3.45***	-0.84
	(16.66)	(14.13)	(14.48)	(13.78)	(5.44)	(-0.94)
Poverty	14.04	17.73	18.21	17.24	-3.69***	0.98
	(9.09)	(9.35)	(9.65)	(9.02)	(-9.61)	(1.66)
Crime	2.20	2.80	2.74	2.87	-0.60***	-0.14
	(2.54)	(2.24)	(2.53)	(1.89)	(-5.78)	(-0.90)
Citizen Deaths	0.00	0.00	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(-1.81)	(-0.29)
Officers	56.85	184.02	170.58	197.54	-127.17**	-26.97
	(164.70)	(1274.99)	(863.97)	(1585.23)	(-3.14)	(-0.33)
N	1379	1001	502	499	2380	1001

Notes: Table shows sample statistics (mean and standard errors) for agencies analyzed in this study by different categories: non-adopters, adopters, early- and late adopters. The demographics data come from 5-year estimates of the American Community Survey of 2013. In the last two columns, I test equality of Column (1) versus (2), and Column(3) versus (4).

Table 1.4: Predicting Adoption of BWCs

	US sample				
	Adopt	Months	Months	Months	Months
Population (10K)	0.000	-0.000	0.019	0.055	0.074
	(0.001)	(0.029)	(0.063)	(0.101)	(0.104)
White	0.001	-0.010	-0.012	-0.011	-0.018
	(0.001)	(0.035)	(0.046)	(0.047)	(0.047)
Black	0.004**	0.012	-0.009	-0.003	-0.022
	(0.002)	(0.039)	(0.050)	(0.051)	(0.051)
Hispanic	0.002**	0.030	-0.000	0.001	-0.005
	(0.001)	(0.020)	(0.025)	(0.025)	(0.025)
Male	0.001	-0.107	-0.115	-0.128	-0.120
	(0.003)	(0.081)	(0.084)	(0.085)	(0.086)
Income (10K)	-0.017**	-0.201	-0.365	-0.212	-0.195
	(0.008)	(0.232)	(0.274)	(0.290)	(0.298)
College	0.001	0.033	0.024	0.021	0.018
	(0.001)	(0.032)	(0.041)	(0.043)	(0.045)
Poverty	0.005***	-0.054	-0.047	-0.034	-0.015
	(0.002)	(0.049)	(0.054)	(0.058)	(0.061)
Crime	0.008*	0.050	0.027	-0.054	0.011
	(0.005)	(0.142)	(0.200)	(0.219)	(0.229)
Citizen Deaths	0.072	-0.139	-0.158	-0.156	-0.131
	(0.048)	(1.008)	(0.979)	(0.981)	(0.979)
Officers	0.000	-0.000	-0.002	-0.005	-0.006
	(0.000)	(0.001)	(0.003)	(0.006)	(0.006)
Sample	All	Adopters	Adopt 30%+	Adopt 40%+	Adopt 50%+
Dep. var. mean	.418	16.6	15.2	15.2	15.1
Observations	2,125	889	581	540	514

Notes: Table reports results from running cross sectional regression on the decision to adopt BWCs in the sample. The dependent variable for the decision to adopt in the first column is a binary variable equal to one if the agency ever adopted BWCs. In the second column, the dependent variable for the timing is the number of months away from January 2014. The sample sizes differ from the main sample of analysis because of data availability for some variables. For the last three columns, I repeat the exercise of the second column with higher thresholds of adoption rates.

Table 1.5: DID Effects of BWCs on Police-involved Homicides

	(1) All	(2) $\geq 30\%$	(3) $\geq 40\%$	(4) $\geq 50\%$
BWC	-0.019**	-0.020*	-0.024**	-0.023**
	(0.009)	(0.011)	(0.011)	(0.012)
Observations	163,326	102,123	94,924	90,485

Notes: Table shows difference-in-differences estimation of the effects on police-caused homicides from body camera adoption using data between 2013 and June of 2016. The dependent variable is the number of homicides per population the agency serves. The first column is a regression including all adopters; in the next columns, I restrict the sample to those that adopt at rate (cameras per officers) over certain thresholds. All regressions include agency FE, time FE, and linear time trends. Standard errors are clustered at the agency level and all regressions are weighted by the population.

Table 1.6: DID Effects of BWCs on Police-involved Homicides by Different Types of Agencies

	(1)	(2)	(3)	(4)
Urban	-0.029**			
	(0.012)			
Not urban	0.001			
	(0.029)			
Low white		-0.031***		
		(0.012)		
High white		-0.004		
		(0.022)		
Required: low			-0.003	
			(0.015)	
Required: high			-0.038***	
			(0.014)	
Low adoption				-0.016
				(0.011)
High adoption				-0.025**
				(0.011)
Observations	94,924	94,924	94,924	163,326

Notes: Table shows difference-in-differences estimation of the effects on police-involved homicides by different types of agencies. For agencies that have adopted by more than 40% rate, I estimates effects for agencies that are in urban versus non-urban areas, agencies with low white populations (less than 70%) versus high white populations, and agencies that are more strict in terms of requirements of when to turn on BWCs (more than median level) versus those are less strict. I also examine whether agencies that have adopted at higher rates (more than 50% in terms of BWCs/officers) have differential effects relative to those that have adopted at lower rates. See Table 1.5 for details.

Table 1.7: DID Effects of BWCs on Police-involved Homicides by Race

	(1) All	(2) $\geq$ 30%	(3) $\geq$ 40%	(4) $\geq$ 50%
White	-0.009	-0.021	-0.021	-0.023
	(0.009)	(0.013)	(0.015)	(0.016)
Minority	-0.026**	-0.026	-0.037**	-0.034*
	(0.013)	(0.020)	(0.018)	(0.020)
Black	-0.042	-0.030	-0.054	-0.040
	(0.026)	(0.042)	(0.042)	(0.045)
Hispanic	-0.024*	-0.032*	-0.034*	-0.037*
	(0.013)	(0.018)	(0.019)	(0.021)
Observations	163,346	102,143	94,944	90,505

Notes: Table shows difference-in-differences estimation of the effects on police-related homicides by race using data between 2013 and June of 2016. All regressions include agency FE, agency-specific time trends and time FE. All regressions are weighted by the relevant population. Outcome measures are the number of citizen fatality of the racial group relative to the total count of that racial group. Standard errors are clustered at the agency level.

Table 1.8: DID Effects of BWCs on Use of Force by Police

	(1) All	(2) Injury
BWC	-0.723***	-0.313***
	(0.261)	(0.092)
Observations	6,929	6,929

Notes: Table shows difference-in-differences estimation of the effects on police-caused use of force from body camera adoption using data between 2012 and 2016. The dependent variable is the occurrences of use of force per population the agency serves. All regressions include agency FE, time FE, and linear time trends. Standard errors are clustered at the agency level and all regressions are weighted by the population.

Table 1.9: Effects of BWCs Police Performance

	(1) All	(2) $\geq 30\%$	(3) $\geq 40\%$	(4) $\geq 50\%$
Arrests	0.008	0.043*	0.036	0.046*
	(0.016)	(0.026)	(0.027)	(0.028)
Observations	154,839	88,438	82,665	78,571
Crimes	0.015	0.020	0.033	0.045
	(0.051)	(0.062)	(0.064)	(0.066)
Observations	154,839	88,438	82,665	78,571

Notes: Table shows difference-in-differences estimation of the effects on police performance measures from body camera adoption using data between 2013 and June of 2016. See Table 1.5 for details.

Table 1.10: Differential Impacts by Benefits Received

	(1)	(2)	(3)
Not agree: Professionalism	-0.012		
	(0.015)		
Agree: Professionalism	-0.029**		
	(0.013)		
Not agree: Misconducts		-0.018	
		(0.014)	
Agree: Misconducts		-0.030**	
		(0.014)	
Not agree: Training			-0.020
			(0.012)
Agree: Training			-0.025*
			(0.013)
Observations	94,924	94,924	94,924

Notes: Table shows difference-in-differences estimation of the impact on police-involved homicides from body camera adoption for different agencies with various perceived benefits. For each column, I test if agencies that mentioned certain benefits of BWCs had different levels of reduction relative to those that did not mention that benefit. See Table 1.5 for details.

Table 1.11: DID Effects of BWCs on Twitter sentiment

	(1) All	(2) $\geq$ 30%	(3) $\geq$ 40%	(4) $\geq$ 50%
Total tweets	-1.101**	-1.254***	-1.267**	-1.101**
	(0.521)	(0.469)	(0.503)	(0.545)
Observations	56,810	19,964	18,607	17,135
Sentiment	0.002	0.004	0.005	0.003
	(0.006)	(0.005)	(0.005)	(0.006)
Observations	56,810	19,964	18,607	17,135

Notes: Table shows difference-in-differences estimation of the effects on Twitter sentiment from body camera adoption using data between 2013 and June of 2016. See Table 1.5 for details.

1.A Appendix Figures and Tables

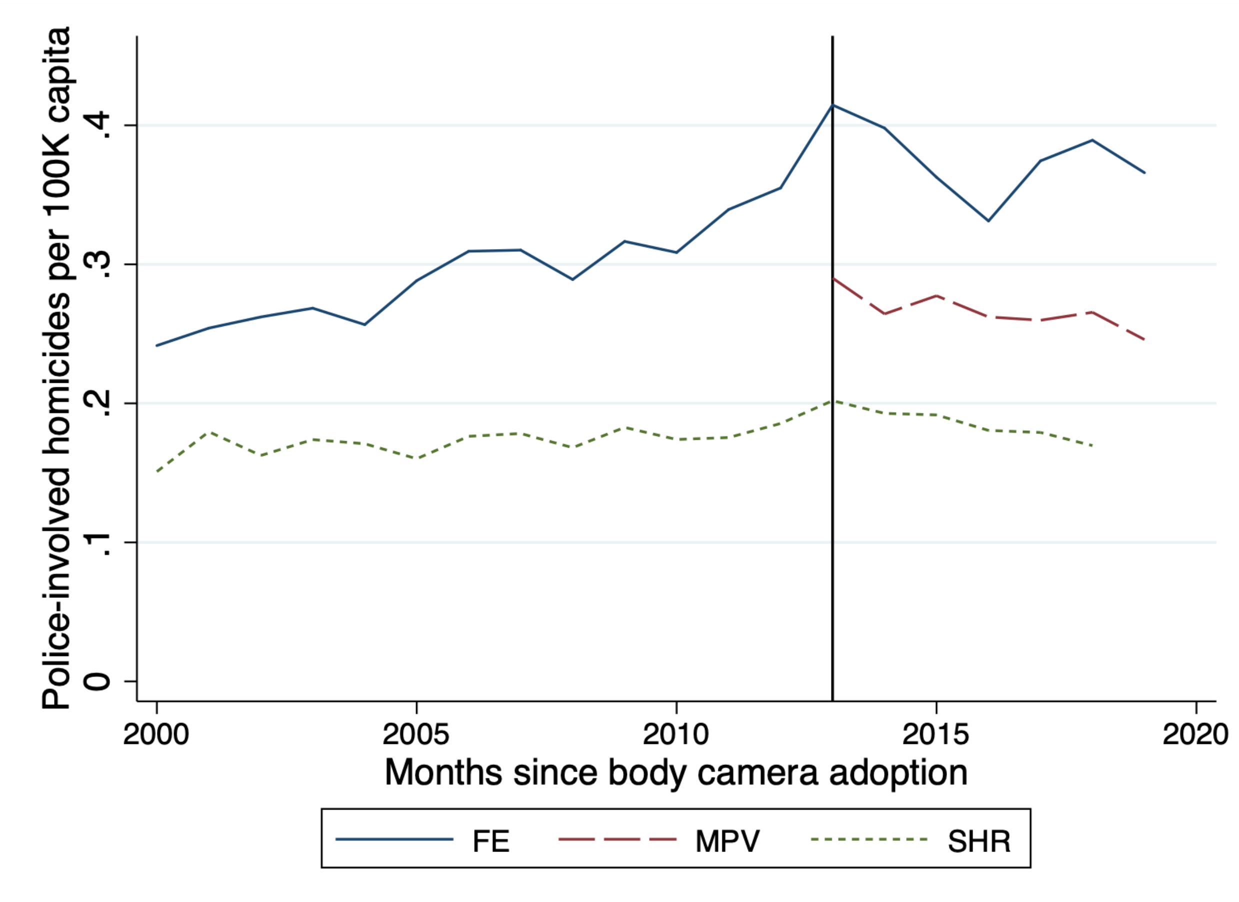
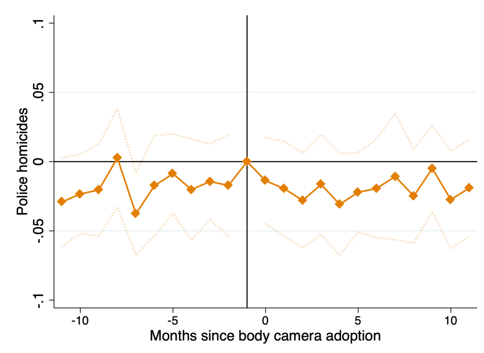
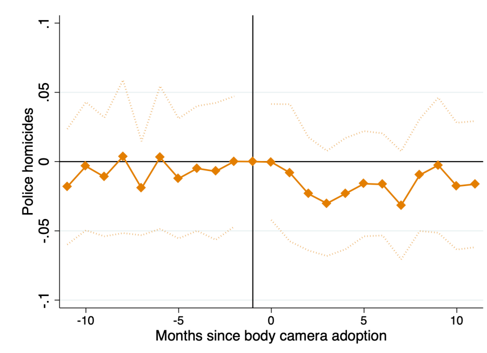


Figure describes the time series of various data sources on police-involved homicides. FE represents Fatal Encounters, MPVD represents Mapping Police Violence which is the data used in this project, and SHR is from the Supplemental Homicides Report (SHR) of the UCR. SHR is voluntary submission of local agencies and have well-documented data quality problems. MPVD compiles the three largest crowd-sourced data including FE and accounts for homicide cases where police are directly involved. For year 2013 (vertical line) and before, FE were gathered retrospectively and have lower quantity and quality of cases because it was difficult to gather and check crowd-sourced cases. The main analysis uses the Mapping Police Violence which synthesizes this data as well as other crowd-sourced data from 2013. The MPV trends in this figure are slightly lower than Figure 1.2, because I have plotted MPV before merging with the LEMAS data.

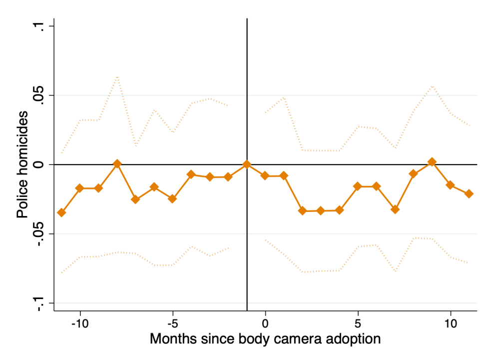
Figure 1.A.1: Time series of various data sources on police-involved homicides



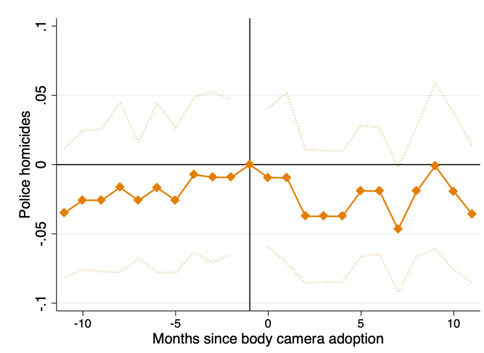
(a) All adoption rate



(b) Adoption rate > 30%



(c) Adoption rate > 40%



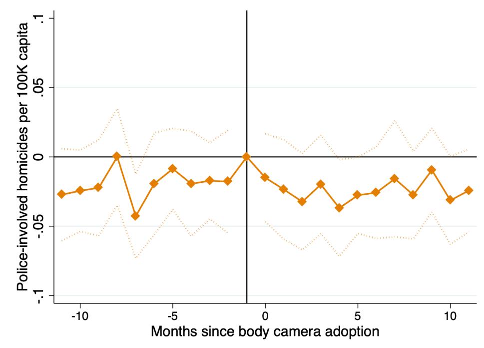
(d) Adoption rate > 50%

Figure 1.A.2: Event study on police-involved homicides including non-adopters

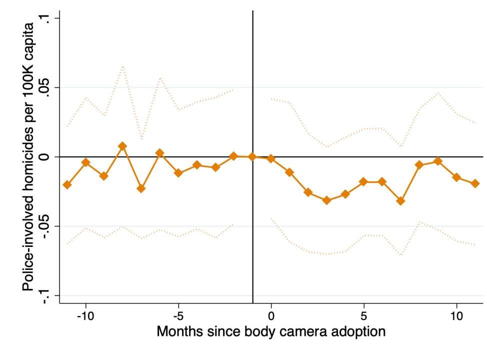
For robustness, I perform an event study estimates around BWC adoption instead of the difference-in-differences design which is my main specification. The main dependent variable is police-related homicides relative the population. The sample includes both adopters and non-adopters. Estimates come from the following specification:

$$Outcome_{jt} = \sum_{\tau=-12}^{12} D_{jt}^{\tau} BWC_j + \phi_j + \delta_t + \phi_j \times f(time)_t + \epsilon_{jt}, \tag{1.8}$$

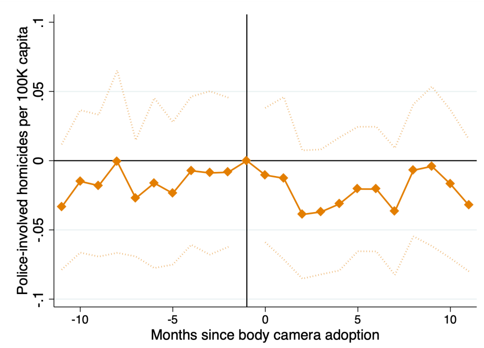
where BWC_j is an indicator for whether the agency ever adopts BWCs. The event study framework compares adopting agencies against non-adopting ones controlling for covariates and secular trends. I present the event study estimates D_{jt}^{τ}. I examine agencies that adopt between January 2014 and June 2015 to ensure I have balanced sample in the event study of 1 year before and after adoption. Event times further away from 11 months from adoption on either side in the event study window are coded as -12 or 12. There are such 501 adopters and 1,878 total agencies including non-adopters. The first panel includes adopting agencies in the sample; the second panel includes adopters with adoption rate greater than 30% and so on. All regressions include time FE, agency-specific time trends, and agency FE and cluster at the agency level. All regressions are weighted by population.



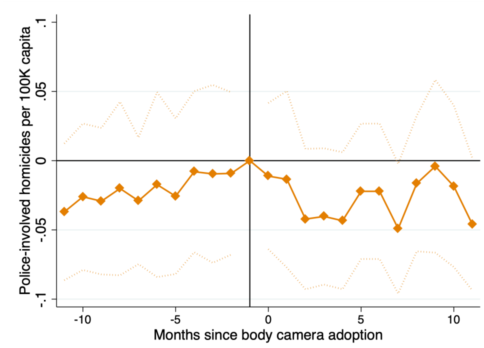
(a) All adoption rate



(b) Adoption rate > 30%

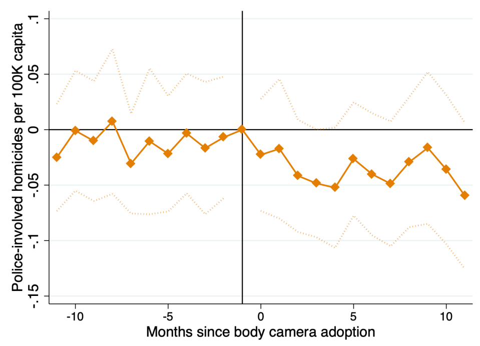


(c) Adoption rate > 40%

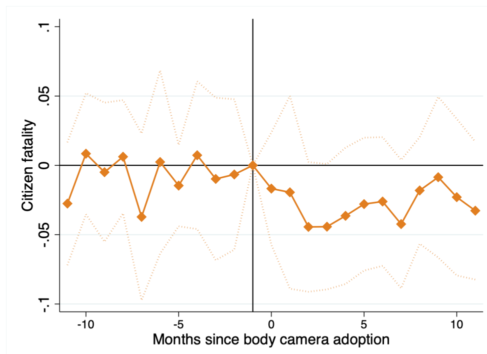


(d) Adoption rate > 50%

Figure 1.A.3: Effects of body cameras on citizen fatality without trends
For robustness, I perform the main analysis of event study comparing early versus late adopters without agency-specific trends. All others remain the same.

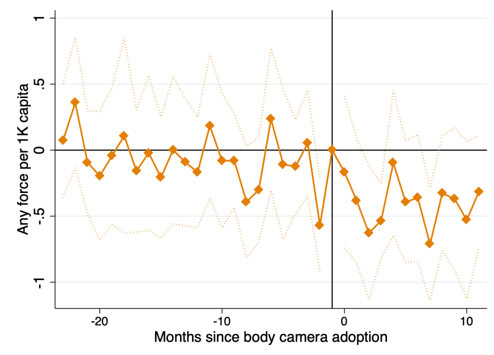


(a) Main estimates

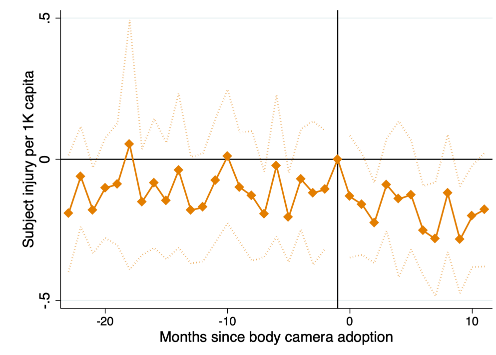


(b) Staggered treatment correction

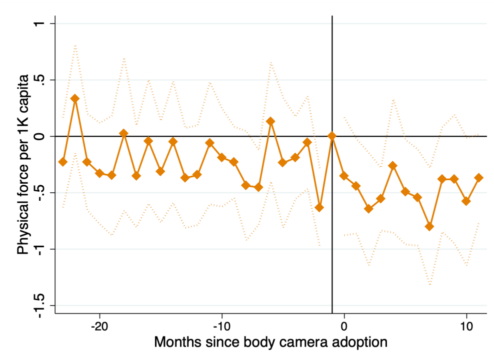
Figure 1.A.4: Effects of body cameras on citizen fatality: staggered treatment correction
Notes: I use the method of Callaway and Sant'Anna (2020) to correct for staggered adoption timing in estimating the effects of BWC adoption. This approach involves estimating difference-in-differences for each adoption month cohort with control groups (never adopters). For comparison, I include main estimates in the left panel that are estimated on agencies that have adopted at 40% or more. The DID coefficients are then averaged across different cohorts with weights using each cohort's share of all treated agencies. The coefficient is normalized at -1 time period, just before BWC adoption.



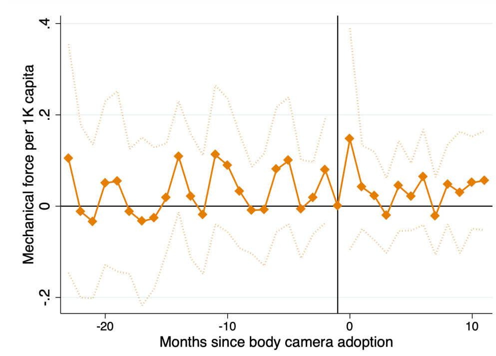
(a) Use of force



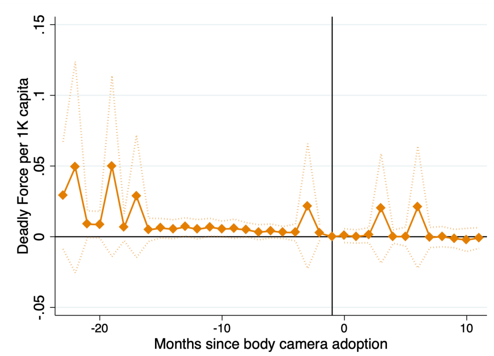
(b) Subject injury



(c) Physical force

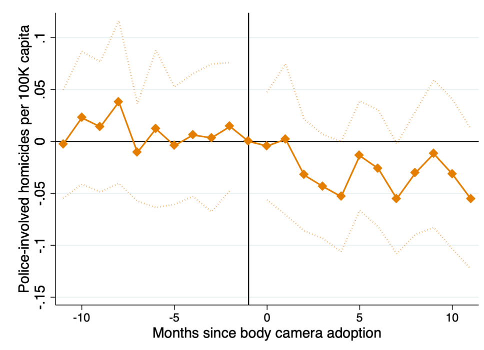


(d) Mechanical force

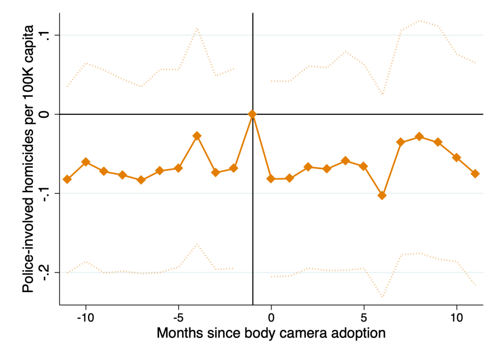


(e) Deadly force

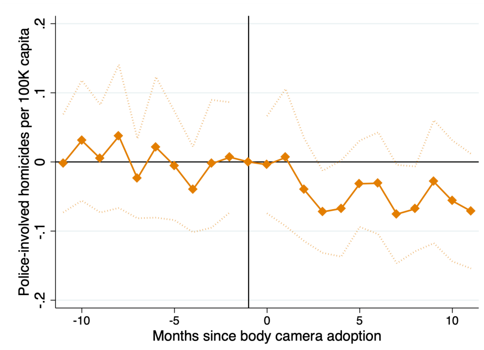
Figure 1.A.5: Effects of body cameras on use of force in New Jersey: monthly estimates
Figure plots the DID monthly estimates of use of force around BWC adoption. Using the use of force data in New Jersey from 2012 to 2016, each early adopters (adopting between 01/2014 and 12/2015) are compared with later adopters adopting one year or later. All event studies include time FE, agency FE, and agency-specific linear trends and are weighted by population. All standard errors are clustered by agency



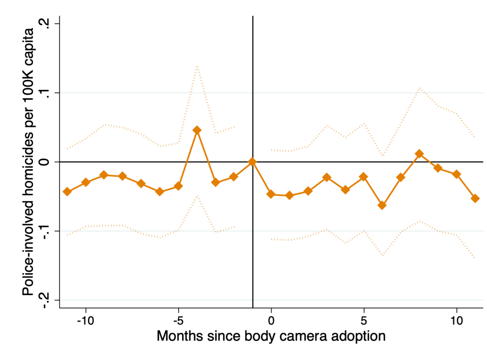
(a) Agree: Professionalism



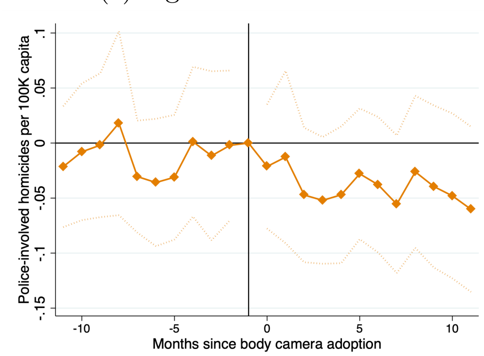
(b) Not agree: Professionalism



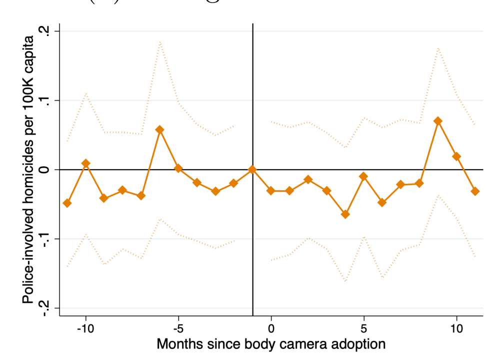
(c) Agree: Misconducts



(d) Not agree: Misconducts



(e) Agree: Training



(f) Not Agree: Training

Figure 1.A.6: Effects of body cameras on citizen fatality without trends

Notes: I show difference-in-differences estimation of the impact on police-involved homicides from body camera adoption for different agencies with various perceived benefits. For each row, I test the effects of BWCs for different stated benefits. In each row, the left panel displays results from agencies that agreed with a stated benefit while the right panel does not. "Professionalism" indicates the statement that BWCs improves officers' professionalism. "Misconducts" indicates the statement that BWCs help identify cases of misconducts. "Training" indicates that statement that BWCs facilitate officer training.

Table 1.12: Testing for Alternative Explanations and Robustness

	(1) All	(2) $\geq 30\%$	(3) $\geq 40\%$	(4) $\geq 50\%$
Removing "Reforms"	-0.006	-0.025	-0.040**	-0.049***
	(0.009)	(0.016)	(0.016)	(0.016)
Observations	64,518	42,990	40,529	39,241
Removing close BLM protests	-0.015*	-0.022*	-0.022*	-0.024*
	(0.009)	(0.012)	(0.012)	(0.013)
Observations	162,863	101,936	94,760	90,344
Not having obstacles	-0.024**	-0.023*	-0.026**	-0.023*
	(0.010)	(0.012)	(0.012)	(0.012)
Observations	160,773	100,007	92,877	88,530
Using FE	-0.022**	-0.020*	-0.023**	-0.025**
	(0.009)	(0.011)	(0.012)	(0.012)
Observations	163,326	102,123	94,924	90,485
Officers adjusted	-8.049**	-9.516*	-12.186**	-12.204*
	(3.790)	(5.539)	(5.863)	(6.276)
Observations	159,227	98,036	90,903	86,464
PPML	-0.627**	-0.762	-0.867	-0.839
	(0.318)	(0.641)	(0.650)	(0.712)
Observations	35,236	9,499	8,982	7,907

Notes: Table conducts robustness checks for the effects on police-involved homicides. Each row indicates the coefficients on the main coefficient of interest (BWC) for different checks. In the first row, I remove reform-based adoption (agencies that adopted to enhance accountability) and only include agencies that have adopted BWCs to improve other police functions like prosecutability. In the second row, I remove agencies that experience Black Lives Matter protests in three months leading up to BWC adoption. In the third row, the estimates come from agencies that did not face logistical obstacles in implementing BWCs. The fourth row uses Fatal Encounters as the data for the outcome variable instead of Mapping Police Violence. In the fifth row, I use an alternative definition of the outcome variable, police-involved homicides per officers, and weight by the number of officers. In the sixth row, I run Poisson psuedo-maximum likelihood estimation (PPML) instead of OLS. As noted in Santos Silva and Tenreyro (2010), maximum likelihood estimates of Poisson models might not exist due to the problem of "statistical separation." To deal with this convergence problem, I drop agencies that never have police-involved homicides. This leaves 175 agencies in the first regression and 152 in the last regression. The dependent variable is the number of homicides per population the agency serves. The first column is a regression including all adopters; in the next columns, I restrict the sample to those that adopt at rate (cameras per officers) over certain thresholds. All regressions include agency FE, time FE, and linear time trends. Standard errors are clustered at the agency level and all regressions are weighted by the population.

1.B Derivation in model on police behavior

In this section, I derive in more detail the predictions from the model on police behavior in Section 2.2. In Period 2, the first order condition with respect to e yields:

$$e = -N'(e) \times q \times \rho_h \times \rho_i \times \Delta + f'(N'(e)) \times N' \tag{1.9}$$

Taking derivatives with respect to the cost of investigation, $\rho_i \times \Delta$, on both sides and rearranging, we have

$$\frac{\partial e}{\partial(\rho_i \times \Delta)} = \frac{N'(e) \times q \times \rho_h}{f''N'' - 1 - N''(e) \times q \times \rho_h \times \rho_i \times \Delta} \tag{1.10}$$

From the second order condition, the denominator is negative and the model structure dictates $N'(e) \times q \times \rho_h > 0$, and so the partial derivative of arrests efforts with respect to costs from possible investigation is negative.

The second prediction that skill investment is decreasing in costs of investment follows from agent's problem in the first period. The agent chooses $k = 1$ if

$$U_2(k = 1) - U_2(k = 0) > \frac{c}{q} \tag{1.11}$$

1.C Crime Data construction

The Uniform Crime Reports (UCR) data are created from monthly self-reports by local agencies on crime and police activities. Previous literature (Chalfin and McCrary, 2018, Evans and Owens, 2007 and Mello, 2019) has noted record errors and extreme outliers in the UCR data. In addition, some agencies choose to report semi-annually or annually, and these choices are not explicitly stated in the data. I follow a data cleaning procedure similar to those used in the literature:

To identify agencies that do not report monthly, I check whether the total number of crimes jump to more than 5 at month 6 or 12 after a series of 0s in the previous 5 months of 0 crimes. If this pattern occurs, I replace all data reported by the agency during this period as missing.

Next, for each city, I fit the monthly time series of index crimes and arrests using a local linear regression with bandwidth two between 2013 and 2016. I then compute the absolute value of the percent difference between the actual and predicted values. To avoid dealing with zeros in the percent calculation, I add one to each variables when running the regressions. I recode the observation as missing if the difference exceeds a threshold.

The thresholds used are 97.5th percentiles of the within- population size group distributions of the percent differences. The population size group categories are the following: less than 2500, 2500-5000, 5000-10,000, 10,000-15,000, 15,000-25,000, 25,000-50,000, 50,000-100,000, 100,000-250,000, and greater than 250,000. Agencies that appear in multiple categories are placed in one that they appear most often.

Observations missing either due to non-reporting or outlier status are imputed using a combination of backwards/forwards fillings and linear interpolation. For example, if the first month of non-missing data appear in March 2013, then that value is imputed in January and February 2013. If a city has a missing data in May 2014, but April and June 2014 are non-missing, May 2014 is linearly interpolated. I choose to impute values so that regression results do not reflect composition changes. Finally, I winsorize crime and arrests per capita at the bottom and top 1% within each size group.

CHAPTER 2

USING PROMOTION INCENTIVES TO IMPROVE POLICE PERFORMANCE: EVIDENCE FROM THE CHICAGO POLICE DEPARTMENT

2.1 Introduction

Governments play a central function in economic growth. Economist in the past decade have systematically studied incentives that bureaucrats face in the public sector. Relative to private sector employees, it is very difficult to incentivize bureaucrats because in the public sector, often the pay is compressed, there are less objective measures of performance, and promotion in the civil service system are rigid (Finan et al., 2017).

In modern bureaucracies, the strict rule-based promotion system arose as a reform to previous systems based on patronage and favoritism. (Xu, 2018, Ornaghi, 2016) Understanding how to incentivize bureaucrat performance while preserving the benefits of the civil service system is a very important area of research (Bertrand et al., 2019).

I examine this question in the setting of the Chicago Police Department (CPD), where in 1998, the management changed the promotion system from one that is strictly based on rank scores in exams to a hybrid system that gave promotion to officers who scored high on the exam but also to those that were deemed by a evaluation board to have leadership potential. Specifically, I examine how promotion incentives in this system affect officers' performance and career choices.

I exploit an eligibility rule that expedited the first promotion opportunity by 7 years for an arbitrarily selected group of officers. In 1998 and 2006, promotion incentives were given to officers who had enough tenure by the evaluation time. Using a regression discontinuity design (RDD), I compare a group of officers who just made the cutoff for promotion eligibility to a similar control group who just missed the cutoff because they had slightly less tenure. I find that those who were eligible for promotion had large increase in actual promotion throughout their career (about 17% of career-long promotion rate for supervisory roles and about 37% for management roles). This substantial rise in promotion incentive reduced misconduct allegations, primarily personnel violations. I do not find effects for arrest performance or retention. Moreover, I find that greater promotion incentives lead officers to switch into positions that can boost their chance of promotion. This paper highlights that merit-based promotion opportunities can potentially be introduced in a rigid civil service system to improve outcomes.

I primarily relate to the literature that studies bureaucratic incentives and potential reforms to improve performance, especially in government agencies of authority where providing incentives are harder than other organizations such as health and education service providers. Various methods for incentivizing bureaucrats have been examined from performance pay, performance-based posting, autonomy, and non-financial incentives (e.g. Khan et al., 2016; Khan et al., 2019; Banerjee et al., 2012; Dal Bo et al., 2013; Deserranno, 2019). I add to these studies by studying promotion incentives based on an unspecified set of performance measures. Incentive schemes that focus on specific performance measures can introduce multi-tasking where the workers only focus on those incentivized portion of their work. When performance measures are not specified, bureaucracies may be able to avoid this problem. Moreover, evaluation by a merit board, instead of a single person, may be able to avoid problems associated with patronage (Xu, 2018).

In addition, I contribute to a recent literature that have examined the incentives and decision-making of police officers. Among different types of government agencies, enforcement agencies face particularly severe challenges because of multitasking problems (Prendergast, 2001, Shi, 2009, Ba, 2017, Rivera and Ba, 2019, Devi and Fryer Jr., 2020). Fryer Jr. (2018) explores racial differences in the police use of force, while Ba (2017) and Rivera and Ba (2019) study the interactions between civilian oversight and the use of force and complaints.

The remainder of the paper proceeds as follows. In Section 2.2, I provide institutional details about the CPD and its promotion system. Section 2.3 describes my data the personnel data and my empirical strategy. In Section 2.4, I present the results. I conclude in Section 2.5.

2.2 Background on the promotion system in the Chicago Police Department

2.2.1 The career of a police officer

As with many public sector organizations, police officers at the CPD remain with the organization for most of their working lives. For example, 80% of officers who joined the organization in 2000 are active in 2018. After basic training at the Police Academy and field training as a probationary police officer for 18 months, officers start their career patrolling in one of 22 districts. About 66% of the force work in the districts. As an officer gains more experience, she has opportunities to work in specialized units such as Gang Enforcement, Canine, Detective, or Information Services. These units are often more prestigious as they are generally regarded to pursue long-term crime control strategies rather than involving routine work and, in some cases, to have better working conditions. They usually require passing special examination processes.

Officers can also advance their careers by climbing the hierarchical ladder. After gaining experience they can move up to supervisory roles, as a sergeant, lieutenant, captain, and so on to the Superintendent. Among active officers in 2018, the vast majority (85%) have the lowest rank – police officer*. About 10% are sergeants, and the rest are lieutenants or above. In this paper, I only consider the promotion process from a police officer to a sergeant, as

∗. I will also use the term "police officer" to denote an individual whose career is in local law enforcement this is the majority of the promotions.

2.2.2 The promotion system

Until 1998, the CPD promoted officers based solely on exam scores. This method was disadvantageous toward minority officers as they tended to score low on the exams, and they struggled to be promoted to leadership roles in the CPD. The CPD was often mired in lawsuits as minority officers claimed that this method of promotion was discriminatory. The leadership also acknowledge the limitation of an exam-based promotion which failed to reward ambitious, capable officer with supervisory potential who may not be a strong test-taker. After an exam held in 1994, which was often involved in a lawsuit, the CPD started a two-track process for promotion: 70% of sergeants are to be promoted by rank order after achieving top scores in the sergeant exam, and the remaining 30% are to be promoted by "merit," after being nominated by a commanding officer and selected by the Merit Board.

A Merit Board consisting of five deputy chiefs and the Director of Human Resources evaluates the nominees. An officer can be considered for merit promotion only after a commanding officer, for example a district commander, nominates for consideration. The Merit Board then reviews the nominee's performance records, disciplinary histories, and complimentary records. The Superintendent makes the final selection among the candidates screened by the Merit Board.

For either of the two tracks for promotion, however, the officer must meet a few requirements to be eligible for promotion. Among other requirements for promotion eligibility, two are central in my empirical strategy. First, in order to be considered for either rank order or merit selection, an officer needs to have taken a qualifying exam, which she can take after 3.5 years of service. The exam is only offered infrequently as explained below. Second, an officer can only be promoted to sergeant after 6.5 years of service as a police officer. This requirement allows me to examine changes in performance and career decisions when an officer has greater promotion opportunities yet did not advance to the higher position.

Most nominations for merit promotion have occurred in specialized units rather than patrol districts as shown in Figure 2.1. Although all officers start as patrols in the city districts, they can apply for a role in specialized units (e.g. narcotics, gang, and SWAT). These units can have more varying work schedules than patrol districts but competitive as officers have realized they help gain promotion. Also there is relatively little difference in nomination rate among patrol districts. The more busy districts, where crime control activities happen more frequently, do not have higher nomination rates than less busy districts. I later examine officers' incentives to make career choices to increase their promotion potential.

2.3 Data and empirical strategy

2.3.1 Data

This study uses personnel data from the CPD that I obtained by submitting Freedom of Information Act (FOIA) requests and collecting data that other organizations have gathered through FOIA. The Invisible Institute maintains a public database of Chicago police personnel obtained through FOIA requests. This includes demographics and unit assignment of all sworn officers at the CPD from 1946 to 2018. The database includes complaints history which becomes reliable starting from 1991. At the incident level, I observe the date and type of allegations. Two most common types of allegations are excess policing and personnel violations. I aggregate the complaints at the officer-month level. I combine this with a database of arrests that I obtained through FOIA. This includes the arresting officers and the circumstances involving the incidents between 2001 and 2016. I consider all arrests as well as Type 1 arrests, which represent arrests for more serious crimes. I also collected promotion data that include all officers who are promoted between 2006 and 2016.

2.3.2 Eligibility for promotion

Promotion exams in the CPD are held infrequently. The most recent four exams were held in intervals of 6 – 7 years. After a qualifying exam is held, the CPD draws up a list of qualifying officers from which it promotes in each promotion review for the next few years. The next exam is held when the CPD determines that there is a need and budget to hold another one. It announces the schedule for the next exam a few months before the exam. In 2006, for example, the exam was announced two months before the exam.

Only officers who have served enough time (3.5 years) by the announcement month are eligible to take the exam. This means that only close to the new exam date officers find out whether they are eligible to take the exam only after having served a few years. I use this arbitrariness in the rules to measure the effects of being given more promotion opportunities. An officer who just makes the cutoff date for qualifying exam has a chance to be promoted 6 – 7 years earlier than her peer who joined the CPD a few months later and missed the cutoff.

In recent years promotion exams were held in 1998, 2006, 2013, and 2019. The main promotion exam that I use for this study is the one offered in January 2006 in which the last eligible class was those that entered the CPD in July 2002. Using this exam allows me to examine the main performance measure of arrests which is available from 2001. I supplement the 2006 exam with the 1998 exam, when I examine measures that go further back such as misconducts and resignation. I do not use the 2013 and 2019 exams because they do not give enough data to the right of the eligibility thresholds. Moreover, the sample of officers that would be in the left of the RD threshold for the 2013 exam entered the organization during the Great Recession, during which classes were sparsely spaced.

2.3.3 Regression discontinuity

My main empirical approach exploits the discrete jump in eligibility of promotion based on the appointed dates of officers. I attribute changes in outcome variables of interest to the additional promotion opportunities that are arbitrarily offered to otherwise similar officers.

In the standard RD specification, I estimate the following regression for officer i:

$$Outcome_i = \alpha + \beta_1 Eligible_i + \beta_2 HireDate_i + \beta_3 Eligible_i \times HireDate_i + \epsilon_i,$$

where $HireDate_i$ is the monthly running variable indicating the month of hire. The dummy variable $Eligible_i$ indicates whether the officer i is eligible to take promotion exams in 2006. β_1 is the main parameter of interest that captures the local discrete jump in the outcome variable.

The main outcome variables are performance as measured by misdemeanor allegations and arrests. I calculate the average monthly measure of these outcomes for each officer from their 42nd month to the 77th month. This start of the window is selected to be the 42nd month because officers whose tenure is at the 42nd month or later would know they are eligible for an exam, should there be one. This also applies to officers whose appointed date is near the eligibility threshold of the 2006 exam. Officers who entered the organization in July of 2002 became aware of their eligibility in January of 2006, which was their 42nd month of tenure, while those who entered in August of 2002 learned they were not eligible in their 42nd month. I set the ending window in the 77th month because officers cannot be promoted until their reach their 78th month of tenure. Ending the observation before the 78th month allows me to compare outcomes when the officers hold the same rank and have similar job responsibilities. Other outcomes of interest are resignation and unit assignment which allow me to examine career choices.

The main assumption for my empirical approach is that officers close to the eligibility threshold are randomly distributed across the threshold. This is plausible for two reasons. First, officers do not have a choice of precisely picking their class because they enter the police academy in a lottery order. Second, the exam and eligibility dates are announced only months before the exam, and precisely selecting class based on those dates is not possible. Nonetheless I carefully check the validity by testing covariate balance.

To obtain RDD estimates, I estimate the mean square error optimal bandwidths for main outcome variables following Calonico et al. (2014) and select the median bandwidth (36) to have a consistent sample across outcomes. I estimate local linear regressions with a triangular weight that puts more weights on points near the threshold.

In Table 2.1, I present results from running the main RD equation using predetermined covariates as the dependent variable. The covariates I test include demographic variables such as birth year, gender, race, as well performance in the first two years of service, such as average complaints and index arrests. I also test whether they were assigned in busy or slow districts. There is anecdotal evidence that top performers in the Police Academy are given the opportunity to pick their first district of assignment. Also officers with political influence within the organization can arrange for desirable assignments in safe districts. I define busy districts as the top 6 districts with highest per capita murder rates or and slow districts as the the bottom 6 lowest per capita murder rates. I find that almost all the variables have small magnitude and do not obtain statistical significance. In a total of 13 variables, I only find that one variable (female) obtains statistical significance at a 5% level. In all my regressions going forward, I control for basic demographic variables, including whether the officer is female, although the results do not change without covariates.

2.4 Results

2.4.1 Discontinuities in promotion chances

How did fast-tracking promotion affect promotion outcomes over the career? Figure 2.2 shows that the rules discontinuously varied promotion opportunities across the threshold. Because the data on promotion dates are available only from 2006, I cut the pre-threshold window shorter and drop classes that may have been promoted earlier than 2006. Those who were appointed in or before July of 2002, and who had the required 3.5 years of time in grade had more than 2 percentage points higher probability of being promoted to sergeants as early as their 7th year after entry into the CPD. Those who are not eligible did not have a chance of getting promoted and had to wait another 4 years until their first promotion opportunity. At the end of the sample period in 2018 (16 years after entry), there was more promotion among the initial ineligible but, there is still a gap in promotion rate by 3 percentage points. This had a knock-on effect on next promotion steps to being a manager (lieutenant or higher). I summarize the results in Table 2.2.

To put these gaps into perspective, I calculate the career-long likelihood of being promoted using officers who have resigned from 2000 to 2018. 18 percent of all officers resigned as supervisors and 6 percent of all officers finished their careers as managers. The 3 percentage points differences in promotion rates by the first 16 years translates to 17% of career-long promotion rate, and the 2 percentage differences in manager promotion rates is 37% of the career-long promotion rate.

Using the earlier 1998 exam, I can examine the difference in promotion outcomes over a longer period of time. After 24 years from entrance into the force, there is less pronounced visual gap in promotion outcomes exists, although the regression still points to a 2 percentage points gap. The gap is more clear for promotion to management roles, where the difference is 1.5 percentage points.

2.4.2 The effects of promotion opportunities on performance

I first present my results using RD figures. The main measures I use for performance are arrests and misconduct allegations. I measure those between 42 and 78 months of tenure, the time period in which eligible officers in my RD sample knew they were eligible but did not have enough tenure to be promoted yet.

Figure 2.3 shows the effects of getting more promotion incentives on arrests. There is not much visual difference across the threshold. A similar pattern holds for index crimes, which are more serious crimes. I use index arrests, which are those for serious crimes, instead of total arrests which include trivial misdemeanors. I use this measure because officers who make arrests for trivial misdemeanors may be acting out of poor judgment. Counter-intuitively, the number of arrests does not go up, and if anything, goes slightly down. For ease of interpretation, I also plot a local linear regression in the figures, fitted to the raw officer-level data separately on either side of the threshold. To test for the significance of the result, Table 2.3 shows the regression results on the dummy for being eligible for promotion. The reduction in index arrests do not obtain statistical significance.

Another important measure of performance is misconduct allegations. Misconduct complaint register was part of the file that the Merit Board has used to examine nominated candidates. In addition to the total number of complaints, I examine complaints related to excess policing (e.g. use of force and illegal search) as well as personnel violations. Because this data goes further than arrests data, I can also examine the 1998 exam as well as the 2006 exam. In Figure 2.4, I do not find discontinuity in total complaints across the threshold of the 2006 exam, but do observe a drop of 0.015 complaints, which represents a 12.5% decline from the control mean. When I examine major categories, personnel violations drive this difference. For the 1998 exam, I find a clear drop of 0.007 from the control mean of 0.03, which is a 23% drop. I also observe a similarly large drop for the 2006 exam sample. The visual evidence for excess policing is more muted, although the RD estimate indicates a drop of 8.5% from the control mean (Table 2.4). The more pronounced result for personnel violations may stem from the limited emphasis the CPD has put on civilian misconducts. The CPD, much like other major metropolitan police departments, have been mired in use of force scandals, and have undergone multiple police reforms in the recent decades.

2.4.3 The effects of promotion opportunities on career choices

The above results find that large promotion incentives do not improve arrest performance, yet do reduce misconducts, especially those related to personnel violations. I next examine how being eligible for more promotion opportunities affects career choice of officers. Promotion opportunities can have the potential to induce officers to remain with the organization, helping it retain talent. Figure 2.5, however, shows that being eligible for promotion hardly changed the resignation rate. For the 1998 exam, the difference in resignation rate actually flips, although the coefficient does not obtain statistical significance (Table 2.5). This absence of difference in resignation highlights the limited set of methods that public sector organizations can use in retaining employees.

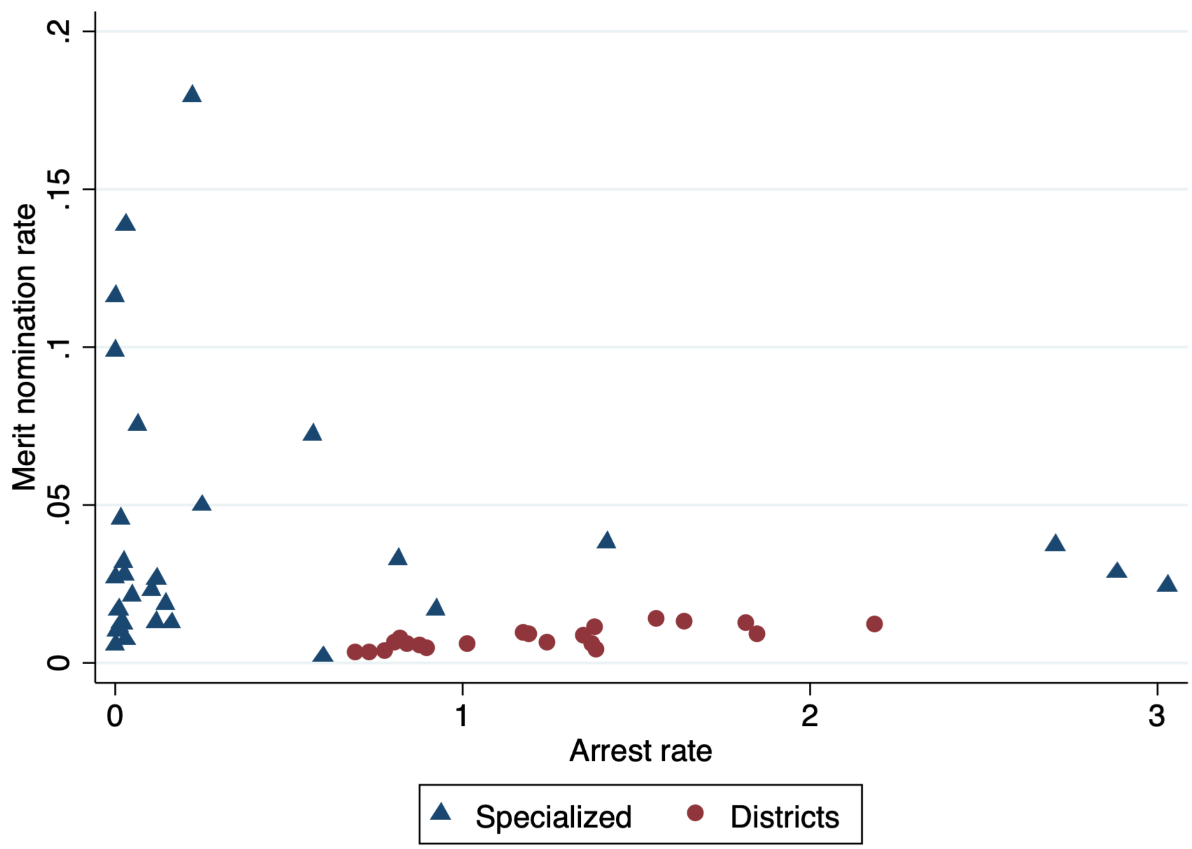
As described in Section 2.2.2, officers are more likely to receive nomination in special units than patrol districts, and faced similar likelihood of nomination regardless of whether officers worked in busy or slow districts. Figure **??** shows that, for the 2006 exam, the promotion incentive induced officers to work more in special units. As summarized in Table 2.5, the increase is 0.046, which is 42% increase from the control mean. Officers are less likely to work in busy districts, which I define as the districts in the top quartile in terms of arrest rates. Officers strategically positioned themselves for promotion opportunities. However, the visual differences of discontinuity vanish when I examine the 1998 exam. This is likely because the 1998 exam was the first promotion exam for which there was merit nomination, and the nomination pattern in different units was not established.

2.5 Conclusion

This paper studies the effects of promotion incentives on the performance and career choice using the setting of CPD where promotion decisions were evaluated by a board based on merit. Using RDD, I am able to obtain the effects of making some officers in CPD eligible for promotion while the control group was given such an opportunity only 7 years later. I find that the promotion incentives have led to large reductions in misconducts while I do not find evidence of changes in arrest performance or retention. I also find that they encouraged officers to invest in their career opportunities to get more promotion. This paper highlights the nuanced way that promotion incentives could be used to improve officer performance which may help improve upon rigid promotion systems that are based on exams or tenure. Also, agencies could use the incentive to encourage officers to invest in skills, such as use-of-force training, that may help both in their promotion chances and performance.

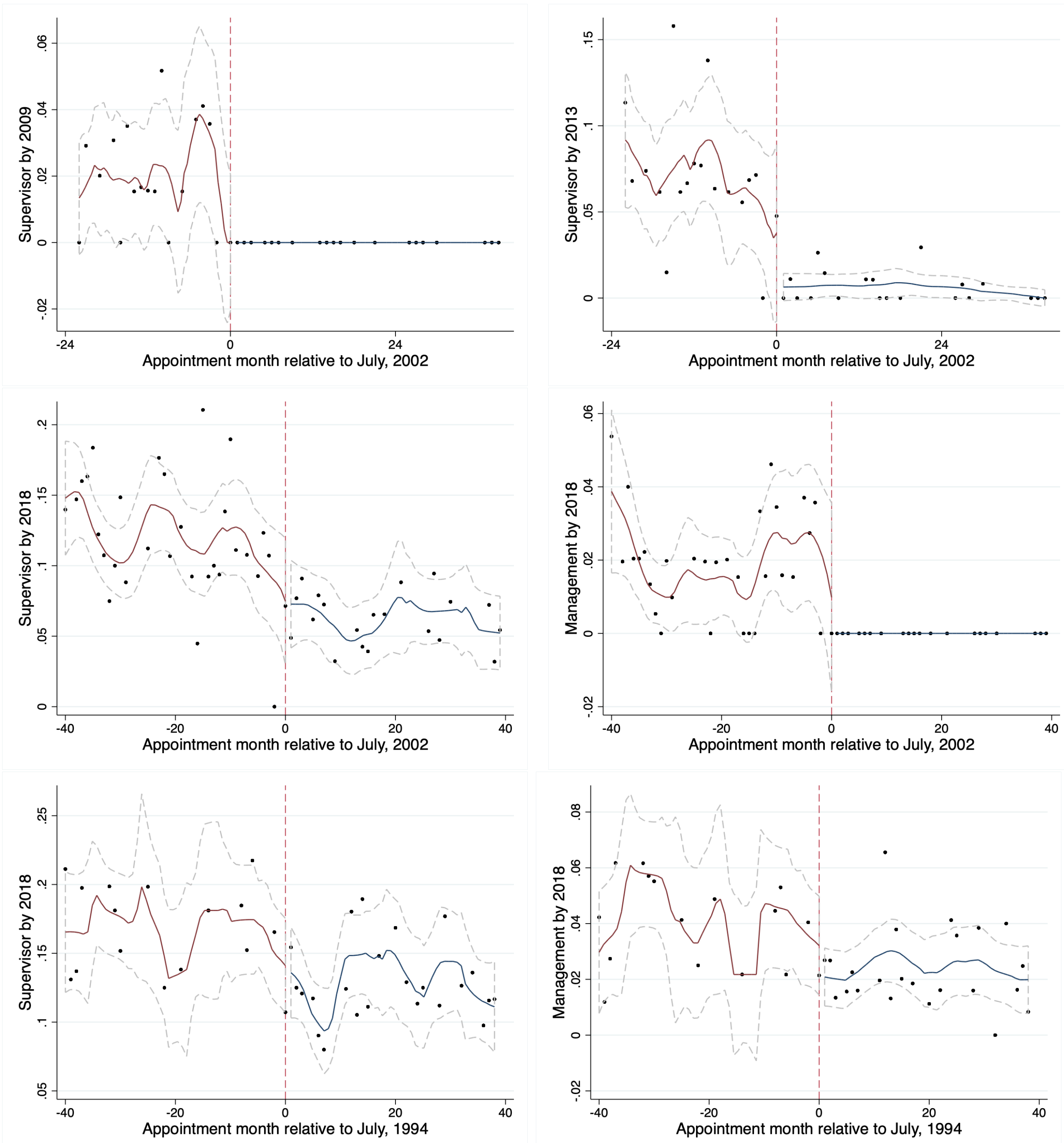
Public sector in developed countries are different from developing countries in important ways. My findings highlight ways in which promotion incentives work differently from developing countries and that they affect individuals in a different way. In the personnel policies of police, my findings suggest that promotion incentives should be tailored for officers' abilities.

Although this study has found that there can be benefits to the merit promotion system, it has recently been stopped after existing for more than 20 years. The main complaint among the rank-and-file was that the nomination process was not transparent. The management has tried recently to improve trust by releasing data on nomination and training nominators on how to more objectively assess subordinates. Future work should examine how to further improve upon a merit-based promotion system and how to help key stakeholders embrace such a reform.



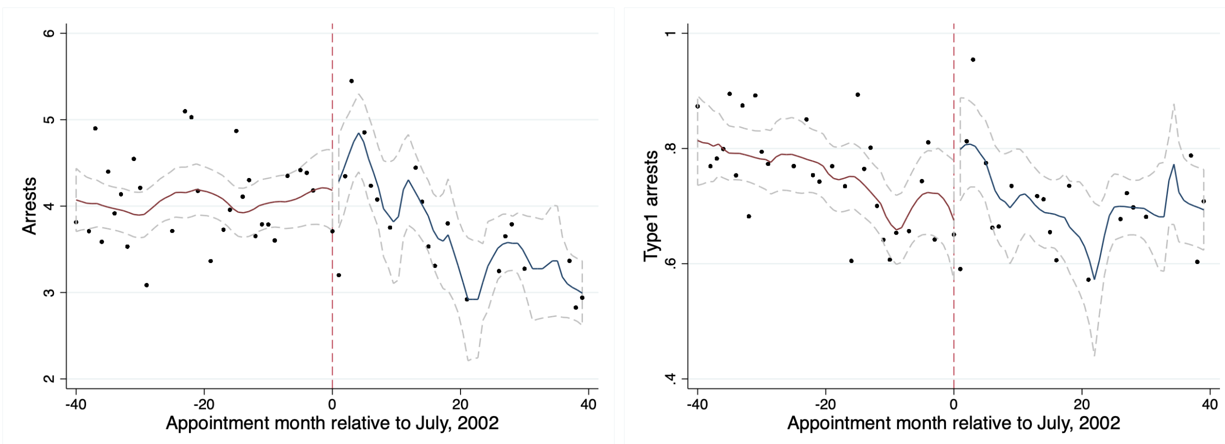
This figure shows average merit nominations by units from 2006 to 2017. Most officers are in patrol districts, and others serve in specialized units such as SWAT and narcotics units.

Figure 2.1: Merit promotion rate in different units



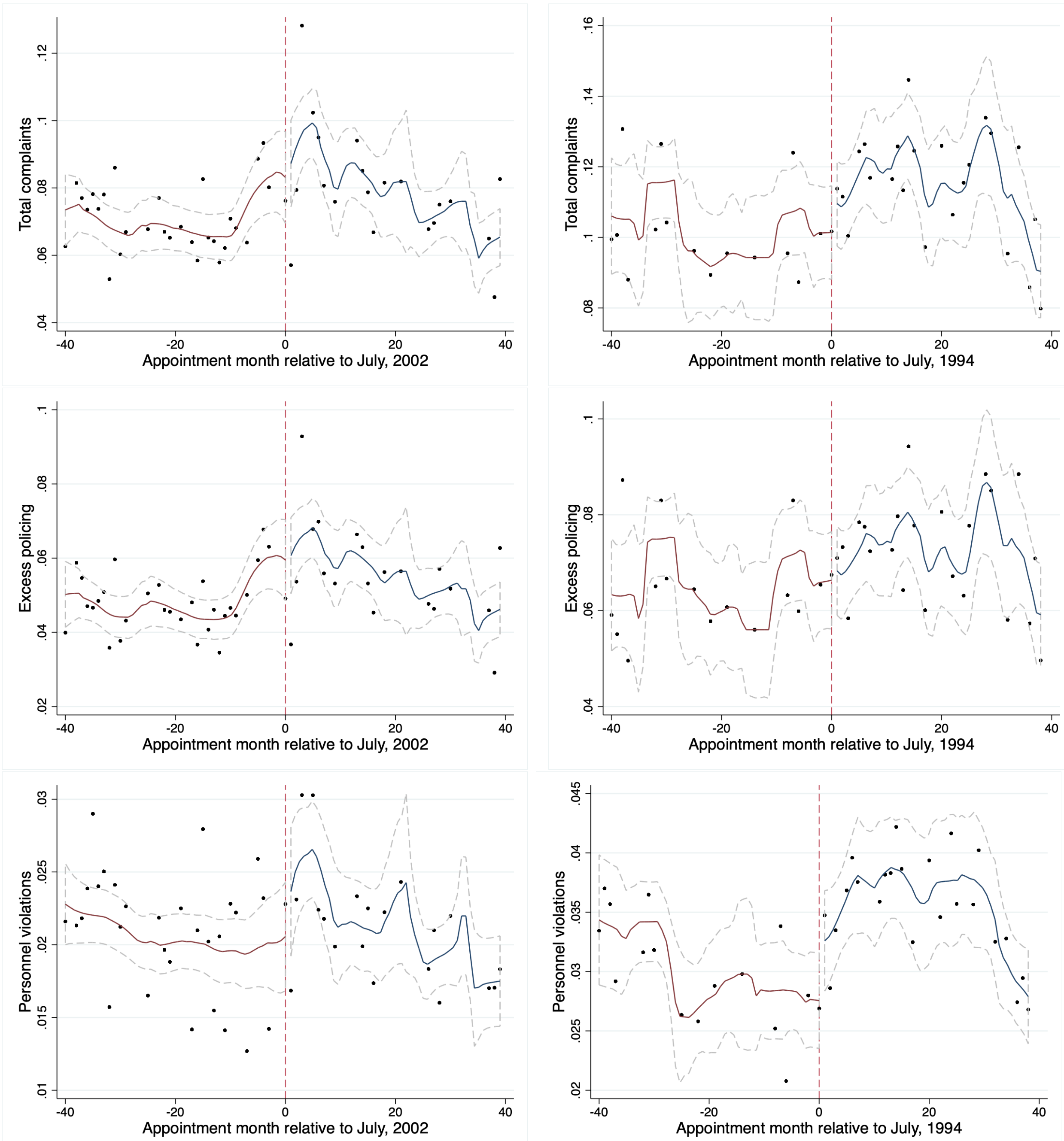
This figure plots the rates of eventual promotion for officers who entered in different classes across the RD thresholds. In the first four rows, I examine the sample for the 2006 exam, where the relevant threshold is appointment month of July, 2002 (officers who entered by that time would have enough tenure to be eligible to take the exam). I examine promotion rates by year 2009 and 2013 in the first row. I also examine promotion rates to supervisory and management roles in the second row by the end of my data period. In the last row, I examine the sample for the 1998 exam, which consists of officers who are at the threshold of being eligible for the exam (appointed in July, 1994) and those in classes around that threshold. Standard errors are clustered at the class level.

Figure 2.2: Promotion chance around eligibility threshold



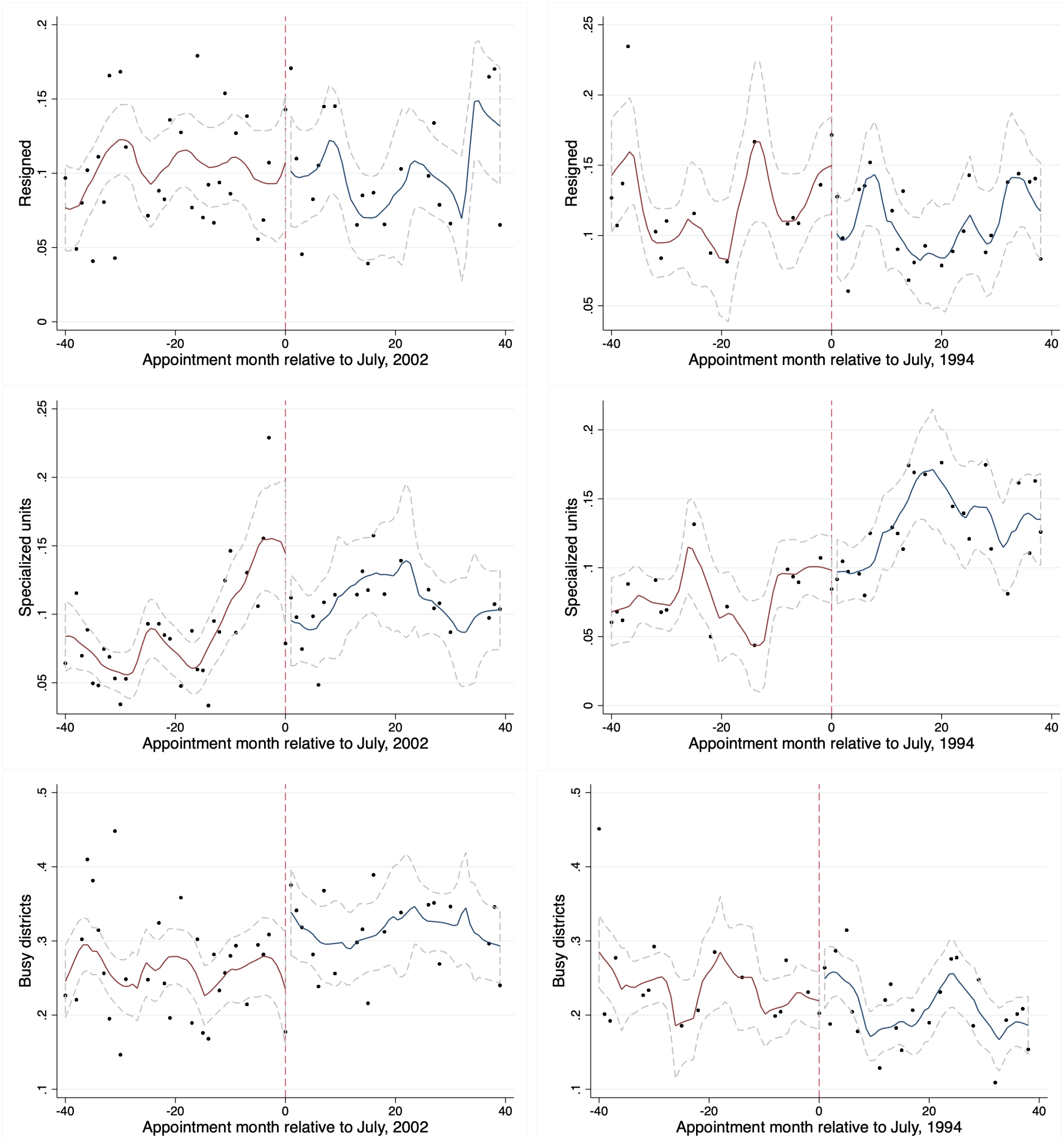
This figure plots average monthly arrest rates from 42nd month of tenure to 77th month of tenure around the RD threshold. During this period, officers know their eligibility status but cannot be promoted yet. I examine the sample for the 2006 exam, where the relevant threshold is appointment month of July, 2002 (officers who entered by that time would have enough tenure to be eligible to take the exam). Standard errors are clustered at the class level.

Figure 2.3: Arrest performance around eligibility threshold



This figure plots average monthly misconduct allegations from 42nd month of tenure to 77th month of tenure around the RD threshold. During this period, officers know their eligibility status but cannot be promoted yet. On the left-hand panels, I examine the sample for the 2006 exam, where the relevant threshold is appointment month of July, 2002 (officers who entered by that time would have enough tenure to be eligible to take the exam). On the right-hand side, I examine the sample for the 1998 exam, which consists of officers who are at the threshold of being eligible for the exam (appointed in July, 1994) and those in classes around that threshold. The first rows analyze total complaints, whereas the second and third rows examine misconducts related to excessive policing and personnel violations. Standard errors are clustered at the class level.

Figure 2.4: Misconduct allegations around eligibility threshold



This figure plots average monthly likelihood of serving in certain police units from 42nd month of tenure to 77th month of tenure around the RD threshold. During this period, officers know their eligibility status but cannot be promoted yet. On the left-hand panels, I examine the sample for the 2006 exam, where the relevant threshold is appointment month of July, 2002 (officers who entered by that time would have enough tenure to be eligible to take the exam). On the right-hand side, I examine the sample for the 1998 exam, which consists of officers who are at the threshold of being eligible for the exam (appointed in July, 1994) and those in classes around that threshold. The first rows analyze total complaints, whereas the second and third rows examine misconducts related to excessive policing and personnel violations. Standard errors are clustered at the class level.

Figure 2.5: Career choices around eligibility threshold

Table 2.1: Covariance balance

	(1) Birth year	(2) Black	(3) White	(4) Hispanic	(5) Female	(6) Busy	(7) Slow	(8) Arrests	(9) Type 1	(10) Excess	(11) Personnel	(12) Off duty	(13) Total complaints
Eligible	0.595	0.001	0.026	-0.022	0.115**	-0.080	0.056	-0.117	-0.037	-0.002	-0.001	-0.001	-0.004
	(0.565)	(0.036)	(0.044)	(0.024)	(0.053)	(0.063)	(0.037)	(0.215)	(0.110)	(0.003)	(0.003)	(0.000)	(0.005)
Dep Var Mean	1972.50	0.23	0.50	0.23	0.24	0.38	0.16	3.23	1.24	0.02	0.03	0.00	0.05
N	3,541	3,541	3,541	3,541	3,541	3,245	3,245	1,628	1,628	3,541	3,541	3,541	3,541

Notes: This table present results from running the main RD equation using predetermined covariates as the dependent variable at the optimal bandwidth of 36, except for arrests performance, where because of the shorter data availability I use 18. Standard errors are clustered at the class level.

Table 2.2: Discontinuity in promotion opportunity

	2006 exam				1998 exam	
	(1)	(2)	(3)	(4)	(5)	(6)
	Promoted: 2009	Promoted: 2013	Promoted: 2018	Manager: 2018	Promoted: 2018	Manager: 2018
Eligible	0.023***	0.046***	0.031**	0.022***	0.026	0.015**
	(0.008)	(0.009)	(0.014)	(0.006)	(0.017)	(0.006)
Dep Var Mean	0.01	0.04	0.10	0.01	0.15	0.03
N	2,593	2,593	3,541	3,541	7,676	7,676

Notes: Table shows the RD estimates in promotion chances by being promoted to a supervisory role and a managerial position by certain years. On the left side, I use the 2006 exam sample and for the last two columns, I use the 1998 exam sample. See Notes on Figure 2.2 for details on the samples. I use the optimal bandwidth of 36 and include basic demographics as controls. I cluster standard errors on class.

Table 2.3: Discontinuity in arrest performance

	2006 exam		1998 exam	
	(1)	(2)	(3)	(4)
	Arrests	Type 1	Arrests	Type 1
Eligible	-0.371	-0.070	-0.148	-0.021
	(0.297)	(0.060)	(0.188)	(0.037)
Control mean	3.88	0.71	3.32	0.80
N	3,541	3,541	5,950	5,950

Notes: Table shows the RD estimates in arrest performance. See Notes on Figure 2.3 for details on the sample. I use the optimal bandwidth of 36 and include basic demographics as controls. I cluster standard errors on class.

Table 2.4: Discontinuity in misconducts

	2006 exam			1998 exam		
	(1)	(2)	(3)	(4)	(5)	(6)
	Misconduct	Excess	Personnel	Misconduct	Excess	Personnel
Eligible	-0.013	-0.007	-0.005*	-0.015***	-0.006*	-0.007***
	(0.009)	(0.007)	(0.003)	(0.006)	(0.004)	(0.002)
N	3541	3541	3541	7676	7676	7676
Control mean	0.08	0.06	0.02	0.12	0.07	0.04

Notes: Table shows the RD estimates in misconduct allegations by being promoted to a supervisory role and a managerial position by certain years. On the left side, I use the 2006 exam sample and for the last two columns, I use the 1998 exam sample. See Notes on Figure 2.4 for details on the samples. I use the optimal bandwidth of 36 and include basic demographics as controls. I cluster standard errors on class.

Table 2.5: Discontinuity in career choices

	2006 exam			1998 exam		
	(1)	(2)	(3)	(4)	(5)	(6)
	Resigned	Special	Busy	Resigned	Special	Busy
Eligible	-0.007	0.046**	-0.052*	0.018	0.017	-0.022
	(0.020)	(0.022)	(0.027)	(0.016)	(0.011)	(0.018)
N	3541	3541	3541	7676	7676	7676
Control mean	0.10	0.11	0.31	0.11	0.13	0.21

Notes: Table shows the RD estimates in misconduct allegations by being promoted to a supervisory role and a managerial position by certain years. On the left side, I use the 2006 exam sample and for the last two columns, I use the 1998 exam sample. See Notes on Figure 2.5 for details on the samples. I use the optimal bandwidth of 36 and include basic demographics as controls. I cluster standard errors on class.

BIBLIOGRAPHY

Abraham, S., & Sun, L. (2018). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. *SSRN Electronic Journal*.

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J., & Zilibotti, F. (2007). Technology, information, and the decentralization of the firm. *Quarterly Journal of Economics*, *122*(4), 1759–1799.

Allcott, H. (2015). Site Selection Bias in Program Evaluation. *The Quarterly Journal of Economics*, 1117–1165.

Annan-Phan, S., & Ba, B. A. (2019). *Hot Temperatures, Aggression, and Death at the Hands of the Police: Evidence from the US*.

Aral, S., Brynjolfsson, E., & Van Alstyne, M. (2007). *Information, Technology and Information Worker Productivity: Task Level Evidence* (No. *9*).

Ariel, B. (2016). Police Body Cameras in Large Police Departments. *Journal of Criminal Law and Criminology*, *106*(4), 729–768.

Ariel, B., Farrar, W. A., & Sutherland, A. (2015). The Effect of Police Body-Worn Cameras on Use of Force and Citizens' Complaints Against the Police: A Randomized Controlled Trial. *Journal of Quantitative Criminology*, *31*(3), 509–535.

Ariel, B., Sutherland, A., Henstock, D., Young, J., Drover, P., Sykes, J., Megicks, S., & Henderson, R. (2018). Paradoxical effects of self-awareness of being observed: testing the effect of police body-worn cameras on assaults and aggression against officers. *Journal of Experimental Criminology*, *14*(1), 19–47.

Athey, S., & Stern, S. (2002). The Impact of Information Technology on Emergency Health Care Outcomes. *RAND Journal of Economics*, *33*(3), 399–432.

Ba, B. A. (2017). Going the Extra Mile : the Cost of Complaint Filing, Accountability, and Law Enforcement Outcomes in Chicago.

Banerjee, A. V., Chattopadhyay, R., Duflo, E., Keniston, D., & Singh, N. (2012). *Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy and Training*.

Banks, D., Blanton, C., Burch, A., Couzens, L., Cribb, D., & Planty, M. (2015). Arrest-Related Deaths Program: Data Quality Profile. *Bureau of Justice Statistics*, (March), NCJ 248544.

Banks, D., Rddule, P., Kennedy, E., & Planty, M. G. (2016). *Arrest-related Deaths Program Redesign Study, 2015-16: Preliminary Findings* (tech. rep. December).

Bernstein, E. S. (2012). The transparency paradox: A role for privacy in organizational learning and operational control. *Administrative Science Quarterly*, *57*(2), 181–216.

Bertrand, M., Burgess, R., Chawla, A., & Xu, G. (2019). The Glittering Prizes : Career Incentives and Bureaucrat Performance. *Review of Economic Studies*, 1–68.

Borusyak, K., & Jaravel, X. (2018). Revisiting Event Study Designs. *SSRN Electronic Journal*, 1–25.

Braga, A., Coldren Jr., J. R., Sousa, W., Rodriguez, D., & Alper, O. (2017). *The benefits of body-worn cameras: New findings from a randomized controlled trial at the Las Vegas Metropolitan Police Department* (tech. rep.).

Braga, A. A., Sousa, W. H., Coldren, J. R., & Rodriguez, D. (2018). The effects of body-worn cameras on police activity and police-citizen encounters: A randomized controlled trial. *Journal of Criminal Law and Criminology*, *108*(3), 511–538.

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-level Evidence. *The Quarterly Journal of Economics*, (February).

Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, (40), 1–31.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica*, *82*(6), 2295–2326.

Chalfin, A., & McCrary, J. (2018). Are U.S cities underpoliced? Theory and evidence. *Review of Economics and Statistics*, *96*(4), 710–728.

Dal Bo, E., Finan, F., & Rossi, M. (2013). Strengthening State Capabilities: the Role of Financial Incentives in the Call to Public Service. *Quarterly Journal of Economics*, 1169–1218.

De Chaisemartin, C., & D'Haultfoeuille, X. (2018). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(17), 2964–2996.

Deserranno, E. (2019). Financial incentives as signals: Experimental evidence from the recruitment of village promoters in Uganda. *American Economic Journal: Applied Economics*, *11*(1), 277–317.

Deshpande, M., & Li, Y. (2019). Who Is Screened Out? Application Costs and the Targeting of Disability Programs. *American Economic Journal: Economic Policy*, *11*(4), 213–248.

Devi, T., & Fryer Jr., R. G. (2020). Policing the Police: The Impact of "Pattern-or-Practice" Investigations on Crime. *National Bureau of Economic Research (NBER)*, 63.

Doleac, J. L. (2017). The effects of DNA databases on crime. *American Economic Journal: Applied Economics*, *9*(1), 165–201.

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

Ellis, T., Jenkins, C., & Smith, P. (2015). *Evaluation of the Introudciton of Personal Issue Body Worn Video Cameras (Operation Hyperion) on the Isle of Wight: Final Report to Hampshire Constabulary* (tech. rep.).

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

Finan, F., Olken, B. A., & Pande, R. (2017). The Personnel Economics of the State. *Handbook of Field Experiments*, (December).

Finch, B. K., Beck, A., Burghart, D. B., Johnson, R., Klinger, D., & Thomas, K. (2019). "Using Crowd-Sourced Data to Explore Police-Related-Deaths in the United States (2010-2017): The Case of Fatal Encounters. *Open Health Data*, *6*(1), 1.

Fryer Jr., R. G. (2018). An Empirical Analysis of Racial Differences in Police Use of Force. *Journal of Political Economy*, 63.

Garicano, L., & Heaton, P. (2010). Information Technology , Organization , and Productivity in the Public Sector : Evidence from Police Departments. *Journal of Labor Economics*, *28*(1).

Goodman-Bacon, A. (2019). Difference-in-differences with Variation in Treatment Timing. *Working paper*, (July).

Harris, D. A. (2010). *Picture This: Body Worn Video Devices ("Head Cams") as Tools for Ensuring Fourth Amendment Compliance by Police*.

Headley, A. M., Guerette, R. T., & Shariati, A. (2017). A field experiment of the impact of body-worn cameras (BWCs) on police officer behavior and perceptions. *Journal of Criminal Justice*, *53*(October), 102–109.

Hedberg, E. C., Katz, C. M., & Choate, D. E. (2017). Body-Worn Cameras and Citizen Interactions with Police Officers: Estimating Plausible Effects Given Varying Compliance Levels. *Justice Quarterly*, *34*(4), 627–651.

Hubbard, T. N. (2003). Information, decisions, and productivity: On-board computers and capacity utilization in trucking. *American Economic Review*, *93*(4), 1328–1353.

Hyland, S. S. (2018). *Body-worn cameras in law enforcement agencies, 2016* (tech. rep. November).

Jennings, W. G., Lynch, M. D., & Fridell, L. A. (2015). Evaluating the impact of police officer body-worn cameras (BWCs) on response-to-resistance and serious external complaints: Evidence from the Orlando police department (OPD) experience utilizing a randomized controlled experiment. *Journal of Criminal Justice*, *43*(6), 480–486.

Khan, A. Q., Khwaja, A. I., & Olken, B. A. (2016). Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors. *The Quarterly Journal of Economics*, 219–271.

Khan, A. Q., Khwaja, A. I., & Olken, B. A. (2019). Making Moves Matter : Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings. *American Economic Review*, *109*(1), 237–270.

Klinger, D. A. (2012). On the problems and promise of research on lethal police violence: A research note. *Homicide Studies*, *16*(1), 78–96.

Levitt, S. D. (1997). Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime. *American Economic Review*, *87*(3), 270–290.

Lum, C., Stoltz, M., Koper, C. S., & Scherer, J. A. (2019). Research on body-worn cameras: What we know, what we need to know. *Criminology and Public Policy*, *18*(1), 93–118.

Maskaly, J., Donner, C., Jennings, W. G., Ariel, B., & Sutherland, A. (2017). The effects of body-worn cameras (BWCs) on police and citizen outcomes: A state-of-the-art review. *Policing*, *40*(4), 672–688.

McCrary, J. (2007). The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police. *The American Economic Review*, *662*, 660–662.

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

Miller, A. R., & Segal, C. (2019). Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence Escalation. *Review of Economic Studies*, (February).

ODS Consulting. (2011). *Body worn video projects in Paisley and Aberdeen: Self Evaluation* (tech. rep. July).

Ornaghi, A. (2016). Civil Service Reforms: Evidence from U.S. Police Departments. *Unpublished Manuscript*, 1–60.

Peterson, B. E., Yu, L., La Vigne, N., & Lawrence, D. S. (2018). *The Milwaukee Police Department's Body-Worn Camera Program* (tech. rep.).

Pierce, L., Snow, D. C., & McAfee, A. (2015). Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science*, *61*(10), 2299–2319.

Police and Crime Standards Directorate. (2007). *Guidance for the Body-Worn Video Devices* (tech. rep. July).

Police Executive Research Forum. (2014). *Implementing a Body-Worn Camera Program: Recommendations and Lessons Learned*.

Police Executive Research Forum. (2017). Research on the Impact of Technology on Policing Strategy in the 21st Century. *Report for National Institute of Justice*
Early adopters used it for specialized positions.

Prendergast, C. (2001). Selection and Oversight in the Public Sector, with the Los Angeles Police Department as an Example. *NBER Working paper series*, (January), 43.

Prendergast, C. (2020). *"Drive and Wave": The Response to LAPD Police Reforms After Rampart*.

Ready, J. T., & Young, J. T. (2015). The impact of on-officer video cameras on police citizen contacts: findings from a controlled experiment in Mesa, AZ. *Journal of Experimental Criminology*, *11*(3), 445–458.

Rivera, R. G., & Ba, B. A. (2019). *The Effect of Police Oversight on Crime and Allegations of Misconduct: Evidence from Chicago*.

Rozema, K., & Schanzenbach, M. (2019). Good Cop, Bad Cop: Using Civilian Allegations to Predict Police Misconduct. *American Economic Journal: Economic Policy*.

Santos Silva, J. M., & Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, *88*(4), 641–658.

Santos Silva, J. M., & Tenreyro, S. (2010). On the existence of the maximum likelihood estimates in Poisson regression. *Economics Letters*, *107*(2), 310–312.

Schwartz, G. L., & Jahn, J. L. (2020). Mapping fatal police violence across U.S. metropolitan areas: Overall rates and racial/ethnic inequities, 2013-2017. *PLoS ONE*, *15*(6 June), 2013–2017.

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

Sun, L., & Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

White, M. D., Todak, N., & Gaub, J. E. (2017). Assessing citizen perceptions of body-worn cameras after encounters with police. *Policing*, *40*(4), 689–703.

Xu, G. (2018). The Costs of Patronage : Evidence from the British Empire. *The American Economic Review*, *108*(11), 3170–3198.

Yokum, D., Ravishankar, A., & Coppock, A. (2019). A randomized control trial evaluating the effects of police body-worn cameras. *PNAS*, (26).