CAGE

Monitoring Technology: The Impact of Body-Worn Cameras on Citizen-Police Interactions

CAGE working paper no. 581

September 2021
(Revised September 2023)

Daniel AC Barbosa
Thiemo Fetzer
Caterina Soto-Vieira
Pedro CL Souza

WARWICK
THE UNIVERSITY OF WARWICK

UKRI Economic and Social Research Council

Monitoring Technology: The Impact of Body-Worn Cameras on Citizen-Police Interactions*

Daniel AC Barbosa   Thiemo Fetzer   Caterina Soto-Vieira
Pedro CL Souza

December 2023

Abstract

We provide experimental evidence that using body-worn cameras (BWCs) for police monitoring improves police-citizen interactions. Dispatches with BWCs show a 61.2% decrease in police use of force and a 47.0% reduction in negative interactions, including handcuff use and arrests. The use of BWCs also improves the quality of officers' record from the dispatches. The rate of incomplete reports dropped by 5.9%, which is accompanied by a 69% increase in the notification of domestic violence. We explore various mechanisms that explain *why* BWCs work and show that the results are consistent with the police changing their behavior in the presence of cameras. Our results stand in contrast with previous experimental literature which used coarser designs and indicated muted or null body-worn camera effects on use of force. Replicating those designs, our data also finds attenuated effects. Overall, our results show that the use of BWCs de-escalates conflicts.

Keywords: police, use of force, technology, field experiment

JEL Classification: C93, D73, D74

*A previous version of this paper was circulated under the title "De-escalation technology: the impact of body-worn cameras on citizen-police interactions." Barbosa is at the University of Oxford. Fetzer is at the University of Warwick. Soto-Vieira is at the London School of Economics. Souza is at Queen Mary University. We thank the Military Police of Santa Catarina and the Igarapé Institute, and especially Cel Tasca, Cel Araújo, Major Pablo, Major Vieira, Emile Badran, and Bárbara Silva, for their outstanding collaboration and support. We thank Oriana Bandiera, Imran Rasul, Marco Manacorda, Robin Burgess, Joana Monteiro, and Cel Cabanas from the Military Police of São Paulo for attentive suggestions and comments in earlier versions of this work. This project was funded through the EGAP Metaketa IV initiative, and it was pre-registered as part of the EGAP Metaketa with registration 20190411AB. The project was registered in the AEA Registry as AEARCTR-0007785.

1 Introduction

The rise of the modern nation-state is linked to the state gradually expanding its reach and asserting its monopoly of violence (Weber, 1946). One of the most visible forms of this process is the expansion and professionalization of police forces that have the right to exercise use of force legitimately. A well-resourced and professionally run police service provides public goods by safeguarding fundamental rights, thus enabling societies to prosper (Besley and Persson, 2010). Yet, the legitimacy and public confidence in the police is under strain worldwide following widespread allegations of excessive use of force.[1]

Police body-worn cameras (henceforth, BWCs) have been hailed as a technological solution to increase scrutiny and oversight of the police. By recording encounters that police officers have with citizens[2] when carrying out their duties, these devices create a public record of interactions that may bring transparency and accountability to the actions of the police. Consequently, this technology has significant potential to safeguard public trust in the police and the state, and as of 2016 had already been adopted by at least 60% of police departments in the United States (Hyland, 2018).

In this study, we provide experimental evidence demonstrating that BWCs used by police significantly enhance the quality of interactions between police officers and citizens. Our first set of results pertain to what officers report from the dispatches.[3] We can address this because we have access to the internal logs of dispatches by police officers, which are not public and are only used as internal registers of police activity. Our findings indicate that dispatches involving cameras were 2.77 percentage points less likely to lack essential descriptive information when submitted to police records. Relative to the mean of control, where events without accurate description represent the majority of internal reports, this equates

[1] See the *New York Times* (2020), Confidence in police is at a record low, Gallup survey finds, August 12, 2020, https://www.nytimes.com/2020/08/12/us/gallup-poll-police.html, accessed 10.08.2021.

[2] In our study, we use the term ”citizens” to describe any individual who is not actively serving in the police force. This usage is a slight abuse of the term because those interacting with the police may not necessarily be Brazilian citizens.

[3] Throughout the paper, we also interchangeably refer to a dispatch as an ”event.”

to a decrease of 5.9% in underreporting. This translates into an increase in the reporting of other crimes, including a substantive increase in reporting of domestic violence. We find that the presence of cameras induce a significant 1.14 percentage point increase in the reporting of domestic violence incidents, which accounts for a 69.2% increase relative to the control mean. We further find that it becomes 9.5% more likely that the dispatch is referred to investigative bodies, and police reports, on average, are 20.1% more likely to include victims. These effects can be also be a downstream consequence of the increase in the incidence of domestic violence reporting. We expand on this in Section 3.2. Overall, we interpret that officers are more diligent in reporting when in the presence of cameras, especially in cases where video footage can be used as evidence to prosecute. Furthermore, accurate reporting by police officers has been demonstrated to improve the rate at which crimes are cleared (Blanes i Vidal and Kirchmaier, 2018).

Our design introduced random variation both on *whether* a BWC was present on a dispatch and on *who* carried it. Hence, it allows us to study whether the characteristics of the officer carrying it mattered. We find evidence of stronger reporting effects and increased compliance with the police's BWC standard operating procedures if the officer wearing the camera is relatively junior. There are several possible explanations for this finding. It is conceivable that junior officers are more adept at using a new technology. It may also suggest that juniors' dynamic incentives and career concerns may be important factors driving their effect and that low-rank officers with BWCs monitor their higher-ranked peers.[4] Yet, because the salience of the cameras to the public is held fixed, any of those mechanisms mostly support the idea that the police are changing their conduct, and not so much the citizens, when in presence of the cameras.

We then turn our attention to the effects of BWCs on use of force. We find that treated dispatches saw a decline in the likelihood of use of force by 61.2%. A Negative Interaction Index following Anderson (2008)—which also combines charges of contempt, disobedience or citizen resistance, and use of handcuffs or

[4]In other contexts, peer monitoring in the workplace has been shown to alleviate principal-agent problems. See Bandiera et al. (2009, 2010), Mas and Moretti (2009), and Ashraf and Bandiera (2018) for a literature review.

arrests—, was reduced by 47%.

Consistent with the results on reporting, we find that the reduction of use of force is higher when junior officers are wearing the cameras. We then dig deeper on other mechanisms that are specifically relevant to use of force. We find that treatment effects are primarily concentrated in events that, *prior to police being dispatched*, were classified as relatively *low risk* by virtue of there being either no weapons reported on the scene, there being no injuries, or there being any material risk of general unrest as judged by the police. This suggests that cameras affect the situation dynamic by *preventing* the escalation of tension that would counterfactually unfold. We also find that BWCs have larger effects in areas with a higher baseline use of force. This suggests that cameras may have the highest benefits in places where police-citizen interactions are relatively strained at the baseline.

Lastly, we document that our estimates are robust to several alternative confounding mechanisms. First, we test if officers with BWCs self-select into specific events when wearing cameras. As a virtue of the implementation, dispatch operators were blind to the treatment status of dispatch units, limiting the potential for such a selection to happen. Yet, we confirm the absence of endogenous sorting through several empirical tests. More specifically, we show that treatment officers are not more likely to patrol less risky or wealthier areas, that they are not less likely to initiate the dispatches, and that they do not take longer to respond to the calls. Second, we find similar results in alternative estimation samples. Third, we re-evaluate the effects of BWCs using only observational variation in a differences-in-differences setting and data from precincts that did not participate in the intervention. Again, we find similar effect sizes.

This paper is hardly the first one to study BWCs (see Lum et al. (2020) and Williams et al. (2021) for reviews of the rather mixed evidence to date). Yet, unlike much of the existing literature, we implemented an exceptionally granular field experiment designed to seed a random allocation of body-worn cameras at the police *dispatch* level. We embedded an experiment that induced experimental variation both at the officer and the shift levels with the aim of generating random variation at the dispatch-level. Contrary to past studies, we view the dispatch-level data as the ”natural” unit of analysis both for treatment delivery as well as outcome measurement, as it is the level at which citizens and police interactions unfold and the use of force or its (de)escalation may occur.

We contribute to the growing literature on the effects of BWCs in a number of ways. First, we produce robust results showing that body-worn cameras reduce use of force, and we uncover the mechanisms as to why that happens. The previous literature, meanwhile, has produced mixed results. Some papers, including meta-studies (Yokum et al., 2019; Lum et al., 2020; Williams et al., 2021) found BWCs to have little or no effects, while other studies showed that such devices are effective in curbing the use of force (for example, Braga et al., 2018; Kim, 2020; Ferrazares, 2023; Monteiro et al., 2022). We shed light on why the literature has found diverging results to date. Naturally, differences could arise because of the different contexts in which the studies were conducted. This paper is among the first to provide evidence of BWCs' effectiveness in the context of a lower-income country in which citizen and police relations may structurally benefit more from BWCs (vis-a-vis the US and the UK, which have been almost exclusively the focus of the existing work). One notable exception is Monteiro et al. (2022), who showed that BWCs reduced police lethal use of force using the staggered adoption across police precincts in the municipality of São Paulo.[5]

We show that a likely explanation for some existing studies and meta-analyses not identifying treatment effects is due to the research designs and, in particular, to the outcome measurement and empirical evaluation strategies adopted. Our research design nests a broad class of commonly used evaluation strategies or outcome measurement approaches that were employed across experimental BWC studies. This allows us to replicate our study at coarser levels of analysis or when employing different empirical strategies for evaluation. We find that the estimated BWC treatment effects attenuate when mimicking the coarser evaluation approaches commonly used in the literature. The exceptionally granular data used in this study enables us to document that contamination, and noise— introduced in the outcome measurement when moving from the event level to coarser designs— are the likely culprit. Our paper makes a contribution by helping

[5] Magaloni (2019) studied a BWC randomized controlled trial in a neighborhood of Rio de Janeiro, Brazil. Overall, they noted very low compliance and little camera footage being produced.

us understand why some of the existing research may have struggled to identify treatment effects— and this is due to *study design*.

We further contribute to the literature by not only showing that BWCs work but also characterizing the mechanisms for why they work. Our randomization design and the granular dispatch data enables us to examine important heterogeneities which have not been explored by the literature. This also marks an important difference between our paper and the recent literature which has found effects of BWCs using spatially explicit designs (Kim, 2021; Monteiro et al., 2022; Ferrazares, 2023). Our paper also contributes to the literature by studying the effects on the reporting margin, which has been understudied thus far.[6]

We also contribute to a broader debate on the productivity effects of monitoring actions, employer-employee agency problems, and alignment of employees' incentives to that of the general organization, for which police officers are just one example. In other settings, a broad literature has shown that those principal-agent problems can be alleviated through monitoring (Kandel and Lazear, 1992; Barron and Gjerde, 1997), peer effects and social incentives (Bandiera et al., 2005, 2010; Mas and Moretti, 2009), and effective management systems (Bandiera et al., 2009; Frederiksen et al., 2020; Fenizia, 2022). For related work, see Ornaghi (2019) on civil service reform and Bertrand et al. (2020) and Xu (2018) on studying bureaucrats. Regarding works about the police, Banerjee et al. (2021), Battiston et al. (2021) and Kapustin et al. (2022) discussed the importance of management quality on police misconduct and other policing outcomes. Additionally, Shi (2009) and Prendergast (2021) studied the response of the police to an increase of monitoring by the public. On a similar note, our paper relates to the literature that explores how career incentives and seniority can affect productivity outcomes (Ba et al., 2021; Battiston et al., 2021).

Finally, this paper contributes to a broad literature about police interventions that aim to build trust, improve citizen relations, or reduce crime—often through deploying new technology. Doleac (2017) studied the impacts of DNA databases on crime, while Rozema and Schanzenbach (2019) specifically investigated mechanisms to prevent police misconduct. Recently, community policing programs were

[6]See also Braga et al. (2022), Boivin and Gendron (2022), and Ferrazares (2023).

shown to have no effects on trust in the police or crime rates (Blair et al., 2021), except in specific study sites (Ferraz et al., 2016). Studies on the effects of investment in police capacity have provided mixed results. Blattman et al. (2021) showed that police presence alone does not generally reduce crime in aggregate, but Bove and Gavrilova (2017) and Harris et al. (2017) showed that investment in police equipment, including weapons, has a positive effect on citizen-officer interactions, reducing both complaints and assaults against officers. This is also important because the use of force is more likely to be experienced by Black and Hispanic minorities (Hoekstra and Sloan, 2022). Our paper shows an effective intervention to reduce the use of force through monitoring of police activity.

We proceed as follows. Section 2 provides the context, presents details about the intervention, and discusses the data and measurement approach. Section 3 provides the main results and robustness checks that could threaten the validity of our estimates. Section 4 situates our results in the light of previous literature and corroborates our experimental evidence using observational data. Section 5 concludes and discusses policy implications of the experiment results.

2 Context, Intervention, and Data

Context Brazil is one of the most violent countries in the world. In 2019, the homicide rate was 20.8 homicides per 100,000 inhabitants compared to 5.0 and 1.2 in the US and the UK respectively.[7] We implemented the BWC intervention in Santa Catarina, Brazil. The state exhibits a homicide rate two times higher than the US and more than nine times higher than the UK. To evaluate the effects of BWCs, we collaborated with the Igarapé Institute and the Santa Catarina state Military Police (PMSC), the main police body responsible for patrolling, responding to emergencies, and operating the 911 hotline. It is the most visible element of the policing institutional infrastructure in Brazil. Five police precincts participated in the study: Florianópolis, São José, Biguaçu, Tubarão, and Jaraguá do Sul. Those sites were chosen given their easy accessibility from the police headquarters in Florianópolis and to represent a variety of settings in terms of socio-demographic

[7] See the United Nations Crime Trends Survey, available at `https://dataunodc.un.org/`.

characteristics and baseline violence levels.[8]

Intervention and design The ideal experiment to evaluate the effects of BWCs would randomize the presence of cameras to dispatches. This is logistically and operationally infeasible in the context of policing. Therefore, we alternatively designed an experiment that seeded random allocation at the dispatch level.

Figure 1 provides an overview of the experimental design, starting with the project timeline in Panel A. The randomization of officers was drawn about a month before the experimental period, the equipment installation and training took place the week before the experiment started, and the intervention lasted for three months.

Panel B illustrates how we implemented the intervention, which was only possible due to our unique depth of integration with the police and its IT infrastructure. In this way, the randomization between officers induced variation at the dispatch level. The table illustrates the dispatch-level data, including in the pre-intervention period. Hypothetical officers {A, B, C, and D} combine into pairs to attend a dispatch (along rows).[9] We considered a dispatch as treated if at least one of the police officers was wearing a camera. Thus, less than half of police officers had to be treated in order to achieve a 50% treatment at the *dispatch level*. For example, in the hypothetical example, if only officer A is treated, that achieves the desired treatment allocation at the dispatch level. If two officers are treated (out of four), almost no variation may be left at the dispatch level.

In practice, our simulation showed that how many officers would have to be treated in order to have around 50% of the events treated using real pre-intervention dispatch data. Panel B in Figure 1 shows the simulation result. They indicate that between one in four and one in three treated officers would amount to half of dispatches being treated (red horizontal line).

In the end, we preferred to adopt a one-in-three design (vertical line). Out of the

[8] A map of the experimental locations is provided in Appendix Figure A1, while Appendix Table A1 shows site demographics. The five selected precincts cover approximately 15% of the state population. In Section 4, we also show that study sites did not present diverging pre-trends from non-experimental precincts.

[9] In reality, the modal dispatch groups size was indeed two, but there was variation in the group size.

roster of sworn police officers per precinct we obtained in July 2018, we randomly selected one third of the officers to be in the treatment group and two thirds to the control group. This resulted in 150 officers assigned to wear BWCs and 295 control group officers. We stratified by precinct, officer activity, rank, previous internal investigations, and gender (40 stratification bins).

Besides randomizing which officers wore a BWC or not, we introduced a second layer of randomization. Every week during the 14 weeks of the experiment, two days were randomly selected to serve as blackout days, with the randomization stratified by day of the week, providing us with across-shift variation to camera exposure. Treated officers always wore a camera if their 12-hour shift fell on days that—due to our second layer of randomization—were not selected to serve as *blackout days*. We leveraged those to assess the persistence of camera effects and have across-shift variation (see Section 4). Control officers were mandated not to wear a camera on any shift.

Thus, the design induced random allocation of cameras at police dispatches, our primary unit of analysis. We considered our treatment to be exposure to cameras at the event level—that is, if there was at least one officer involved wearing a camera. Since the vast majority of dispatches involve more than one officer, our sparser one-in-three officer-level randomization was calibrated such that approximately half of the dispatches post-treatment would have BWCs present, maximizing statistical power and standing in sharp contrast with the existing literature, which had typically assigned cameras to more than 50% of the officers participating—we will elaborate on this in our discussion about the literature in Section 4.

Figure A2 displays the number of dispatches by day over the project period along with a moving average of both the total number of dispatches and the number of treated dispatches, showing that we successfully induced around half of the events treated. Out of the population of events that did not occur on blackout days, around 58% had an officer present wearing a BWC, in line with our simulations.

Integrity of the research design The integrity of the research design was protected by a host of precautions. Cameras and docking stations were kept in the armory of the police precincts that officers visit at the start and end of each shift to collect and return their service weapons and equipment. Further, the blackout days were randomly selected at the beginning of the experiment but only communicated directly to the armory the evening before to avoid potential selection around the blackout days. Moreover, dispatch operators were blind to whether dispatch units were carrying a BWC. We find no evidence suggesting significant non-compliance or other issues that could affect the integrity of the experiment, which we discuss in the robustness checks. We further describe the implementation details in Appendix A.

Throughout the implementation, the research team had strong backing from the police leadership. Following recommendations on best practice that were informed by past research suggesting low compliance with BWC use (which we discuss in Section 4 in more detail), a standard operating protocol (SOP) was developed mandating that every dispatch involving an interaction with a citizen should be recorded, with few exceptions, such as sensitive or covert operations. Officers were required to inform citizens verbally that "the dispatch was being recorded, according to police protocol," whenever the situation allowed.

The research team never had access to recordings due to individuals' privacy concerns and the sensitive nature of such data. We were nevertheless able to measure compliance at the individual dispatch level, as we outline further below.

Data and outcome measurement We primarily draw on *dispatch-level data*, which was facilitated by PMSC's fully digital backend called PMSC mobile. The data captures the universe of all events that were attended by any PMSC officers. One important aspect of the dispatch-level data is that PMSC mobile is a state-wide technology and the definitions of variables are always the same, both across space and over time. The system was already in place before our experiment, so officers did not have to learn how to use a new system together with the use of cameras. The dataset is kept for internal use by the police, and they have granted us access for this study.

Events in the PMSC mobile system typically originate from self-initiated calls due to routine operations (such as patrolling) or due to scheduled activities (e.g., the execution of court orders), although 91.8% of them are the result of police being dispatched to a 911 call of service through the central dispatch service. Our main outcome dataset contains a total of 17,665 events that span the experimental period ranging from September 3, 2018 to December 10, 2018.

The dispatch data includes a number of useful fields. It reports the time of the dispatch (arrival at the scene and end of the event), hashed officer identifiers (allowing us to link to the treatment status) and precise geo-coded locations and addresses. It further contains the information about the serial number of the cameras that officers were wearing during their dispatches, if any. Those are merged with the individual camera log files, which provide both the serial number of the device and all information on when and for how long the camera was activated. From this process, we created our measure of whether recordings actually took place.

At the end of the dispatch, officers fill in a report with their observations and descriptions about what unfolded in the operation. Our first measure of reporting is whether the police officers filled in a complete report.[10] Conditional on observing information about the nature of the events, these reports describe the crime type (out of those that have any crime type reported, the most common are verbal attrition or threat, noise complaints, burglary, assault, and domestic violence), an indicator for the presence of victims, and a field that indicates whether the dispatch was later handed over to the investigative unit ("Policia Civil"). We also made use of these reporting measures.

We also observed the types of force that were deployed in the dispatch (physical, non-lethal, or lethal force). Note that use of force is self-reported by the police officers through the police systems.[11,12] We observe if handcuffs or arrests were deployed. Our final measure of use of force pertains to contempt, disobedience or resistance charges towards police officers. We further created an inverse covariance-weighted index combining these three outcomes following Anderson

[10]In fact, this is important, as the events frequently had key details empty and/or incomplete (47.3%).

[11]As we discuss in Section 3, it is conceivable that the presence of the cameras force officers to report the facts more accurately. In this case, we would expect that the presence of cameras reduce the under-reporting of use of force, and would thus bias the estimates *against* finding any effect.

[12]The reports disclose if use of force is deployed, but they do not contain information about which officers have used force.

(2008), which we call Negative Interaction Index.[13]

Important to some heterogeneous analyses that follow, the central dispatch service pre-classifies the risk of the event. This assessment is done prior to officers being dispatched to the event, so it is not contaminated by the presence of cameras. An event was classified as high risk if any of the following conditions are met: there are individuals with life-threatening injuries, the suspect is still on site, the suspect was armed, and there is a general risk of broader disturbance to peace. An event was considered low risk if the response is negative to all these questions.

In addition to the dispatch data, we further observe a range of officer characteristics, such as their job title, rank, gender, the date of admission to the police, and the number of internal investigations that have involved the specific officer. These characteristics were also used to inform the stratification of the camera random assignment and were further used to explore pre-registered heterogeneous effects.

3 Empirical Framework and Results

In this section, we present the empirical framework and the main results and discuss range of further empirical tests to speak to the robustness of the results to guide our interpretation of the main mechanisms driving the effects. In what follows, we use the following empirical specification to study the effect of the presence of BWCs across a set of outcomes measured at the police-dispatch level:

$$y_{ibdw} = \beta \times \text{Treated Event}_i + \eta_{bw} + \tau_d + z_{ibdw} + \sum_{j=1}^{n} \phi_{o_j(i)} + \epsilon_{ibdw} \qquad (1)$$

In this specification, i indicates an event attended by a police dispatch, b is the police precinct, d is the day of the week, and w is the week of intervention. For our main specifications, $\text{Treated Event}_i = 1$ if at least one officer forming the dispatch that attended event i was assigned to wear a camera. This implies that our estimates capture an intention-to-treat effect. We include police precinct-by-week fixed effects (η_{bw}), day-of-the-week fixed effects (τ_d), and officer o stratification bins

[13] The use of force outcomes were registered in the pre-analysis plan. The reporting margin is considered part of an exploratory analysis. We detail the pre-registration of this study in Appendix Section C.

($\sum_{j=1}^{n} \phi_{o_j(i)}$). We also control for the number of officers involved in the event z_{ibdw}. In our initial specifications, we exclude blackout days and focus exclusively on comparing treated with control events.[14] The disturbance ϵ_{ibdw} is clustered at the police precinct-by-week level. We also include randomization inference p-values that are free from assumptions about the structure in the disturbance term.

3.1 Recording

Column (1) of Table 1 shows the frequency in which the cameras were activated by police officers in their dispatches. We show that, on average, 24% of the treated events were recorded. Virtually none of the control group events had any camera recording linked to them. To interpret those coefficients, note that the SOP required the use of cameras only if there were interactions with citizens, which does not occur in all dispatches. To keep some perspective, 47.3% of dispatches in the control group were reported without any information regarding the nature of the infraction. Assuming that there was no interaction with citizens in those dispatches, only the 52.7% of events could be recorded if every dispatch was treated.

Heterogeneity by officer rank Panel B of Table 1 reflects whether the effect is heterogeneous by officer rank. We seek to understand how supervision and managerial relations are key to understanding workers' performance in our context (Bandiera et al., 2009; Frederiksen et al., 2020; Fenizia, 2022). We leverage the fact that not only the event-level exposure to the camera was random but also the officer who was carrying it, allowing us to explore heterogeneous effects by their characteristics. Given the military hierarchy structure at PMSC, we classify police officers in two categories, either a low-ranked "soldier" category or a higher-ranked category for corporal or above ranks. Such rank distribution, however, is not observable to the average civilian. Importantly, compliance with the protocol was lower when higher-ranked officers carried the camera: 25% of dispatches were recorded when junior officers ported the camera, as compared to 18.9% when senior officers wore the device, a difference of 6.1 percentage points or 32.3%.

[14] Our main sample excludes shifts without cameras to correctly estimate the average treatment effect of the BWC. We later include the blackout days in the sample to study if officers change their behavior even in the absence of a camera.

There are several possible explanations that can rationalize this effect. First, it's conceivable that junior officers are more adept at using technology and adjust more swiftly to new technological tools. Second, this observation aligns with the officers' roles in the dispatch scenario, where junior officers tend to be more actively engaged with citizens, whereas more experienced officers, though at a similar physical distance, often take on more of a supervisory roles. Third, these effects are in line with a career-concerns motive: early-career officers may be more likely to show behavioral improvements and protocol compliance when in the presence of a camera. Additionally, this insight contributes to existing research by suggesting that the practice of reverse monitoring (junior officers monitoring their senior counterparts) might also play a role in alleviating principal-agency in the delivery of public services.

In common, the three main mechanisms support the idea that police behavior is likely to change when wearing cameras. This does not rule out that citizens also change their conduct when in presence of a camera, but it suggests that it is not *exclusively* citizens changing their conduct. We will return to this point when interpreting the effects on the reporting margin (Section 3.2) and use of force (Section 3.3).

Heterogeneity by treatment intensity Panel C documents that the treatment effects are larger with more on-site cameras. This suggests that the extensive and intensive margins of monitoring matter. We find that, relative to dispatches with one camera, dispatches with two or more cameras were recorded 8 percentage points more often, suggesting, as we will see below, that more intense use of cameras produces higher effects, although with decreasing returns to scale.

3.2 Reporting Effects

Police officers complete a report sheet after each incident. They describe the incident's details, including its nature, any potential victims, and information about possible perpetrators. They also collect initial evidence at the scene. This information is then stored in police systems and can be accessed later for investigations and possible legal action. Essentially, this data serves as the starting point for criminal prosecution (Blanes i Vidal and Kirchmaier, 2018). Columns (2)–(9) of Table

1 examine how the use of BWCs influenced what police officers report in their dispatches.

Column (2) of Table 1 shows a record of how frequently events were reported with missing the information about their nature. This can naturally occur when, for instance, the dispatch team cannot identify the source of the event, locate individuals, or find evidence related to the reported crime. In practice, instead of assuming that information is purposely missing, this might also indicate that officers are making a conscientious effort to adhere to the established reporting standards of the police department. This accounted for 47.3% of control dispatches. The estimates reveal that the presence of BWCs reduced the occurrence of events being reported without data collection by 2.77 percentage points, which is equivalent to a 5.9% decrease, amounting to 368 dispatches during the treatment period.

Columns (3)–(7) deepen the analysis and investigate which crime types were reported more often in treated events. We observe the effects across the five more recurring crime types, of which we report the mean of control events in parentheses: verbal attrition or threat (9.9% of dispatches), noise complaints (8.6%), burglary (4.8%), assault (3.6%), and domestic violence (1.6%). According to the randomization inference p-values, we find no significant effects across the first three outcomes, and a marginally significant effect at 10% on assault (with a p-value of 9.6%). In contrast, we find an increase in the reporting of domestic violence by 1.14 percentage points, highly significant at 1% significance level. Compared to the control group (domestic violence being reported for 1.6% of events), this represents a shift of 69.2% more cases of domestic violence being reported, or 151 cases. It also suggests that $\frac{1.14}{2.77} = 41.2\%$ of the reduction in under-reporting is translated into domestic violence cases. We interpret these findings as further evidence that BWCs improve reporting by police officers. This is particularly relevant in cases such as domestic violence, in which, without hard evidence, the prosecution is considerably more challenging.[15]

Column (8) examines the impact of cameras on the share of incidents with reported victims. We discovered that, on average, about 13.8 percent of reports

[15] Insufficient evidence is identified as one of the underlying causes of the high attrition rates on gender-based violence in the criminal justice process (UNODC, 2014, pp. 38).

include victims, and the presence of cameras increased this number by 2.8 percentage points, which corresponds to a 20.1% increase. Moving to Column (9), we investigate the effects of police officers wearing BWCs on whether the incident they respond to results in the creation of an investigative report by the investigative police unit ("Polícia Civil"). We observed a significant 9.5 percentage point increase in the likelihood of an investigative report being generated as a result of using BWCs during police operations.

Now, we seek to offer suggestive evidence about how the observed effects on the number of victims and the generation of investigative reports might be linked to a substantial increase in domestic violence cases. It is conceivable that the more frequent reporting of these cases could naturally result in an increase in the number of reported victims and a higher likelihood of investigation. This aspect is crucial in the broader conversation about the overall welfare implications of the reporting impacts of BWCs. As mentioned in the introduction, such increases in reporting are not necessarily welfare enhancing. For instance, as demonstrated by Monteiro et al. (2022), by reducing officers' discretion, the use of BWCs may also lead to more reporting of less significant crimes, like small drug possessions, which could potentially divert police resources from addressing more socially significant crimes and/or increase the prosecution of misdemeanors, leading to demonstrated negative consequences for the lives of those who are prosecuted (Agan et al., 2023).

To shed light on this issue, we observe that a large share of domestic violence cases reported a victim (73.4%) and were brought to further investigations (75.3%) at the baseline. Thus, it is conceivable that the effects in Columns (8) and (9) of Table 2 are indirect or downstream consequences of the increase in the reporting of domestic violence. To quantify the extent to which this is the case, we employed the following back-of-the-envelope calculation. As reported above, the rescaled treatment effects suggest that an additional 151 domestic violence cases were reported. Under the baseline parameters, this translates to an additional 110.8 cases with victims.[16] In turn, the point coefficients in Column (8) indicate that an additional 369.4 cases reported a victim. Thus, roughly $\frac{110.9}{369.4} = 30.0\%$ of the main

[16]This is under the assumption that those shares are unchanged in the period after the intervention, which is likely a lower bound considering that cameras improved reporting.

effects could be attributed to a downstream consequence of the increase in domestic violence reporting alone. A similar calculation suggests that 27.4% of the coefficient in Column (9) regarding investigative reports could be attributed to the increase in domestic violence. Overall, this suggests that a sizeable portion of the amplification of the demands for investigative services were driven by these cases.

Heterogeneous treatment effects The heterogeneous effects broadly mirrored those reported in Column (1) and explained in Section 3.1 when exploring the effects on compliance. Panel B shows that the rank of the officer assigned to hold the camera matters in explaining the treatment effects. The BWC treatment effects in the reporting and interaction margins were only present when an officer with a *soldier* rank was holding a camera in the dispatch unit. Treatment effects were null when senior officers were the only ones able to activate the camera.[17] This is coherent with low compliance effects and is, again, suggestive of office rank being an important mediator of the camera effects. Panel C documents that a higher number of cameras on the site was associated with better effects on reporting, although with likely decreasing returns to scale.

3.3 Effects on Use of Force

Table 2 displays the primary findings regarding police use of force along different measures. Column (1) shows the impact of the police deploying physical, non-lethal, or lethal force. We note a significant decrease in the use of force in the treated events; 0.69% of the control events reported use of force in comparison to 0.26% of the treated events, corresponding to a reduction of 0.43 percentage points or 61.2%.

Although self-reported use of force data can be subject to underreporting, the presence of cameras is likely to mitigate this issue, as suggested by the improved reporting mentioned earlier. In this case, the reduction of underreporting would be a bias *against* finding treatment effects. That is, if police officers were underreporting their use of force and the BWCs work as an incentive for them to report truthfully, we would interpret the results of Panel A as a lower bound to the true

[17]The coefficients on reporting of verbal attrition and threat is even negative in those cases.

effect of BWCs on use of force.[18] The substantive decline in use of force marks a notable contrast with the existing literature which has mostly found mixed effects. We revisit this divergence from previous work in Section 4.

Columns (2) and (3) substantiates the effects along other margins of use of force between the police officers and citizens. The impact of BWCs on handcuffs or arrests and charges of contempt, resistance, or disobedience. The main effects find negative point coefficients (of 5.9% and 28.2% respectively), although not statistically significant. We then combine all three indicators on the "Negative Interaction Index," which revealed a significant and relevant causal estimate of 47.0% of reduction of adverse interactions between citizens and the police.

We interpret these effects as ensuing from the accountability and diligence of the police actions and on-the-scene dynamics promoted by the camera. We next delve into the mechanisms for why this effect takes place, as evidenced by the heterogeneous treatment effects.

Heterogeneity by officer rank and treatment intensity We first reproduce the same heterogeneity as in Panel B of Table 1 in order to observe how the camera effects on use of force are mediated by the officer rank structure.

When interpreting this variation, it is important to note that the salience of the camera to the citizen remains consistent regardless of which officer is wearing it. Both junior and senior officers have close interactions with the public. Therefore, since the camera's salience to the public remains constant, any differing effects can be attributed to and understood as likely outcomes of changes in police behavior.

In line with Section 3.1, we only find significant effects on all use of force measures when junior officers were required to wear the cameras, with point coefficients generally smaller in absolute value and non-significant for all outcomes. Echoing the compliance and reporting effects, this confirms that police officer behavior was a likely mediator of the changes in use of force (without eliminating the possibility that there was a behavioral change by the citizens, too). In other words, these results are suggestive that citizen pacifism cannot be the *only* explanation for why BWCs produce these strong reductions in use of force.

[18] See also Monteiro et al. (2022) for an observational study for São Paulo which estimates comparable effects on the use of force using independent measures of police violence.

Panel C documents that the treatment effects were larger with more on-site cameras. Relative to dispatches with one camera, dispatches with two or more cameras had a reduction in the Negative Interaction Index – promoting a further drop of 25.9%. In particular, the use of force fell by 79.8% when two cameras were present, also representing how increasing the treatment intensity increased the magnitudes of the effects, although with decreasing marginal returns to scale. This effect, however, is only statistically significant at the 10% level.

Heterogeneity by ex-ante risk assessment We now extend our analysis beyond the previous tables' heterogeneity margins to focus on aspects relevant to use of force cases. Panel D studies whether effects were primarily concentrated in events classified *ex-ante* as low risk.

The results presented in Panel D of Table 2 suggest that the effects of BWCs were driven by events that were ex-ante classified as low risk. For those events, the negative interactions index was reduced by 51%. Columns (1)–(4) show that the point coefficient is negative for all the index components. Only handcuffs or arrests are not statistically significant at the 10% level. No BWC effects are detected among events that were considered to be high-risk ex-ante: the index points to a much smaller and non statistically significant reduction of 8.8%. Only the point coefficient on the use of force is negative, and it is non-statistically significant with large standard errors. This suggests that BWCs may mitigate the escalation of situations. In high-risk events that already escalated prior to dispatch, however, the camera's presence itself may not have affected the situational dynamic. This occurred even though the reporting margin was strongly affected by the presence of cameras in high-risk events. Those results suggest that cameras indeed serve as a way to de-escalate conflicts, diffuse tensions, and ensure a better cooperative environment on both sides.

Baseline use of force We explore the extent to which the camera effects are higher in areas with higher likelihood of use of force in dispatches at the baseline. To do so, for each census tract, we counted the events with use of force in the 13 weeks before the experiment and split the areas along the median for each municipality.[19] The results, depicted in Panel E of Table 2, interact the measure of baseline use of force with the treatment indicator and suggest that stronger treatment effects are observed in areas of the municipalities that experienced higher use of force at the baseline, despite compliance being unchanged across the two areas. The Negative Interaction Index suggests a reduction in use of force that is nearly five times larger in absolute terms in areas with a historically higher propensity to involve use of force compared to areas with a lower historic use of force. This is consistent with cameras being effective deterrent devices, especially in places and situations where use of force would be more likely to unfold counterfactually. This indicates that BWCs may be particularly suitable to benefit citizen-police dynamics in areas that historically involved a higher degree of use of force. We consider this to be particularly important when considering how the roll-out of such technology may need to be prioritized in areas with relatively strained citizen-police interactions.

Blackout specifications We next discuss whether BWCs affect the behavior of police officers *after* they had worn the devices in previous days. Our analyses that follow could be suggestive of learning effects taking place. To examine this, we replicate Equation (1) for the blackout days. That is, we compare events that would be treated against those that would be control if counterfactually not on blackout days.[20]

The results are presented in Appendix Tables A2 and A3 for, respectively, effects on reporting and on use of force. Column (1) of Appendix Table A2 highlights that, on blackout days, hardly any event were recorded due to the experimentally induced absence of cameras. We still find that a minority of events were recorded (3.3%), which may have occurred because some officers started their shifts on regular days, and ended during blackout periods. According to experimental protocol, in those cases they would continue to use the cameras due to operational constraints of returning the devices while attending dispatches. Despite the much lower recordings, we still find significant effects on decreasing the frequency of re-

[19]The results are robust to using alternative measures of baseline use of force or geographies.

[20]This is preferable to exploring the variation across days: throughout our sample period, we had only 24 blackout days.

ports with no information, and continue to observe increased reporting of verbal attrition and threats. Again mimicking the main results, these events generated investigative effects (although on the margin of significance).

Table A3 demonstrates that all BWC treatment effects on the outcomes measuring citizen and police use of force disappeared. Yet changes in the reporting behavior persisted. This could indicate that officers who were previously exposed to the use of cameras behaved differently, even in the absence of the camera. The learning effects primarily affected the reporting margin but did not appear to have an effect on use of force or other citizen-police interaction margins.

3.4 Robustness

No endogenous allocation of dispatches The observed decrease in the interaction margins could be confounded by a change in the pattern of policing rather than cameras inducing a change in officers' behavior. This could occur, for example, if the treated officers chose to patrol safer areas or areas with less potential for the use of force compared to the control dispatch units. We show that this hypothesis finds little support in our data in multiple ways.

We test for the absence of endogenous sorting as a function of treatment in an econometric setting. We estimate Equation (1) with characteristics of the event as outcomes and evaluate if the treatment affected them. The results are presented in Table 3. In column (1), we test if officers with cameras avoided geographic locations with a higher baseline use of force.[21] The results show that cameras were not allocated to events as a function of the baseline level of use of force. Overall, we find no reasons to believe that cameras hinder officers from working in areas where citizen-police interactions are more likely to escalate.

To test the allocation of dispatches in space, we regress treated events against latitude and longitude measures in Columns (2) and (3). We find that treated and control events were statistically similar concerning latitude. The point coefficient for longitude is 0.002 of a degree. This distance is negligible, representing roughly 200m measured at the equator; further, while the point estimate appears significant

[21] This follows the division of the sample between above and below within-precinct median baseline use of force, the same as used in Table 2 Panel E.

using conventional inference methods, it is insignificant using a randomization inference (p-value of 0.374).

Column (4) captures if the event occurred in a census tract above median levels of income at baseline. A positive coefficient would indicate that officers were more likely to tend to dispatches in relatively wealthy neighborhoods when wearing cameras. The effect is statistically insignificant with a randomization inference p-value of 54.1%.

Columns (5) and (6) reflect whether the treatment affected measures such as time to dispatch and an indicator if the time to dispatch was greater than five minutes. The interval between an incident being reported and the officer arriving at the scene was the same between treated and control events. Therefore, in summary, treated and control events occurred in the same places and had the same baseline level of use of force, and treated officers did not take longer to get there.

A potential concern could be that cameras may change the way dispatches occur. Most dispatches were initiated by the central dispatch and call handlers, who were blind to the underlying treatment status. Nevertheless, there is a potential concern that officers carrying a camera may not initiate events at the same rate as officers without cameras: the regression outcome in Column (7) is a dummy variable equal to one if the police self-initiated the dispatch. We also test whether treatment status is uncorrelated with call-handler-induced dispatches: Column (8) confirms this. Columns (7) and (8) together confirm there were no compositional effects in the method through which dispatches were initiated. Finally, Column (9) shows that there was no differential assignment of officers with cameras to events based on their ex-ante risk level as measured in Panel D of Table 2.

Appendix Table A4 goes further and explores if the null effects could be heterogeneous by officer rank. It is possible that more experienced officers are more able to anticipate the monitoring effects of the camera and react more strongly by altering their policing patterns. We find that, except for latitude and longitude, the heterogeneous effects are non-statistically significant using randomization inference p-values. The effects on latitude are only significant when the dispatch composition had both high- and low-ranked officers but not significant with only high-rank types. Effects on longitude are .009 and .012 degrees with one and two high-ranked officers respectively, significant with randomization inference p-values of 7.0% and 2.9%. The point coefficients are interpretable as a shift in dispatches of 900m and 1.2km to the east. As the other columns suggest, this spatial change is not correlated with shifting patterns toward places with less baseline use of force, higher income levels, higher baseline ex-ante risk, a reduction in time to dispatch, or other compositional effects. This difference is not visible in Appendix Figure A3 which shows the dispatch heatmap by treatment status. In summary, the results suggest that there was no substantial selection in space as a function of the treatment and that officers' behavior with respect to patrolling, as arriving at the event and the working location did not seem to be altered as a function of the cameras.

No endogenous allocation of senior officers to riskier events One potential concern is if high-rank officers are more likely to be dispatched to higher-risk events, which could have a lower potential for the de-escalation of conflicts. The heterogenous treatment effects documented in Panel B of Table 2 show that treatment effects were driven by events in which a lower-ranked officer was carrying a BWC. Therefore, this could raise a concern about senior officers being allocated to attend to higher-risk events, thereby confounding the results. Again, our experimental protocol ruled this out as dispatch operators were blind to the respective treatment status of any dispatch unit. Reassuringly, Appendix Table A5 shows that the presence of a higher-ranked officer was not correlated with an event being classified ex-ante as high risk.

Alternative sample composition Appendix Table A6 demonstrates that the results are robust to changes in the estimation sample. Panel A reproduces the main effects for reference. In Panel B, we include data from blackout days. Not surprisingly, the treatment effect is still present but is smaller in magnitude since the sample includes days when officers were randomly not handed out cameras.

Panel C looks at dispatches with two officers, which was the modal dispatch size. The results show that when we restrict the sample to these events, the effect on use of force loses precision, even though it remains negative and sizable. The effects of the Negative Interaction Index on adverse citizen behavior remain strong and statistically significant. Finally, Panel D excludes dispatches with more than four police officers, and the results remain virtually the same. Overall, our results remain qualitatively unchanged in this exercise.

Exploiting only observational variation In Section 4, we position the findings from this paper in the context of the much broader literature on BWCs, and we discuss an analysis that can be seen as a further robustness check. We estimate treatment effects exploiting only observational variation, exploring the spatially explicit dimension of our intervention. Using the group of precincts that did not participate in the experiment, we assess how use of force evolved in the experimental precincts vis-a-vis those in which no officer wore a camera. This approach can allay some concerns about potential unobservable within-precinct spillovers. As discussed in the section on spatially explicit designs covered in section 4, we find similar point estimates when replicating our analysis exploiting such a non-experimental research design in our data.

4 Reconciling BWC Effects with the Literature

The literature on BWCs' effects on use of force has produced mixed results thus far. While recent meta-studies (Lum et al., 2020; Williams et al., 2021) have failed to detect effects of BWCs,[22] some recent evidence is more encouraging and points to detectable and meaningful effects on use of force (see, for example, Braga et al., 2018; Kim, 2020; Ferrazares, 2023; Monteiro et al., 2022).

The results of our paper aligns with the most recent literature, which has detected that BWCs can decrease use of force, but there is a more significant portion of the literature, closely related to our paper in methods, that has failed to detect meaningful effects of BWC on use of force. Appendix B and Appendix Table A7 provide an overview of the literature, including description of the results and methodologies.

In this section, we attempt to make sense of these differences and reconcile

[22]The meta-analysis found two overall significant BWC effects from a total of 12 outcomes: a decrease in complaints against officers and an increase in non-traffic citations. The other outcomes including stop and frisk, calls for service, assaults on officers, resistance, and others, showed no meta-analytic effects.

the mixed results in the literature. We have classified the papers as shift-centric-, officer-centric or spatially explicit according to the experimental design and evaluation and measurement strategy adopted for the implementation of BWCs. While most shift- and officer-centric papers are RCTs, the majority of papers that study the implementation of BWCs across space use difference-in-differences methods.

Our paper is unique since we could study the effects of BWCs using the three possible available designs in the literature. Moreover, the granularity of our data and the variation across officers, shifts, and space allows us to replicate the evaluation designs used in most past studies *in our data*. Our study directly nested shift- and officer-centric research designs, and we also could leverage spatial variation in the BWC implementation.[23]

We use these analyses to develop two interconnected arguments. First, spillover and contamination effects may have attenuated some previous evaluations of BWCs which relied exclusively on officer or shift variation. The potential for spillover effects in this context is considerable, as, for example, control officers patrolling with treated officers were indirectly treated and very likely to alter their behavior. This may be an important SUTVA assumption violation. Replicating the design of those studies in our data uncovers BWC treatment effects that do not reject the null of no effects and with treatment effect estimates of similar magnitude.

Second, we also investigate how the data aggregation might have affected the BWC effect estimates. Analysis at the event level helps account for the aforementioned contamination issues that can arise if data was aggregated at the officer- or shift-level. Further, we find that more disaggregated data — as used in this study — provides for increased precision by increasing sample sizes and enabling detailed controls, e.g., fixed effects at very fine levels. In other words, these findings are consistent with the interpretation that past studies using aggregate data may have suffered from bias and power issues.

We also compare our estimates to those relying on observational variation in a difference-in-differences framework (similar to Braga et al., 2020, and Monteiro

[23]Appendix Table A7 overviews the main features of 30 papers we looked at in some detail that investigated the effects of BWCs. There were 7 shift-centric papers, 11 officer-centric, and 12 spatially explicit designs.

et al., 2022) and suggest that both can exhibit relatively similar BWC effects.

Finally, different findings could have naturally arisen due to the different settings in which the experiments were conducted. For instance, our study is among the first to evaluate the effects of BWCs in a middle-income, high-crime environment (compared to existing studies being mostly conducted in the UK or the US). We cannot ignore that contextual effects might have taken place. Yet the exercises — which we describe next — naturally suggest that the mixed results can be at least partially explained due to the empirical design rather than the absence of true effects.

4.1 Unit of Randomization and Analysis

We first focus on specific dimensions of the randomization. Doing so allows us to reproduce the experimental setting of other studies by reanalyzing the data at the officer or shift levels.[24] The level of detail of our dataset and the two-layer design of our experiment allows us to replicate shift- and officer-centric designs in our data and compare results with the treatment effects at the event level – our unit of analysis. This allows us to investigate the extent to which the BWC camera effect estimates are sensitive to the experimental design.

In Panel A of Figure 2, we present the results of this comparison. The point estimate displayed in red shows the 61.2% reduction in use of force between treated and untreated events, estimated from the nominal effect size of Table 2, along with the 95% confidence interval.[25] The event-level estimate is also indicated with the red dashed horizontal line across all panels for ease of comparison with other designs.

As we noted, officer- and shift-centric designs make up the vast majority of experimental BWC studies, which have detected mixed results regarding BWCs' effect on use of force. Using our data, we explore how changing the randomization unit impacts the BWC treatment effect estimates. We start with mimicking an *officer-centric* study, which randomizes officers into treatment and control groups.

[24]Table A7 organizes 30 papers that were surveyed. Out of those, 16 papers are RCTs, and they either randomize at the shift level or at the officer level.

[25]We normalize our coefficients in percentage reductions relative to the baseline incidence of use of force to render the estimates comparable across studies.

As outcome variable y_{od}, we measure the share of incidents in which an officer o used force over a time period – say, a day of the week d. We then explore the experimental variation in officer allocation to the treatment and control groups in the following specification:

$$y_{od} = \beta_{\text{officer}} \times \text{Treated Officer}_o + \eta_{bw} + \tau_d + \phi_o + \epsilon_{od} \qquad (2)$$

excluding blackout days, so we solely rely on the between-officer variation. As in our main specification, we include police precinct-by-week fixed effects (η_{bw}) along with day-of-the-week fixed effects (τ_d). We also include stratification bins fixed by officer o effects ϕ_o. The disturbance ϵ_{od} is clustered at the police precinct-by-week level. Treated Officer$_o$ equals 1 if the officer was assigned to wear a camera. We are interested in the estimated β_{officer}.

The results plotted in Panel A of Figure 2 suggest an attenuation: the effect size is reduced from our event-level benchmark of 61.2% to 26.5%. The estimated treatment effect size capturing a decline in use of force of 26.5% is 57% smaller compared to the estimated treatment effect when carrying out the analysis at the event level. The attenuation is not surprising: in our design, one third of the officers were randomly selected to wear a camera. Due to dispatches typically involving more than one officer, this indirectly resulted in around 58% of the events being treated with at least one camera present. Since a noticeable share of events attended by control group officers was, in fact, indirectly treated due to the presence of other experimental officers carrying cameras, this downward biases the treatment effect estimate since a large share of events coded as being attended by control-group officers were treated. Such contamination-induced attenuation bias may affect many existing studies designed at the officer level that use difference-in-means econometric frameworks. Almost all existing studies could not directly test or measure contamination due to a lack of detailed event-level data.[26]

Furthermore, the extent of contamination-induced attenuation bias is likely in-

[26]Of the 12 studies that opted for an officer-centric research design, we were able to identify from the papers whether officers were dispatched in teams for only six studies — out of those, 66% reported that officers were dispatched in groups of two or more officers. These studies may thus have been vulnerable to such attenuation bias.

creasing in the share of officers that wear a camera. In Column (10) of Table A7, we see that virtually all officer-centric studies opted for a design with around 50% of officers assigned to wear a BWC. Assuming a similar dispatch composition as in our context, this implies that 75% of all events were treated with at least one camera (see Panel B of Figure 1), undermining the statistical power and downward-biasing the treatment effect estimate when considering officer-level data. Therefore, the attenuation of the results is consistent with spillover effects since the analysis at the officer level did not account for the fact that control officers would sometimes mechanically tend to dispatches with treated officers.

The final estimate in Panel A of Figure 2 presents the treatment effect estimate implied in our data carrying out the analysis when we solely exploit treatment and control variation across shifts. In this case, we collapse the data at the precinct-by-day level, and we exploit the fact that our research design allows us to contrast blackout and non-blackout days to give us treated and untreated shifts. This is close to the experimental design of shift-centric papers because a day comprises approximately two consecutive 12-hour police shifts. In the following specification, the outcome variable y_{bd} is the share of events in which force was used at police precinct b during day of the week d,

$$y_{bd} \quad = \quad \beta_{\text{shift}} \times \text{Treated Shift}_d + \eta_{bw} + \tau_d + \epsilon_{bd}. \tag{3}$$

The fixed effects we control for are police precinct-by-week and day-of-the-week. The error term ϵ_{bd} is clustered at the police precinct-by-week level. The estimated effect sizes are around 16%, a substantial attenuation from the 61.2% reduction in use of force that was estimated from the event-level specification, and statistically insignificant. This effect size is comparable with studies that used variation at the shift level. For example, Ariel et al. (2016b) also used data at the precinct-shift level, explored shift randomization, and is within our confidence interval.

4.2 Temporal Resolution of Outcome Measurement

Accounting for unobserved time effects may be important as well. Our design at the event level allows us to control for granular time effects that might confound the effects. When comparing to the literature and focusing on the officer-level variation to uncover BWC effects, we aggregate the data either at officer by day level, officer by month level, or pooling officer observations during the whole experimental period. Specifications at coarser levels may introduce a broad range of biases, as they would imply that we cannot control for the potential confounding effect of time fixed effects which are likely very relevant. Only a few studies have considered time fixed effects as additional control variables in their respective econometric framework, with the majority of studies either ignoring time, reducing the time-dimension to before-and-after comparisons, or simply estimating differences-in-means without control variables (see Appendix Table A7).

Panel B of Figure 2 documents what happens to our point estimates with various data aggregations. The first effect size — outcomes at the officer-day level — is replicated from Equation (2), the most granular aggregation of the event-level data when exploiting our experimental variation at this unit of analysis. The second model aggregates the data to the officer-month level. In this case, the outcome variable is the share of the events with use of force by police officer o during month m. We estimate the following specification:

$$y_{om} = \beta_{\text{officer-month}} \times \text{Treated Officer}_o + \eta_{bm} + \phi_o + \epsilon_{om} \tag{4}$$

where index o refers to an officer, while index m indicates the month. We include police precinct-by-month and stratification bin by officer o fixed effects. Although the effect size does not change considerably, the precision decreases substantially. This can have two main reasons: first, naturally, we have a smaller sample size, which implies that with conventional inference, the standard error estimates are less precise.[27] Furthermore, a coarser design does not allow for the inclusion of other relevant controls, such as granular fixed effects which, while uncorrelated with the treatment, in our experimental setup would improve the precision of the point estimates.

We further aggregate the data for each officer and consider the whole experi-

[27] Clustering the data at the precinct-month level would not be adequate due to the low number of groups that this combination provides, so we instead use heteroskedastic-robust standard errors.

mental period. Such simple group comparisons are often found in the BWC literature, accounting for at least one third of the existing experimental and non-experimental studies we surveyed. We only exploit cross-sectional variation arising from the randomization of the treatment status. We refer to this as the "pooled" specification. The estimating equation is:

$$y_o = \beta_{\text{officer-pooled}} \times \text{Treated Officer}_o + \eta_b + \phi_o + \epsilon_o \quad (5)$$

where we include only precinct and stratification bin fixed effects. Standard errors are heteroskedastic robust. The effect size from this equation is smaller in magnitude and also statistically insignificant. As a reference from the literature, Yokum et al. (2019) also used data at the officer level pooled during their experimental period. The effect size they found for use of force was virtually zero in both magnitude and statistical significance. Our pooled result is comparable to theirs, and, again, their point estimates shown in the horizontal black dashed line fall within our confidence intervals.

4.3 Difference-in-Differences Designs

In this section, we use difference-in-differences (DiD) frameworks with our data. DiD empirical frameworks are widespread in observational studies of BWC effects and typically come in two forms: either to study treated and untreated officers or treated and untreated spatial units over time. In the officer-centric DiD evaluation framework, the main concerns that may cause biased estimates are spillovers from control officers working with treatment officers and measurement error. In the spatially explicit DiD design, the prime concern is statistical power, especially in the context of low compliance and coarse outcome measurement.

Officer-centric DiD design We first present point estimates that emerge in our data when employing a DiD design that compares the changes in outcomes associated with officers assigned to wear BWCs with those officers who never wear cameras. We estimate the following specification:

$$y_{obd} = \beta_{\text{officer-did}} \times \text{Treated Officer}_o \cdot \text{Post}_t + \eta_{bw} + \tau_d + \phi_o + \epsilon_{bdw} \quad (6)$$

where b is the police precinct, d is the day of the week, and w is the week of intervention. We include police precinct-by-week (η_{bw}), day-of-the-week (τ_d) and stratification bins fixed effects (ϕ_o). The disturbance term is clustered at the precinct-by-week level.

The first estimate in Panel C of Figure 2 presents the treatment effect estimate, which suggests that among officers assigned to wear BWCs, use of force decreased by 32.5% relative to untreated officers. This point estimate is still around 47% smaller in absolute magnitude compared to the point estimate obtained when carrying out the analysis at the event level, and it remains consistent with the difference-in-means presented in Panel B. This is explained by the confounding effect of spillovers arising from treatment and control officers being dispatched together. Panel C illustrates the point estimate from Braga et al. (2018), which was one of the few existing studies that opted for such an evaluation approach and found smaller but statistically significant treatment effects of BWC reducing use of force.

Spatially explicit design We then move to a spatially explicit design where we compare outcomes of experimental against non-experimental police precincts. In our study, we obtained data from non-experimental precincts that were included in parallel research on the effects of community policing (see Blair et al., 2021; Barbosa et al., 2022). We leverage the data from those study sites to estimate a DiD design using non-experimental precincts as the control group. Naturally, these estimates may also suffer from some attenuation due to the sparsity of the treatment: only 58% of the events in experimental municipalities in the post-period were treated as per our randomization protocol. Such attenuation bias in treatment effect estimates would not arise if all officers in treated precincts were given BWCs, as is common practice in some existing studies leveraging spatially explicit research designs. We aggregate our main outcome variables measured at the event level to the precinct-by-day level by calculating the share of events in a given precinct and day that involved use of force. We estimate the following equation:

$$y_{bdw} = \beta_{\text{precinct-did}} \times \text{Treated Precinct}_b \cdot \text{Post}_t + \eta_b + \eta_w + \tau_d + \epsilon_{bdw} \qquad (7)$$

where b stands for police precinct, d is the day of the week, and w is the week of intervention. Standard errors ϵ_{bdw} are clustered by precinct-week.

The second point estimate in Panel C of Figure 2 presents the results. We find that treated precincts presented a 0.17-percentage-point reduction in use of force, equivalent to a 46.1% decline of average use of force.[28] This treatment effect estimate is imprecisely estimated, suggesting that the research design may struggle with power and measurement error introduced with the aggregation, and that a large share of events in "treated" precincts were untreated. Nevertheless, the point estimate gets closest to the event-level estimate, just around 25% smaller in absolute value. Out of 12 spatially explicit studies, most often exploiting non-experimental variation, only Kim (2021) and Ferrazares (2023) found a negative treatment effect suggesting that BWCs may reduce use of force — albeit with less of an effect than what we document here.

Furthermore, we also consider this exercise as a robustness check to the treatment effect estimates, given they are obtained solely by exploiting observational variation. It also serves as corroborating evidence for the estimates presented in Section 3.

De-policing hypothesis We use the spatially explicit design with untreated precincts to test if the adoption of BWCs could have negative effects on police activity. One could argue that BWCs would make officers change behavior and decrease oversight, which would in turn affect the number of events registered in the police database. If the adoption of BWCs had an effect of making officers overlook crime, we would expect that the number of officer-initiated events would decrease in precincts that adopted BWCs. Relatedly, one could also argue that a decrease in police activity could increase total crime, which would in turn increase the number of telephone-initiated events.

We run a DiD in a specification similar to equation 7, but using the number of events, the number of telephone-initiated events and the number of police-initiated events as dependent variables. Table A8 shows no effects of BWC adoption on police activity in any of the three dependent variables. These results rule out the

[28] Appendix Figure A4 shows that pre-trends were absent for both DiD designs in this section.

hypothesis that there was de-policing due to bodycam adoption.

5 Conclusion

The issue of police violence is a global problem, and there is an urgent need to find ways to increase accountability. In 2021, Brazil suffered from high levels of police brutality, resulting in 6,145 deaths caused by police action.[29] Through a large-scale experiment in Santa Catarina, Brazil, we have revealed that the implementation of BWCs by police can significantly lower instances of force by an average of 61.2%. While our results did not measure deaths in police action, it is conceivable that some of the benefits in reduction of use of force could translate into curbing loss of life during police encounters.

We find evidence that implies that the change in conduct is not solely on the part of the citizens, and that some behavioral change can be taking place on the part of the police. We also have revealed the impacts of the devices on improving police reporting, especially concerning domestic violence (the reporting of which grew by 69%), which goes in tandem with a reduction of officers filing empty reports. This is important, as accurate reporting by police officers has been demonstrated to improve the rate at which crimes are cleared (Blanes i Vidal and Kirchmaier, 2018).

Taken together, our results indicate that using BWCs can increase the accountability of police officers. This has important policy implications. The positive results of this research paper have already proven to impact the policymaking and the adoption of police BWCs. Following this study, the government of Santa Catarina purchased 2,425 cameras in August 2019, and now virtually all dispatches are recorded in the state.[30] Since then, the Military Police of São Paulo state has also adopted BWCs and Monteiro et al. (2022) evaluated the effect of cameras in this context using observational and aggregated data, finding similar effects in the reduction of use of force. More recently, BWC policy has been formulated at the

[29]Source: Forum Brasileiro de Segurańca Publica, http://forumseguranca.org.br:3838/, accessed November 2023.

[30]Source: https://www.pm.sc.gov.br/noticias/policia-militar-lanca-cameras-policiais-individuais, accessed March 22nd 2023

national level through the work of the Ministry of Justice.[31] The Supreme Court of Brazil mandated that the police of the state of Rio de Janeiro should adopt BWCs in the wake of allegations of excessive use of force.[32] More generally, BWCs feature as a key program of the Bureau of Justice Assistance of the United States' Department of Justice.[33] The growing take-up of BWCs suggests that there is a latent demand for interventions that increase scrutiny and oversight of the police; and that cameras can provide one such solution with visible impacts.

Our final point is concerned with external validity. While our study provides evidence of strong effects of BWCs, it was also conducted under experimental conditions with relatively strong involvement of police supervisors. Other successful implementations, such as in São Paulo, also stressed the role of supervisors in ensuring that street-level officers, who directly interact with the public, adhere to camera-use mandates and protocols.

This suggests that BWCs should be seen as part of a broader set of incentives set out by the institutions, and any understanding of BWCs' effects should account for how monitoring devices interplay with these incentives. The role of monitoring in the workplace as a broader topic has been stressed both in the context of private- and public-sector agents (see Bandiera et al. (2010), Frederiksen et al. (2020), and Fenizia (2022), among others). This may corroborate the view that strong supervisions may be necessary in order for similar effects to materialize in future implementations. This further highlights that advancing our understanding about the importance of monitoring becomes more fundamental when the outcomes in question are police violence, use of force, and, in extreme cases, the death of citizens resulting from these actions.

[31] Source: https://www1.folha.uol.com.br/cotidiano/2023/02/governo-lula-prepara-programa-de-cameras-em-uniformes-policiais-para-o-1o-semestre.shtml, accessed March 22nd 2023.

[32] Source: https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=508510&ori=1. Accessed November 28th 2023.

[33] Source: https://bja.ojp.gov/program/bwc, accessed March 22nd 2023.

References

Agan, A., J. L. Doleac, and A. Harvey (2023). Misdemeanor prosecution. *The Quarterly Journal of Economics 138*(3), 1453–1505.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention : A reevaluation of the abecedarian , perry preschool , and early training projects. *Journal of the American Statistical Association 103*, 1481–1495.

Ariel, B., A. Sutherland, D. Henstock, J. Young, P. Drover, J. Sykes, S. Megicks, and R. Henderson (2016). Wearing body cameras increases assaults against officers and does not reduce police use of force: Results from a global multi-site experiment. *European Journal of Criminology 13*, 744–755.

Ashraf, N. and O. Bandiera (2018). Social incentives in organizations. *Annual Review of Economics 10*, 439–463.

Ba, B., P. Bayer, N. Rim, R. Rivera, and M. Sidibé (2021). Police officer assignment and neighborhood crime. Technical report, National Bureau of Economic Research.

Bandiera, O., I. Barankay, and I. Rasul (2005). Social preferences and the response to incentives: Evidence from personnel data. *The Quarterly Journal of Economics 120*(3), 917–962.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social connections and incentives in the workplace: Evidence from personnel data. *Econometrica 77*(4), 1047–1094.

Bandiera, O., I. Barankay, and I. Rasul (2010). Social incentives in the workplace. *The review of economic studies 77*(2), 417–458.

Banerjee, A., R. Chattopadhyay, E. Duflo, D. Keniston, and N. Singh (2021). Improving police performance in rajasthan, india: Experimental evidence on incentives, managerial autonomy, and training. *American Economic Journal: Economic Policy 13*(1), 36–66.

Barbosa, D. A. C., T. Fetzer, C. Soto-Vieira, and P. C. L. Souza (2022). Can trust be built through citizen monitoring of police activity? In *Crime, insecurity, and community policing: Experiments on building trust.* Cambridge Press.

Barron, J. M. and K. P. Gjerde (1997). Peer pressure in an agency relationship. *Journal of Labor economics 15*(2), 234–254.

Battiston, D., J. Blanes i Vidal, and T. Kirchmaier (2021). Face-to-Face Communication in Organizations. *The Review of Economic Studies 88*(2), 574–609.

Bertrand, M., R. Burgess, A. Chawla, and G. Xu (2020). The glittering prizes: Career incentives and bureaucrat performance. *Review of Economic Studies 87*, 626–655.

Besley, T. and T. Persson (2010). State capacity, conflict, and development. *Econometrica 78*(1), 1–34.

Blair, G., J. M. Weinstein, F. Christia, E. Arias, E. Badran, R. A. Blair, A. Cheema, A. Farooqui, T. Fetzer, G. Grossman, D. Haim, Z. Hameed, R. Hanson, A. Hasanain, D. Kronick, B. S. Morse, R. Muggah, F. Nadeem, L. L. Tsai, M. Nanes, T. Slough, N. Ravanilla, J. N. Shapiro, B. Silva, P. C. L. Souza, and A. M. Wilke (2021). Community policing does not build citizen trust in police or reduce crime in the global south. *Science 374*(6571).

Blanes i Vidal, J. and T. Kirchmaier (2018). The effect of police response time on crime clearance rates. *The Review of Economic Studies 85*(2), 855–891.

Blattman, C., D. P. Green, D. Ortega, and S. Tobón (2021). Place-based interventions at scale: the direct and spillover effects of policing and city services on crime. *Journal of the European Economic Association 00*, 1–30.

Boivin, R. and A. Gendron (2022). An experimental study of the impact of body-worn cameras on police report writing. *Journal of Experimental Criminology 18*(4), 747–764.

Bove, V. and E. Gavrilova (2017). Police officer on the frontline or a soldier? the effect of police militarization on crime. *American Economic Journal: Economic Policy 9*, 1–18.

Braga, A. A., L. M. Barao, G. M. Zimmerman, S. Douglas, and K. Sheppard (2020). Measuring the direct and spillover effects of body worn cameras on the civility of police–citizen encounters and police work activities. *Journal of Quantitative Criminology 36*, 851–876.

Braga, A. A., J. M. MacDonald, and J. McCabe (2022). Body-worn cameras, lawful police stops, and nypd officer compliance: A cluster randomized controlled trial. *Criminology 60*(1), 124–158.

Braga, A. A., W. H. Sousa, J. R. Coldren, and D. Rodriguez (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*.

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

Fenizia, A. (2022). Managers and productivity in the public sector. *Econometrica 90*(3), 1063–1084.

Ferraz, C., J. Monteiro, and B. Ottoni (2016). Monopolizing violence in ungoverned spaces : Evidence from the pacification of rio's favelas. *Preliminary Draft*.

Ferrazares, T. (2023). Monitoring police with body-worn cameras: evidence from chicago.

Journal of Urban Economics, 103539.

Frederiksen, A., L. B. Kahn, and F. Lange (2020). Supervisors and performance management systems. *Journal of Political Economy 128*(6), 2123–2187.

Harris, M. C., J. Park, D. J. Bruce, and M. N. Murray (2017). Peacekeeping force: Effects of providing tactical equipment to local law enforcement. *American Economic Journal: Economic Policy 9*, 291–313.

Hoekstra, M. and C. Sloan (2022). Does race matter for police use of force? evidence from 911 calls. *American Economic Review 112*(3), 827–60.

Hyland, S. (2018). *Body-worn cameras in law enforcement agencies, 2016*. US Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Kandel, E. and E. P. Lazear (1992). Peer pressure and partnerships. *Journal of political Economy 100*(4), 801–817.

Kapustin, M., T. Neumann, and J. Ludwig (2022). Policing and management. Working Paper 29851, National Bureau of Economic Research.

Kim, T. (2020). Facilitating police reform : Body cameras , use of force , and law enforcement outcomes. *Working Paper December*, 1–70.

Kim, T. (2021). Facilitating police reform : Body cameras , use of force , and law enforcement outcomes. *Working Paper May*, 1–70.

Lum, C., C. S. Koper, D. B. Wilson, M. Stoltz, M. Goodier, E. Eggins, A. Higginson, and L. Mazerolle (2020). Body-worn cameras' effects on police officers and citizen behavior: A systematic review. *Campbell Systematic Reviews 16*, 1–40.

Magaloni, B. (2019). How body-worn cameras affect the use of gunshots , stop-and searches and other forms of police behavior : A randomized control trial in rio de janeiro. *Stanford Poverty Violence Governance Lab*, 1–55.

Mas, A. and E. Moretti (2009). Peers at work. *American Economic Review 99*(1), 112–145.

Monteiro, J., L. Piquet, E. Fagundes, and J. Guerra (2022). Avaliacão do impacto do uso de câmeras corporais pela polícia militar do estado de são paulo. Technical report.

Ornaghi, A. (2019). Civil service reforms : Evidence from u.s. police departments. *Working Paper July*, 1–55.

Prendergast, C. (2021). 'drive and wave': The response to lapd police reforms after rampart. *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2021-25).

Rozema, K. and M. Schanzenbach (2019). Good cop, bad cop: Using civilian allegations to predict police misconduct. *American Economic Journal: Microeconomics 11*, 225–268.

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

UNODC (2014). Handbook on effective prosecution responses to violence against women and girls.

Weber, M. (1946). *Essays in Sociology*. Oxford University Press.

Williams, Morgan C, J., N. Weil, E. A. Rasich, J. Ludwig, H. Chang, and S. Egrari (2021). Body-worn cameras in policing: Benefits and costs. Working Paper 28622, National Bureau of Economic Research.

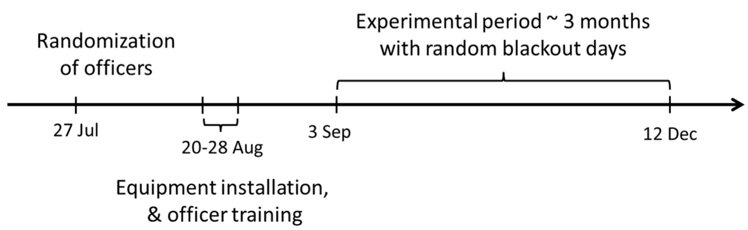
Xu, G. (2018). The costs of patronage: Evidence from the british empire. *American Economic Review 108*, 3170–3198.

Yokum, D., A. Ravishankar, and A. Coppock (2019). A randomized control trial evaluating the effects of police body-worn cameras. *Proceedings of the National Academy of Sciences of the United States of America 116*, 10329–10332.

Tables and Figures

Figure 1: Timeline and experimental design

Panel A. Experimental timeline

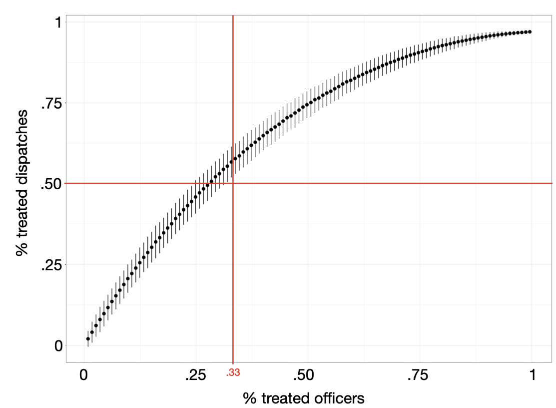


Panel B. Embedding variation into the police dispatch data

Pre-intervention baseline data...

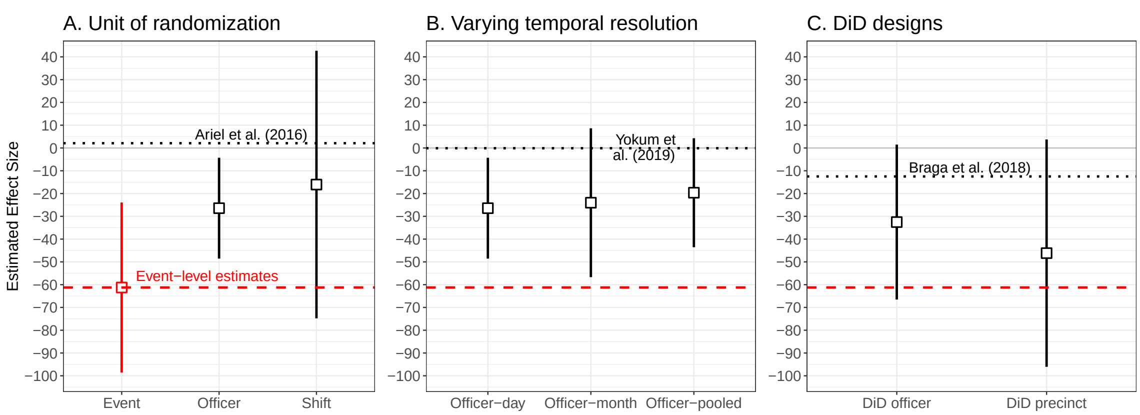
... was used to induce treatment allocation at the event level

			Treated		
Dispatch	Off. 1	Off. 2	A	A and B	Shift
1	A	B	1	1	1 (Regular)
2	A	C	1	1	1 (Regular)
3	A	D	1	1	1 (Regular)
4	B	C	0	1	1 (Regular)
5	B	D	0	1	1 (Regular)
6	C	D	0	0	1 (Regular)
		⋮			
25	A	B	1	1	0 (Blackout)
26	A	C	1	1	0 (Blackout)
27	A	D	1	1	0 (Blackout)
28	B	C	0	1	0 (Blackout)
29	B	D	0	1	0 (Blackout)
30	C	D	0	0	0 (Blackout)



Notes: This figure presents the experimental timeline. Panel A provides the timeline of the experiment that was conducted in 2018. The table in Panel B illustrates that a sparse (i.e., less than half) camera allocation to officers is required to induce a 50% treatment – 50% control variation at the dispatch level. The figure in Panel B shows the exact simulations that were conducted with actual police pre-intervention dispatch data. It shows that optimal treatment allocation would be achieved by allocating cameras to between one fourth and one third of police officers.

Figure 2: Comparing the distribution of effects with different designs and the literature



Notes: This figure presents the results on how the estimated treatment-effect sizes vary if we reanalyze the data using different commonly used evaluation strategies. Benchmark results from this paper exploit event-level variation and are presented in red. Estimates of effect sizes from reference studies in the literature using such designs are annotated as a horizontal dashed line and are closer to the most comparable estimate from our data. Panel A explores how changing the unit of randomization affects the results, exploring experimental variation between treated and control officers and between treated and control shifts. Panel B explores varying the temporal resolution in the aggregation of the outcome data while keeping the experimental variation of officers constant. Panel C explores two difference-in-differences models, the first exploring experimental variation between officers and the second exploring the spatially explicit implementation of BWCs.

Table 1: Effects of BWCs on reporting

	Event Recorded	Event Registered with No Info	Verbal Attrition/ Threat	Noise Complaint	Burglary	Assault	Domestic Violence	Share of Reports with Victims	Generated Investigative Report
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Main Effects									
Treated Event	24.043***	-2.770**	0.237	0.126	0.842*	0.709*	1.138***	2.783***	3.101**
	(1.873)	(1.239)	(0.645)	(0.550)	(0.427)	(0.388)	(0.351)	(0.805)	(1.190)
	p = 0.000	p = 0.164	p = 0.774	p = 0.870	p = 0.131	p = 0.096	p = 0.000	p = 0.004	p = 0.060
Panel B. Heterogeneity by Officer Rank									
Treated Event by Officer(s) with Soldier Rank	24.974***	-2.840**	1.130*	-0.047	1.106**	0.761*	1.269***	3.240***	3.247**
	(2.040)	(1.329)	(0.672)	(0.608)	(0.478)	(0.393)	(0.364)	(0.871)	(1.289)
	p = 0.000	p = 0.147	p = 0.200	p = 0.953	p = 0.065	p = 0.099	p = 0.001	p = 0.004	p = 0.049
Treated Event by Officer(s) with Higher-than-Soldier Rank	18.906***	-2.230	-3.073***	0.528	-0.164	0.699	0.619	1.503	2.201
	(2.244)	(2.417)	(1.103)	(1.173)	(0.730)	(0.725)	(0.616)	(1.349)	(2.077)
	p = 0.000	p = 0.634	p = 0.037	p = 0.693	p = 0.883	p = 0.394	p = 0.273	p = 0.391	p = 0.609
Treated Event by Officers of Both Types	24.788***	-3.524	-3.675***	1.503	-0.217	0.068	0.697	-0.108	3.957
	(2.678)	(3.054)	(1.214)	(1.260)	(1.102)	(1.200)	(0.722)	(1.880)	(3.217)
	p = 0.000	p = 0.426	p = 0.022	p = 0.371	p = 0.848	p = 0.947	p = 0.341	p = 0.958	p = 0.394
Panel C. Heterogeneity by Treatment Intensity									
Treated Event by 1 Camera	22.430***	-2.626**	0.143	0.520	1.045**	0.427	1.038***	2.485***	2.825**
	(1.881)	(1.272)	(0.642)	(0.559)	(0.454)	(0.384)	(0.342)	(0.875)	(1.275)
	p = 0.000	p = 0.186	p = 0.869	p = 0.522	p = 0.066	p = 0.321	p = 0.004	p = 0.014	p = 0.082
Treated Event by 2 or More Cameras	30.473***	-3.345*	0.609	-1.446	0.034	1.834**	1.535**	3.972***	4.201**
	(2.632)	(1.890)	(1.144)	(0.947)	(0.616)	(0.784)	(0.655)	(1.184)	(1.599)
	p = 0.000	p = 0.238	p = 0.617	p = 0.265	p = 0.968	p = 0.007	p = 0.007	p = 0.007	p = 0.098
Mean Dep. Var.	0.000	47.268	9.940	8.661	4.787	3.618	1.644	13.832	32.761
N	13274	13274	13274	13274	13274	13274	13274	13274	13274

Notes: This table documents the effects of BWCs on recording and reporting measures, including criminal typology. The dependent variables are "event recorded", indicating that the dispatch was partially or fully recorded using the BWC; "event registered with no info"indicating no criminal typology was recorded; the five most frequent criminal typologies reported: "verbal attrition/threat", "noise complaint", "burglary", "assault"and "domestic violence"; and "share of reports with victims" and "generated investigative report", when officers reported events to the Civil Police, who would proceed with investigations. Panel A presents the main results capturing the average intent-to-treat effect. Panel B explores rank heterogeneity of who wore the camera. Panel C investigates treatment intensity heterogeneity based on the number of officers wearing a camera in events. All dependent variables are multiplied by 100. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Shifts without a camera are excluded from the regression, which follows specification (1). Standard errors are clustered at the precinct-by-week level. *** p<0.01; ** p<0.05; * p<0.1. The randomization inference p-values are indicated below the standard errors.

Table 2: Effects of body-worn cameras on use of force

	Use of Force	Handcuff and/or Arrest	Contempt, Resistance and/or Disobedience	Negative Interaction Index
	(1)	(2)	(3)	(4)
Panel A. Main Effects				
Treated Event	-0.425***	-0.320	-0.263	-0.371**
	(0.157)	(0.471)	(0.196)	(0.149)
	p = 0.009	p = 0.584	p = 0.280	p = 0.030
Panel B. Heterogeneity by Officer Rank				
Treated Event by Officer(s) with Soldier Rank	-0.471***	-0.294	-0.390*	-0.444***
	(0.166)	(0.504)	(0.219)	(0.161)
	p = 0.007	p = 0.619	p = 0.121	p = 0.015
Treated Event by Officer(s) with Higher-than-Soldier Rank	-0.304	-0.520	0.095	-0.174
	(0.325)	(1.156)	(0.413)	(0.311)
	p = 0.311	p = 0.734	p = 0.871	p = 0.606
Treated Event by Officers of Both Types	-0.065	-0.001	0.604	0.154
	(0.541)	(1.707)	(0.641)	(0.502)
	p = 0.909	p = 1.000	p = 0.522	p = 0.813
Panel C. Heterogeneity by Treatment Intensity				
Treated Event by 1 Camera	-0.392***	-0.091	-0.207	-0.330**
	(0.143)	(0.533)	(0.205)	(0.145)
	p = 0.017	p = 0.884	p = 0.390	p = 0.048
Treated Event by 2 or More Cameras	-0.554*	-1.233	-0.487	-0.535**
	(0.298)	(0.804)	(0.357)	(0.265)
	p = 0.083	p = 0.209	p = 0.293	p = 0.111
Panel D. Heterogeneity by Ex-ante Event Risk Assessment				
Treated Event x Low Risk	-0.414**	-0.433	-0.381*	-0.403**
	(0.161)	(0.501)	(0.216)	(0.160)
	p = 0.004	p = 0.445	p = 0.095	p = 0.010
Treated Event x High Risk	-0.489	0.718	0.777	-0.070
	(0.703)	(1.474)	(0.811)	(0.640)
	p = 0.452	p = 0.636	p = 0.487	p = 0.928
Panel E. Heterogeneity by Baseline Use of force				
Treated Event x Below Median Use of Force	-0.265**	-0.265	-0.087	-0.206*
	(0.133)	(0.517)	(0.170)	(0.123)
	p = 0.079	p = 0.675	p = 0.714	p = 0.189
Treated Event x Above Median Use of Force	-1.059**	-0.527	-0.962*	-1.025**
	(0.436)	(0.827)	(0.541)	(0.402)
	p = 0.006	p = 0.590	p = 0.092	p = 0.009
Mean Dep. Var. Control Events	0.694	5.427	0.932	0.790
N	13274	13274	13274	13274

Notes: This table shows the effect of BWCs on use of force. The dependent variables are "use of force", which indicates if there was any deployment of physical, non-lethal (mechanical), or lethal force by the police, not considering use of handcuff or arrest; "arrest and/or the use of handcuffs", which is an indicator for if handcuffs were used or if any arrests made; "contempt, resist, and/or disobey" which is an indicator for if charges of contempt, disobedience, or resistance toward the police were registered; "Negative Interaction Index" is the standardized inverse-covariance weighted average of the three indicators in the group. Panel A presents the main results capturing the average intent-to-treat effect. Panel B explores rank heterogeneity of who wore the camera. Panel C investigates treatment intensity heterogeneity based on the number of officers wearing a camera in events. Panel D explores heterogeneity by the ex-ante risk level of the events, which characterizes an event as low risk if it had no weapons on the scene, if there were no injuries, if the suspect was not on site and if there was no material risk of general unrest. Panel E explores the heterogeneity by baseline use of force in areas of the municipalities. All dependent variables are multiplied by 100. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Shifts without camera were excluded from the regression. Standard errors are clustered at the precinct-by-week level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$. The randomization inference p-values are indicated below the standard errors.

Table 3: Testing for endogenous allocation of BWC to Events

	High Baseline Use of Force	Latitude	Longitude	High Baseline Income	Time to Dispatch (Minutes)	Time to Dispatch Greater than 5 min.	Active Policing	Telephone Initiated Dispatch	High Ex-Ante Risk
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated Event	-0.170	0.001	0.002**	1.726	-1.719	-1.348	-0.201	0.235	-0.341
	(0.896)	(0.001)	(0.001)	(1.313)	(1.269)	(1.100)	(0.573)	(0.574)	(0.587)
	p = 0.869	p = 0.708	p = 0.374	p = 0.541	p = 0.223	p = 0.275	p = 0.838	p = 0.812	p = 0.586
Mean Dep. Var. Control Events	20.080	-27.468	-48.787	48.639	10.701	43.868	7.637	92.180	10.378
N	13274	13274	13274	13274	13274	13274	13274	13274	13274

Notes: This table presents tests for the characteristics of the dispatch that could suggest endogenous allocation with respect to treatment assignment. The dependent variables are: high baseline use of force, which indicates whether the event happened in census tracts with an above median baseline use of force; latitude and longitude, both measured in degrees; high baseline income, which indicates whether the event happened in census tracts with an above median income in the baseline; time to dispatch, a measure of the interval between communication and dispatch arrival in minutes and a dummy to whether this interval was higher than five minutes; active policing, a dummy indicating if police self-initiated the event rather than being dispatched to it; telephone-initiated dispatch, which is an indicator for if the event was communicated through the telephone central; and high ex-ante risk, which is a measure of ex-ante risk that characterizes an event as low risk if it had no weapons on the scene, if there were no injuries, if the suspect was not on site, and if there was no material risk of general unrest. Sample includes all events in the experimental period and excluded blackout shifts. Specification includes police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Standard errors are clustered at the precinct-by-week level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$. The randomization inference p-values are indicated below the standard errors.

Online Appendix

"Monitoring technology: the impact of body-worn cameras on citizen-police interactions"

For Online Publication

Barbosa, Fetzer, Soto-Vieira, and Souza

December 14, 2023

A Implementation Details

Preparations and randomization We obtained the full roster of citizen-facing police officers in the beginning of July 2018. Officers without citizen-facing duties, such as those with administrative roles, were not considered eligible for camera use. We used pre-intervention dispatch data to validate if the list that was sent to us had contained all citizen-facing police officers. We then confirmed that there was no self-selection of officers into the study sample.

During this period, all tests with the cameras and docking stations were conducted to ensure that the information necessary for the experiment was correct and to minimize technical issues during the experiment period. Prior to the start of the experiment, all officers were briefed how to use the equipment, and how to adjust standard operating procedures allowing for the use of a camera (in particular, it was made clear that officers were required to verbally communicate to citizens that the events were being recorded). Importantly, *all* officers were briefed — irrespective of their treatment status — to avoid the briefing itself confounding the BWC treatment effects.

The implementation timeline is depicted in the Figure 1 above. We randomized officers and blackout days on July 7, 2018. Shift-level treatment allocation was randomized before the start of the experiment, but we only communicated to the police precincts on the preceding evening through dedicated WhatsApp groups established for this purpose. This was supposed to avoid the potential for the endogenous selection of any aspect of the policing activity with respect to the anticipation of blackout days. Importantly, the blackout applied to officers starting their shifts. That is, officers already on duty at midnight of the start of the blackout would continue to use their cameras until the end of their shift; conversely, any shift that started during blackout, even if it continued beyond the blackout, would not be recorded. This feature was necessary for logistic reasons: the police deemed it not practical or desirable to interfere in the apparatus of the dispatch units after they had left the precinct headquarters.

Intervention step by step Once the experiment period started, the intervention progressed as follows. At the start of their shifts, treated officers would obtain their camera, from the armory section of the police precincts — from where they obtained their gun, radio, and other equipment for regular and special use. The armory sections are usually very secluded and considered to be a high-security environment due to the nature of the material that is stored therein, and only a few high-ranking officers have access to those rooms. Importantly, the docking stations, which both downloaded the videos at the end of every shift and recharged the cameras, were located inside the armory rooms. This ensured that not only the equipment was maintained, regularly inspected, and kept in good working order throughout the experiment, but also ensured that docking stations and cameras themselves were not interfered with or violated during the experiment.

The docking stations were remotely accessible from the PMSC headquarters. Videos were stored locally for 30 days and transferred to the central HQ on demand due to bandwidth issues. The research team established routines to consolidate the cameras' automatic logs in a central database. In this way, it was possible to observe if a given dispatch generated a video recording, as well as the corresponding docking station and filenames. After finishing their shift, police officers would hand back their cameras to the armory section, which would then be docked in the station and readied for their next use. This recycling process usually took between four and six hours for a full battery charge that lasted at least eight hours in continuous regular use.

On the preceding night before control shifts, the research team would message the officers responsible for the armory sections in each police precinct telling them to **not** give cameras to treated officers. So the officers who started their shifts on the blackout day would receive from the armory all their equipment but the cameras.

On any given day, dispatch units would be composed by on average two officers. If any of those was assigned to wear a camera at the officer-level randomization, this dispatch (as well as the events they tended to) was classified as a treatment one. Thus, the average treatment effect of BWC implementation over police activity and police-citizen interactions was identified by comparing events attended by dispatches with at least one officer assigned to wear a camera with events with no officers with a camera.

As for blackout shifts, all treated officers were not allowed to wear cameras. Therefore, we can compare events attended by treated dispatches with events attended by control dispatches in days in which no treatment officer was allowed to wear cameras, allowing us to identify if the effects persisted when the treatment technology was not present. Importantly, the dispatch operators were blind to whether dispatch units were manned with officers wearing a BWC. This prevented the endogenous allocation of dispatch calls to be recorded (or, conversely, to avoid being recorded).

B The BWC Literature and Use of Force

Experiments on the effects of BWCs on use of force do not consistently show that cameras effectively work to decrease excessive use of force, as shown by Lum et al. (2020) and Williams et al. (2021). Appendix Table A7 lists the main BWC papers in the literature.[1] We include their main features, e.g., the number of citations, country, sample size, share of treated units, and whether any effect of use of force was detected. As can be seen, the literature is not conclusive on BWCs' effects on use of force. We argue below that most of the experiments with BWCs were potentially affected by methodological issues which attenuated the cameras' effects.

We start in Panel A with the studies that allocated cameras on the basis of shifts. All of the seven papers listed were RCTs, and they allocated cameras on the basis of treatment and control shifts. We argue that this design is potentially problematic, as a given officer may be allocated to both a treatment and a control shift over time. This may be an important SUTVA assumption violation if, for example, officers alter their behavior after using a camera, e.g., through learning, or if there are across-officers spillover effects (Ariel et al., 2017). Out of the seven studies that used a shift analysis, five had use of force as an outcome and only one found statistically significant results (at the 5% level) that suggest that BWCs affect use of force. Ariel et al. (2015) conducted the first experiment on BWCs and it is by far the most cited paper in the literature. The shifts were randomized

[1] This is not intended as a literature review but a selective and partial read on the studies that we found to be most prominent in the literature.

to be conducted with and without cameras, and the results suggest that BWCs reduce use of force by the police. However, these effects were barely significant at the 10% level. Following that, Ariel et al. (2016b) repeated the same design across multiple sites, and the results show null effects of BWCs on use of force. Ariel et al. (2016a) suggested that one potential explanation for these muted results stemmed from compliance with the protocol. They showed that use-of-force rates were higher in sites where the compliance with the protocol was lower, and vice-versa. Magaloni (2019) did not find any effects of BWCs on use of force, and the experiment faced issues with low compliance as well. With an experiment in the UK, Henstock and Ariel (2017) used shift randomization and found that BWC were effective in reducing use of force, in particular physical restraint and non-compliant handcuffing.

We move to officer-centric designs in Panel B. The literature shifted to officer-level allocation to ensure officers were always in the same assigned group throughout the duration of the experiment. This design also presents its challenges. First, contamination is a substantial concern: among the officer-centric papers we could identify, 60% routinely had more than one police officer per dispatch, which can mechanically result in contamination between officers if both a treated and a control officer are in the same dispatch. Moreover, all officer-level studies treated half of the police officers, which resulted in a much higher share of treated events — if an event was considered as treated if one or more cameras were present — given that most dispatches are tended by more than one police officer. In our data, simulations show that treating half of the officers would lead to 75% of the events being treated (see Panel B of Figure 1), leading to a considerably smaller control group and potentially undermining the statistical power. A study by Braga et al. (2020), who used officer-level randomization combined with spatial selection of districts, indeed shows evidence of large contamination from treated officers to control officers. Finally, some papers only included in the experiment officers that volunteered to wear a BWC (Jennings et al., 2015; Ready and Young, 2015; White et al., 2017; Jennings et al., 2017; Headley et al., 2017; Braga et al., 2018). This can introduce self-selection bias and compromise the identification of the effects. Taken together, these design characteristics can result in muted estimated effects of BWCs on police operations.

Finally, Panel C listed papers that used of spatially explicit empirical designs. Out of 12 studies, only four looked at use of force as an outcome, and only Kim (2021) and Ferrazares (2023) found evidence of the impact of BWCs. Kim (2021) used the variation in the timing of the adoption across US agencies to assess the effects of BWCs on a national level using a difference-in-differences with two-way fixed effects. While this strategy does not have to deal with the spillover that can occur between officers, it relies on the strong identifying assumption that adoption timing is independent of agency characteristics. Similarly, Miller and Chillar (2021) explored the staggered adoption of BWCs to study the effects on fatalities that arise from citizen-police interactions. While this is not the typical use-of-force outcome considered by us and the literature, it represents an extreme and infrequent case of when escalation unfolds. Bollman (2021) studied the effects of BWCs on court outcomes, also using a spatially explicit difference-in-difference. She found a significant reduction in new case fillings for offenses initiated during a citizen-police interaction, suggesting an improvement in these encounters.

Overall, some papers in this panel did not follow rigorous program evaluation techniques, and some did not even perform statistical inference methods. Nonetheless, meta-analyses of the existing studies have found no statistically significant effect of BWCs on use of force, even though the point estimate is negative (Lum et al., 2020; Williams et al., 2021).

C PAP Registration

Submission history Our initial study design was pre-registered on the "Evidence in Governance and Politics" (EGAP) repository as part of a broader project on the Metaketa IV round of funding that analyzed the effects related tocommunity policing. Later in 2020, the EGAP repository fully migrated with the OSF repository and can now be accessed through the link `https://osf.io/yzpva/`. File dates in the OSF system refer to the migration date, not the original date we submitted to the EGAP registry. The pre-analysis plan (PAP) associated with this project was registered in November 2018, before we had access to most of the experimental data.[2]

We had access to the majority of the data after substantial delays in December 2019. We registered an update to the PAP in January 2020.[3] The updates from the first version are not relevant to this project, as they mostly pertain to the parallel study on the effects of the community policing program. We further amended the PAP, including a specification appendix specifically for this project, before we undertook any data analysis in June of 2020.[4] The analysis was also registered at the AEA Registry with registration code `AEARCTR-0007785`. We did so when we decided to submit the paper to AEA journals as mandated by editorial guidelines.[5]

Hypotheses In the November 2018 PAP, we registered the hypotheses to be tested that we reproduce here in Appendix Table A9. Our understanding at that point was that we would be able to distinguish which officer took specific actions within the dispatch. For example, we believed we would obtain data regarding which officer in the dispatch was responsible for use of force and who conducted arrests. We postulated the hypotheses based on this understanding. We later learned that it was not possible to distinguish which officer in a dispatch had been responsible for each outcome. Our outcome data is instead at the level of the *dispatch*, not the officer within the dispatch. Following on the example above, we observe if there was use of force or arrests at the level of the dispatch, but not the specific officer who undertook those actions. This made it impossible to test `H1-a` and `-b`, `H2-a` and `-b`, `H3-a` and `-b`, and `H4-a` and `-b`. Moreover, the low quality of the citizen complaints dataset made it impossible to test `H2` altogether. Our definition of the treatment follows the hypotheses `H1-c` and `H3-c` and extrapolates the same definition to be able to test `H4`.

Outcomes The use-of-force outcomes were registered in the PAP. Outcomes of the reporting margin are understood as part of our own exploratory analysis. Nonetheless, given the effect sizes detected (see Table 1), we believe them to be

[2] https://osf.io/j2p5y/

[3] Available at https://osf.io/j2p5y/.

[4] Available at https://osf.io/f923e/.

[5] From January 2018, the submission policy to AEA Journals makes mandatory registration in the AEA RCT registry.

of high importance for the understanding of the impact of the policy. The fact that the increases in reporting of certain types of crimes are larger for domestic violence and, to a lesser extent, burglary and assault are important findings, and for this reason we opted to include in the main tables.

Specifications The specifications we use in the paper are largely consistent with those that were registered in the PAP. Below we detail and explain the reasoning behind minor differences in some specifications. In all cases, the versions in the paper and as registered in the PAP produce very similar results. The latter are available upon request.

We reproduce Equation (1) from the PAP below

$$y_{ibdw} = \beta \times \text{Treated Event}_i + \eta_{bw} + \tau_t + z_{ibdw} + \epsilon_{ibdw} \tag{8}$$

where η_{bw} precinct-by-week fixed effect, τ_d is day-of-the-week fixed effect, and z_{ibdw} is the number of officer fixed effects. For convenience, the notation and subscripts were harmonized with those used elsewhere in the paper. In the main specification, Equation (1), we added stratification bins to account for the stratified random assignment of cameras to officers. This is necessary to account for the fact that the camera assignment is random, conditional on the stratification bin. Results are shown in Panel A of Tables 1 and 2. The inclusion of officer stratification bin fixed effects was also reproduced in the blackout specification, corresponding to the PAP Equation (4). The event risk pre-assessment by the police was also pre-registered within the paragraph for heterogeneities. In line with the PAP, we also explore the effects of treatment intensity in Panel C. Other specifications registered in the PAP pertain to the event-study design and variations of the main equation using less granular variation aggregated at the officer and precinct levels. We use those specifications in Section 4 and Appendix Section B, as they allow us to compare the results with previous studies in the literature which implemented similar designs. Once more, we replicated all the analysis with the exact versions in the PAP and found results that are very consistent with those reported in the paper, and which are available upon request.

References

Ariel, B. (2016a). Increasing cooperation with the police using body worn cameras. *Police Quarterly Vol. 19(3)*, 326–362.

Ariel, B. (2016b). Police body cameras in large police departments. *The Journal of Criminal Law & Criminology 106*, 729–768.

Ariel, B., W. A. Farrar, and A. Sutherland (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*, 509–535.

Ariel, B., A. Sutherland, D. Henstock, J. Young, P. Drover, J. Sykes, S. Megicks, and R. Henderson (2016a). Report: increases in police use of force in the presence of body-worn cameras are driven by officer discretion: a protocol-based subgroup analysis of ten randomized experiments. *Journal of Experimental Criminology 12*, 453–463.

Ariel, B., A. Sutherland, D. Henstock, J. Young, P. Drover, J. Sykes, S. Megicks, and R. Henderson (2016b). Wearing body cameras increases assaults against officers and does not reduce police use of force: Results from a global multi-site experiment. *European Journal of Criminology 13*, 744–755.

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

Ariel, B., A. Sutherland, D. Henstock, J. Young, J. Sykes, S. Megicks, and R. Henderson (2017). "contagious accountability" : A global multisite randomized controlled trial on the effect of police body-worn cameras on citizens' complaints against the police. *CRIMINAL JUSTICE AND BEHAVIOR 44*, 293–316.

Bennett, R. R., B. Bartholomew, and H. Champagne (2019). Fairfax county police department's body-worn camera pilot project: an evaluation.

Bollman, K. (2021). The effects of body-worn cameras on policing and court outcomes: Evidence from the court system in virginia. *Working Paper November*.

Braga, A. A., L. M. Barao, G. M. Zimmerman, S. Douglas, and K. Sheppard (2020).

Measuring the direct and spillover effects of body worn cameras on the civility of police–citizen encounters and police work activities. *Journal of Quantitative Criminology 36*, 851–876.

Braga, A. A., W. H. Sousa, J. R. Coldren, and D. Rodriguez (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*.

Ferrazares, T. (2023). Monitoring police with body-worn cameras: evidence from chicago. *Journal of Urban Economics*, 103539.

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

Hedberg, E. C., C. M. Katz, and D. E. Choate (2017). Body-worn cameras and citizen interactions with police officers: Estimating plausible effects given varying compliance levels. *Justice Quarterly 34:4*, 627–651.

Henstock, D. and B. Ariel (2017). Testing the effects of police body-worn cameras on use of force during arrests: A randomised controlled trial in a large british police force. *European Journal of Criminology 14*, 720–750.

Jennings, W. G., L. A. Fridell, M. Lynch, K. K. Jetelina, and J. M. R. Gonzalez (2017). A quasi-experimental evaluation of the effects of police body-worn cameras (bwcs) on response- to-resistance in a large metropolitan police department a quasi-experimental evaluation of the effects of police. *Deviant Behavior 38*, 1332–1339.

Jennings, W. G., M. D. Lynch, and L. A. Fridell (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*, 480–486.

Katz, C. M., D. E. Choate, J. R. Ready, and L. Nuno (2014). Evaluating the impact of officer worn body cameras in the phoenix police department. *Center for Violence Prevention & Community Safety, Arizona State University*, 1–43.

Kim, T. (2021). Facilitating police reform : Body cameras , use of force , and law enforcement outcomes. *Working Paper May*, 1–70.

Koslicki, W. M., D. A. Makin, and D. Willits (2020). When no one is watching: evaluating the impact of body-worn cameras on use of force incidents. *Policing and Society 30*, 569–582.

Lum, C., C. S. Koper, D. B. Wilson, M. Stoltz, M. Goodier, E. Eggins, A. Higginson, and L. Mazerolle (2020). Body-worn cameras' effects on police officers and citizen behavior: A systematic review. *Campbell Systematic Reviews 16*, 1–40.

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

Magaloni, B. (2019). How body-worn cameras affect the use of gunshots , stop-and searches and other forms of police behavior : A randomized control trial in rio de janeiro. *Stanford Poverty Violence Governance Lab*, 1–55.

Miller, J. and V. F. Chillar (2021). Do police body-worn cameras reduce citizen fatalities? results of a country-wide natural experiment. *Journal of Quantitative Criminology*, 1–32.

Mitchell, R. J., B. Ariel, M. E. Firpo, R. Fraiman, F. del Castillo, J. M. Hyatt, C. Weinborn, and H. B. Sabo (2018). Measuring the effect of body-worn cameras on complaints in latin america: The case of traffic police in uruguay. *Policing: An International Journal 41*, 510–524.

Monteiro, J., L. Piquet, E. Fagundes, and J. Guerra (2022). Avaliacão do impacto do uso de câmeras corporais pela polícia militar do estado de são paulo. Technical report.

Morrow, W. J., C. M. Katz, and D. E. Choate (2016). Assessing the impact of police body-worn cameras on arresting , prosecuting , and convicting suspects of intimate partner violence. *Police Quarterly 19*, 303–325.

Peterson, B. E., L. Yu, N. L. Vigne, and D. S. Lawrence (2018). The milwaukee police department's body-worn camera program evaluation findings and key takeaways. *Urban Institute May*, 1–11.

Ready, J. T. and J. T. N. Young (2015). The impact of on-officer video cameras on police-citizen contacts: findings from a controlled experiment in mesa , az. *J Exp*

Criminol 11, 445–458.

Wallace, D., M. D. White, J. E. Gaub, and N. Todak (2018). Body-worn cameras as a potential source of depolicing: testing for a camera-induced passivity. *Criminology 56*, 481–509.

White, M. D., J. E. Gaub, and N. Todak (2017). Exploring the potential for body-worn cameras to reduce violence in police – citizen encounters. *Policing*, 1–11.

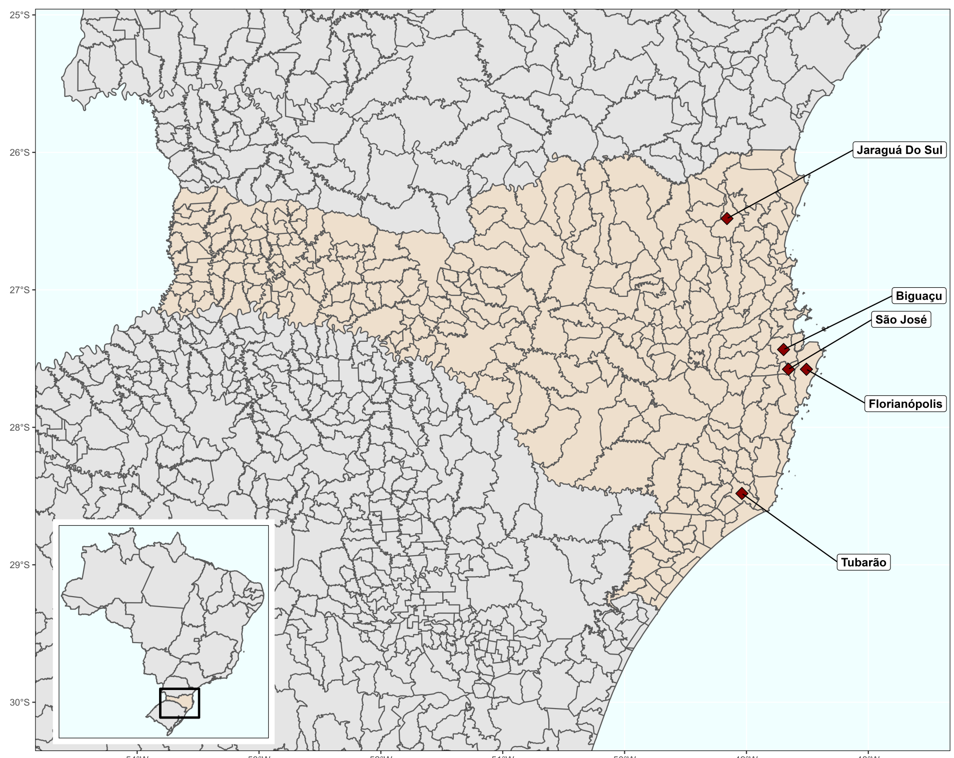
Williams, Morgan C, J., N. Weil, E. A. Rasich, J. Ludwig, H. Chang, and S. Egrari (2021). Body-worn cameras in policing: Benefits and costs. Working Paper 28622, National Bureau of Economic Research.

Yokum, D., A. Ravishankar, and A. Coppock (2019). A randomized control trial evaluating the effects of police body-worn cameras. *Proceedings of the National Academy of Sciences of the United States of America 116*, 10329–10332.

Çubukçu, S., N. M. Sahin, E. Tekin, and V. Topalli (2021). Body-worn cameras and adjudication of citizen complaints of police misconduct. *NBER Working Paper No. 29019 July*.

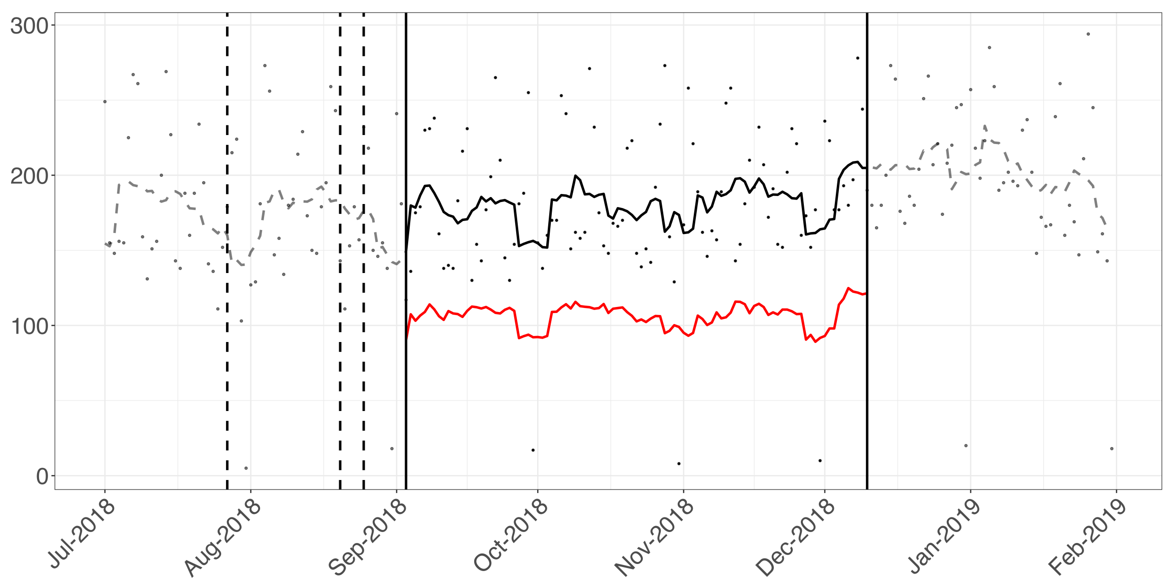
Appendix Tables and Figures

Figure A1: State of Santa Catarina and the experimental sites where the BWC intervention was implemented



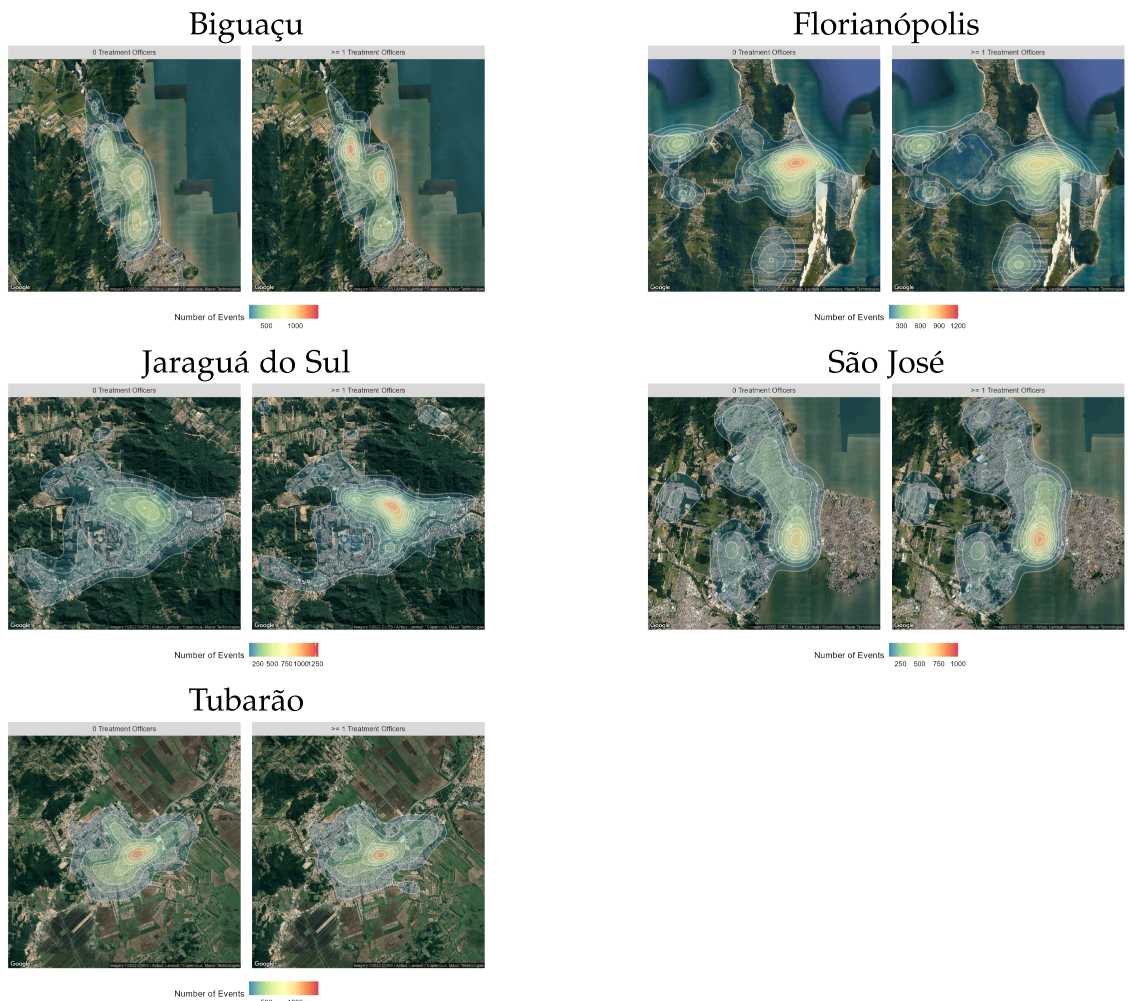
Notes: This figure presents the experimental sites where the experiment took place. These are the catchment areas of the 24^{th} Police Precinct in Biguaçu, the 21^{st} in Florianópolis, the 14^{th} in Jaraguá do Sul, the 7^{th} in São José, and the 5^{th} in Tubarão.

Figure A2: Time series of treated and control events



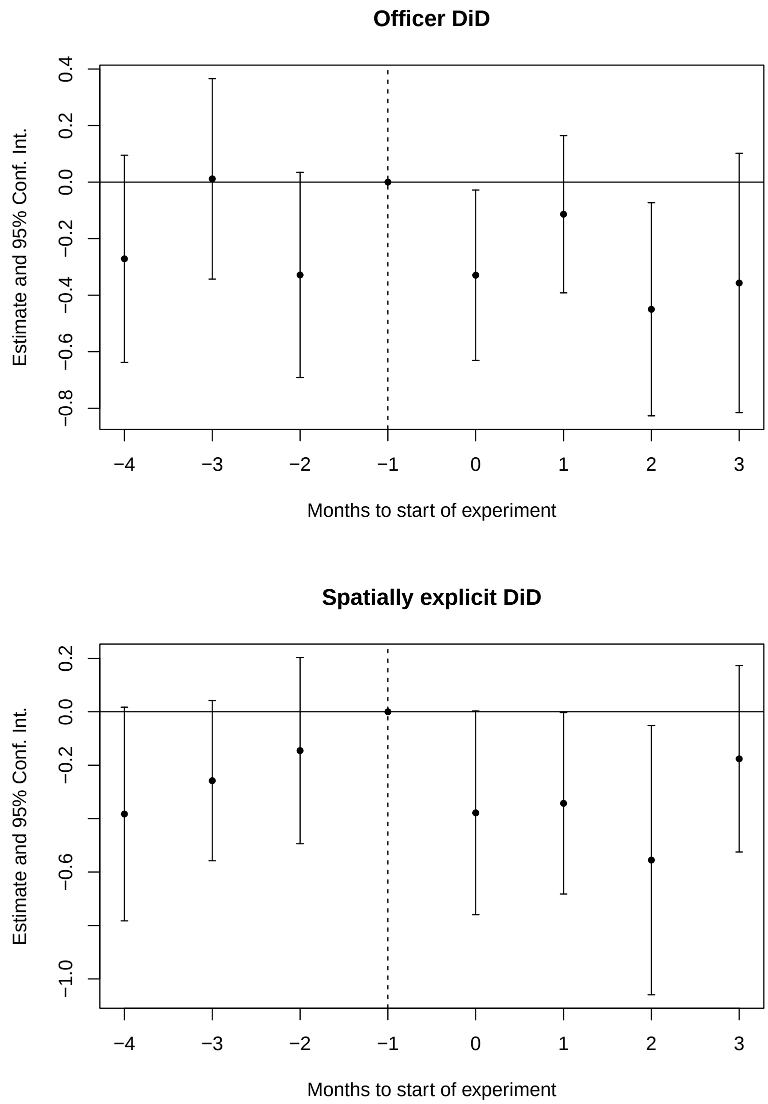
Notes: This figure presents the time series of the events by treatment status. The solid lines measure the seven-day moving average of the number of treated events (in red) and the overall number of events, illustrating that, on average, 50% of events had the presence of a BWC. Each dot measures the number of events in a given day. The dashed vertical lines represent important dates of the experiment design, as illustrated in Figure 1.

Figure A3: Spatial distribution of treatment and control dispatches



Notes: This figure presents kernel density estimates of the spatial distribution of the treatment and control events across the five cities that were part of the experiment. It highlights that the spatial distribution of both the treatment and control event dispatches were very similar throughout and highlights the different topographies of the study area.

Figure A4: The effects of BWCs on use of force: two event-studies



Notes: These figures show event-study estimates of the effects of BWCs. The first explores the variation between treated and control officers and the second the variation between treated and control precincts. The point estimates are the coefficient of the treatment unit interacted with period dummies. The officer-level DiD regression uses officer, precinct, week and weekday fixed effects and the standard errors are clustered at the precinct-week level. The spatially explicit DiD uses precinct, week, and weekday fixed effects and the standard errors are clustered at the precinct-week level.

Table A1: Summary statistics of study sites

	Biguaçu	Florianópolis	Jaraguá Do Sul	São José	Tubarão	SC average
Panel A. Socioeconomic Characteristics						
Population	58,206	421,240	143,123	209,804	97,235	18,468
						(42,990)
Urban (%)	0.904	0.964	0.932	0.989	0.907	0.599
						(0.231)
Income	1,208.22	2,578.28	1,586.99	1,692.74	1566.36	1,127.35
						(236.72)
White (%)	0.836	0.846	0.864	0.844	0.908	0.829
						(0.103)
Primary school or less (%)	0.292	0.623	0.594	0.574	0.656	0.571
						(0.082)
High school or less (%)	0.797	0.959	0.941	0.937	0.965	0.940
						(0.025)
Water access (%)	0.995	0.999	0.996	0.999	0.995	0.987
						(0.023)
Computer (%)	0.490	0.727	0.585	0.661	0.569	0.365
						(0.110)
Internet (%)	0.391	0.650	0.427	0.564	0.462	0.248
						(0.101)
Panel B. Violence and Use-of-Force Incidence						
Use-of-Force Incidents	23	52	34	62	22	-
Crime Events	739	2135	2622	3097	1309	-
Homicide Rate per 100k	22.9	17.16	5.38	16.9	9.65	-
Use of Force - Yearly Rate per 100k	106.90	33.39	64.27	79.95	61.21	-
Crime Events - Yearly Rate per 100k	3,435.05	1,371.276	4,956.55	3,993.78	3,642.28	-

Notes: This table presents socio-demographic characteristics and baseline violence across the five study sites and the average in Santa Catarina State. The sociodemographic data comes from 2010 IBGE Census, the homicide rate from the 2016 IPEA Atlas da Violência, and use-of-force and crime-events incidence from the authors' calculations using PMSC data from March to July 14, 2018. Income is in Brazilian reais per month. Standard errors are in parentheses.

Table A2: Effects of BWCs on reporting — blackout specifications

	Event Recorded	Event Registered with No Info	Verbal Attrition/ Threat	Noise Complaint	Burglary	Assault	Domestic Violence	Share of Reports with Victims	Generated Investigative Report
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated Event on Blackout Day	3.343***	-6.631***	3.193***	0.178	-0.940	0.664	0.254	0.435	3.233*
	(0.785)	(1.802)	(0.971)	(0.795)	(0.845)	(0.606)	(0.513)	(1.220)	(1.908)
	p = 0.272	p = 0.000	p = 0.000	p = 0.813	p = 0.088	p = 0.121	p = 0.499	p = 0.659	p = 0.049
Mean Dep. Var.	0.000	48.937	7.237	8.329	6.433	3.848	1.895	17.231	34.406
N	4391	4391	4391	4391	4391	4391	4391	4391	4391

Notes: This table documents the effects of BWCs on recording and reporting measures, including criminal typology, in control shifts (when BWCs were not handed out to officers). The dependent variables are “event recorded”, indicating that the dispatch was partially or fully recorded using the BWC; “event registered with no info”indicating no criminal typology was recorded; the five most frequent criminal typologies reported: “verbal attrition/threat”, “noise complaint”, “burglary”, “assault”and “domestic violence”; and “share of reports with victims” and “generated investigative report”, when officers reported events to the Civil Police, who would proceed with investigations. All dependent variables are multiplied by 100. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. The sample only includes shifts without a camera and the regression follows specification (1). Standard errors are clustered at the precinct-by-week level. *** p<0.01; ** p<0.05; * p<0.1. The randomization inference p-values are indicated below the standard errors.

Table A3: Effects of BWCs on use of force — blackout specifications

	Use of Force	Handcuff and/or Arrest	Contempt, Resistance and/or Disobedience	Negative Interaction Index
	(1)	(2)	(3)	(4)
Treated Event on Blackout Day	0.077	0.084	-0.001	0.051
	(0.234)	(0.809)	(0.310)	(0.178)
	p = 0.632	p = 0.879	p = 0.997	p = 0.768
Mean Dep. Var. Control Events	0.517	5.399	0.804	0.629
N	4391	4391	4391	4391

Notes: This table shows the effect of BWCs on use of force in shifts when cameras were not present. The dependent variables are "use of force", which indicates if there was any deployment of physical, non-lethal (mechanical), or lethal force by the police, not considering use of handcuff or arrest; "arrest and/or the use of handcuffs", which is an indicator for if handcuffs were used or if any arrests made; "contempt, resist, and/or disobey", which is an indicator for if charges of contempt, disobedience, or resistance toward the police were registered; "Negative Interaction Index" is the standardized inverse-covariance weighted average of the three indicators in the group. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. The sample only includes shifts without a camera. Standard errors are clustered at the precinct-by-week level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$. The randomization inference p-values are indicated below the standard errors.

Table A4: Testing for endogenous allocation of BWC to events — heterogeneity by officer rank

	High Baseline Use of Force	Latitude	Longitude	High Baseline Income	Time to Dispatch (Minutes)	Time to Dispatch Greater than 5 min.	Active Policing	Telephone Initiated Dispatch	High Ex-Ante Risk
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated Event by Officer(s) with Soldier Rank	-0.676	0.000	0.000	0.821	-1.859	-1.615	-0.310	0.314	-0.292
	(0.969)	(0.001)	(0.001)	(1.444)	(1.275)	(1.198)	(0.570)	(0.573)	(0.691)
	p = 0.541	p = 0.892	p = 0.913	p = 0.792	p = 0.238	p = 0.198	p = 0.744	p = 0.745	p = 0.674
Treated Event by Officer(s) with Higher-than-Soldier Rank	2.176	0.000	0.009***	6.961***	-1.475	-2.000	0.125	0.027	-0.112
	(1.556)	(0.001)	(0.002)	(2.208)	(1.288)	(2.045)	(1.281)	(1.299)	(1.250)
	p = 0.382	p = 0.907	p = 0.068	p = 0.211	p = 0.477	p = 0.560	p = 0.960	p = 0.993	p = 0.937
Treated Event by Officers of Both Types	0.560	0.006***	0.012***	0.659	-0.462	4.224	0.692	-0.535	-1.654
	(2.067)	(0.001)	(0.002)	(3.144)	(3.311)	(2.580)	(1.847)	(1.837)	(1.600)
	p = 0.825	p = 0.037	p = 0.026	p = 0.915	p = 0.846	p = 0.180	p = 0.826	p = 0.862	p = 0.396
Mean Dep. Var. Control Events	20.080	-27.468	-48.787	48.639	10.701	43.868	7.637	92.180	10.378
N	13274	13274	13274	13274	13274	13274	13274	13274	13274

Notes: This table presents tests for the characteristics of the dispatch that could suggest endogenous allocation with respect to treatment assignment. The dependent variables are defined as in Table 3. The sample includes all events in the experimental period and excluded blackout shifts. The specification includes police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Standard errors are clustered at the precinct-by-week level. *** p<0.01; ** p<0.05; * p<0.1. The randomization inference p-values are indicated below the standard errors.

Table A5: Correlation between event characteristics and officer rank

	High Risk (1)
Event with High-Ranking Officer	0.939
	(1.069)
Mean Dep. Var. Events Only with Soldiers	9.820
N	13274

Notes: This table shows the correlation between the presence of a high-ranking officer in an event and the ex-ante level of risk of an event. High-risk is the ex-ante risk assessment indicator used in Table 2. The specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Standard errors are clustered at the precinct-by-week level. *** p<0.01; ** p<0.05; * p<0.1.

Table A6: Sample Robustness

		Reporting Margin							
	Event Recorded	Event Registered with No Info	Domestic Violence	Share of Reports with Victims	Generated Investigative Report	Use of Force	Handcuff and/or Arrest	Contempt, Resistance and/or Disobedi-ence	Negative Interaction Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Main Results									
Treated Event	24.043***	-2.770**	1.138***	2.783***	3.101**	-0.425***	-0.320	-0.263	-0.371**
	(1.873)	(1.239)	(0.351)	(0.805)	(1.190)	(0.157)	(0.471)	(0.196)	(0.149)
Mean Dep. Var.	0.000	47.268	1.644	13.832	32.761	0.694	5.427	0.932	0.790
N	13274	13274	13274	13274	13274	13274	13274	13274	13274
Panel B. Including Blackout Days									
Treated Event	18.852***	-3.647***	0.899***	2.180***	3.100***	-0.307**	-0.171	-0.190	-0.268**
	(1.521)	(1.046)	(0.286)	(0.652)	(1.097)	(0.120)	(0.429)	(0.172)	(0.113)
Mean Dep. Var.	0.000	47.671	1.705	14.652	33.158	0.652	5.420	0.901	0.751
N	17665	17665	17665	17665	17665	17665	17665	17665	17665
Panel C. Two Officers - Modal Dispatch Size									
Treated Event	24.140***	-3.641**	1.102***	3.365***	4.036***	-0.222*	-0.158	-0.371*	-0.271**
	(1.930)	(1.416)	(0.336)	(0.897)	(1.328)	(0.114)	(0.474)	(0.205)	(0.125)
Mean Dep. Var.	0.000	49.147	1.575	13.173	30.569	0.416	3.807	0.810	0.557
N	9928	9928	9928	9928	9928	9928	9928	9928	9928
Panel D. At Most Four Officers									
Treated Event	24.061***	-3.024**	1.164***	2.947***	3.196**	-0.344***	-0.427	-0.285	-0.325**
	(1.859)	(1.268)	(0.339)	(0.848)	(1.236)	(0.128)	(0.433)	(0.177)	(0.125)
Mean Dep. Var.	0.000	47.572	1.619	13.600	32.335	0.595	5.153	0.856	0.698
N	12546	12546	12546	12546	12546	12546	12546	12546	12546

Notes: This table presents intention-to-treat specifications. The unit of observation is a police event. The dependent variables defined as in Tables 1 and 2. The specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Shifts without cameras are excluded from the regression. Standard errors are clustered at the precinct-by-week level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table A7: Characteristics of BWC Studies in the Literature

Paper	Year	# Citations	Country	RCT	BWC varies across:	T	C	Analysis Unit	Analysis sample size	Share of Treated Units	Avg # of officers per dispatch	UoF as outcome?	Effects on UoF	Empirical strategy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Shift-centric studies (7 studies)														
Ariel et al. (2015)	2014	633	US	Yes	Shift	489	499	Shift	988	0.495	1	Yes	Null	Poisson regression
Ariel et al. (2016b)	2016	287	Multisite	Yes	Shift	2,447	2,468	Shift	4,915	0.498	-	Yes	Increase	Cohen's d
Ariel et al. (2016a)	2016	218	Multisite	Yes	Shift	2,447	2,468	Shift	4,915	0.498	-	Yes	Mixed	Cohen's d
Ariel et al. (2017)	2017	243	Multisite	Yes	Shift	1,908	1,974	Shift	3,882	0.491	-	No	-	Cohen's d
Henstock and Ariel (2017)	2017	107	UK	Yes	Shift	215	215	Shift	430	0.5	1	Yes	Decrease	Odds-Ratio
Ariel et al. (2018)	2018	86	Multisite	Yes	Shift	3,637	3,697	Shift	7,334	0.495	-	No	-	Odds-Ratio
Magaloni (2019)	2019	2	BR	Yes	Unit-shift	16,390	18,642	Shift	35,032	0.468	1+	Yes	Null	OLS
Panel B: Officer-centric studies (11 studies)														
Jennings et al. (2015)	2015	299	US	Yes*	Officer	46	43	Officer	89	0.517	1+	Yes	Unclear	t-test
Ready and Young (2015)	2015	288	US	Yes*	Officer	50	50	Contact report	3,698	0.5	1+	No	-	Logit
White et al. (2017)	2017	167	US	Yes*	Officer	82	67	Officer*time	298	0.55	-	Yes	Null	DiD
Jennings et al. (2017)	2017	81	US	No*	Officer	60	60	Officer	120	0.5	-	Yes	Decrease	% change **
Headley et al. (2017)	2017	89	US	Yes*	Officer	26	25	Officer*time	102	0.51	-	Yes	Null	t-test
Braga et al. (2018)	2018	180	US	Yes*	Officer	218	198	Officer*time	832	0.524	1	Yes	Decrease	DiD
Peterson et al. (2018)	2018	54	US	Yes	Officer	252	252	Officer	504	0.5	-	Yes	Null	DiD
Wallace et al. (2018)	2018	107	US	Yes	Officer	82	67	Call-officer	228,220	0.55	1+	No	-	DiD
Yokum et al. (2019)	2019	91	US	Yes	Officer	1,189	1,035	Officer	1,922	0.535	-	Yes	Null	OLS
Koslicki et al. (2020)	2020	20	US	No	-	-	-	Officer * Month	-	-	-	Yes	Decrease	Time series analysis
Braga et al. (2020)	2020	23	US	Yes	Officer + District	140	141	Officer	562	0.498	1	Yes	Decrease	DiD
Panel C: Spatially explicit designs (12 studies)														
Katz et al. (2014)	2014	217	US	No	Area	1	1	Area	2	0.5	-	No	-	t-test
Morrow et al. (2016)	2016	149	US	No	Area	1	1	IPV events	2,063	-	-	No	-	t-test**
Ariel (2016a)	2016	109	US	No	District	1	5	Street segment	17,726	0.167	2	No	-	Unclear
Ariel (2016b)	2016	129	US	No	District	1	5	Call	924,457	0.167	-	Yes	Null	Odds-Ratio
Hedberg et al. (2017)	2017	209	US	No	Precinct	1	1	Incident	44,380	0.499	1+	No	-	GLM
Mitchell et al. (2018)	2018	19	UY	No	Region	5	14	Region	38	0.263	-	No	-	Time series analysis
Bennett et al. (2019)	2019	-	US	No	Squad areas	1	1	Squad * Week	Unclear	0.5	-	Yes	Null	Diff. in trends test
Kim (2021)	2021	11	US	No	Agencies	25	75	Agency * Month	100	-	-	Yes	Decrease	TWFE
Miller and Chillar (2021)	2021	10	US	No	Agencies	1,346	1,030	Agency * Year	10,448	-	-	No	-	DiD
Bollman (2021)	2021	-	US	No	Courts	70	33	Court * Quarter	4,141	-	-	No	-	TWFE
Çubukçu et al. (2021)	2021	9	US	No	District	-	-	Complaints	2,117			No		TWFE
Ferrazares (2023)	2023	1	US	No	District	-	-	District * Day	46,011	-	-	Yes	Decrease	TWFE

Notes: This table provides a non-exhaustive overview of some of the existing empirical literature on BWC. The overview does not claim to be comprehensive but has aimed to include all empirical studies evaluating BWCs across a broad range of fields from criminology to economics. In column (4), we indicate with * if the evaluation includes volunteer officers. The table focuses on the respective randomization design, the outcome measurement approach, empirical strategy employed and whether effects on use of force (UoF) have been identified. Empirical strategies that do not perform statistical inference have **. Not in all cases was it possible to infer all required input and only two papers have replication data available.

Table A8: The effects of BWC adoption on the number of events

	Number of Events	Number of Telephone Initiated Events	Number of Officer Initiated Events
	(1)	(2)	(3)
Treated Precinct x Post	0.705	1.033	-0.226
	(0.762)	(0.726)	(0.180)
Mean Dep. Var. Control Precinct	19.163	17.145	2.000
N	7372	7372	7372

Notes: This table shows the effects of adopting BWCs on police activity at the precinct level. The data is at the precinct-day level, and there are 38 total precincts, of which five adopted BWC. The specifications include precinct, week and day of the week fixed effects. The regression uses number of officers in each precinct as weights. Standard errors are clustered at the precinct-by-week level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table A9: Registered Hypotheses in the Nov 2018 PAP

	Main Hypothesis	Sub-hypothesis
H1	BWC reduces use of force incidents	by a) officers wearing a camera; b) officers in the same patrol group as those wearing a camera; c) officers attending an event where at least one officer was wearing a camera.
H2	BWC reduce citizens complaints against officers	by a) officers wearing a camera; b) officers in the same patrol group as those wearing a camera; c) officers attending an event where at least one officer was wearing a camera.
H3	BWC reduce use of force incidents by police officers that had in the past	a) worn a camera; b) patrolled with officer that had worn a camera; c) attended an event where one officer was wearing a camera.
H4	BWC reduce dispatch time	a) wearing a camera; b) patrolling with an officer wearing a camera.