DISCUSSION PAPER SERIES

DP16578
(v. 2)

De-escalation technology: the impact of body-worn cameras on citizen-police interactions

Thiemo Fetzer, Pedro CL Souza, Daniel Barbosa and Caterina Vieira

DEVELOPMENT ECONOMICS

PUBLIC ECONOMICS

ISSN 0265-8003

De-escalation technology: the impact of body-worn cameras on citizen-police interactions

Thiemo Fetzer, Pedro CL Souza, Daniel Barbosa and Caterina Vieira

Discussion Paper DP16578
First Published 24 September 2021
This Revision 24 September 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

De-escalation technology: the impact of body-worn cameras on citizen-police interactions

Abstract

We provide experimental evidence that monitoring of the police activity through body-worn cameras reduces use-of-force, handcuffs and arrests, and enhances criminal reporting. Stronger treatment effects occur on events classified ex-ante of low seriousness. Monitoring effects are moderated by officer rank, which is consistent with a career concern motive by junior officers. Overall, results show that the use of body-worn cameras de-escalates conflicts.

JEL Classification: C93, D73, D74

Keywords: police citizen interaction, use-of-force, technology, field experiment

Thiemo Fetzer - thiemo.fetzer@gmail.com
University of Warwick and CEPR

Pedro CL Souza - pedroclsouza@gmail.com
University of Warwick

Daniel Barbosa - daniel.adrianocb@gmail.com
Pontifical Catholic University of Rio de Janeiro

Caterina Vieira - caterina.vieira@gmail.com
London School of Economics

Acknowledgements
We thank the Military Police of Santa Catarina and the Igarape Institute, and especially Ten Cel Tasca, Major Pablo, Major Vieira, Emile Badran and Barbara Silva for the outstanding collaboration and support. We thank Joana Monteiro and Cel Cabanas from the Military Police of Sao Paulo for attentive suggestions and comments in earlier versions of this work. This project was funded through EGAP Metaketa IV initiative.

De-escalation technology: the impact of body-worn cameras on citizen-police interactions*

Daniel AC Barbosa Thiemo Fetzer Caterina Soto
Pedro CL Souza

September 2021

Abstract

We provide experimental evidence that monitoring of the police activity through body-worn cameras reduces use-of-force, handcuffs and arrests, and enhances criminal reporting. Stronger treatment effects occur on events classified *ex-ante* of low seriousness. Monitoring effects are moderated by officer rank, which is consistent with a career concern motive by junior officers. Overall, results show that the use of body-worn cameras de-escalates conflicts.

Keywords: police citizen interaction, use-of-force, technology, field experiment
JEL Classification: C93, D73, D74

*Fetzer is at the University of Warwick. Barbosa is at PUC-Rio, Brazil. Souza is at Queen Mary University. Soto is at the London School of Economics. We thank the Military Police of Santa Catarina and the Igarapé Institute, and especially Ten Cel Tasca, Major Pablo, Major Vieira, Emile Badran and Bárbara Silva for the outstanding collaboration and support. We thank 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 EGAP Metaketa IV initiative, and it was pre-registered as part of the EGAP Metaketa with registration `20190411AB` `https://osf.io/uqv8j`. The project was registered in the AEA Registry as `AEARCTR-0007785`

1 Introduction

The idea that the state alone has the right use or authorize the use-of-force is one of its defining characteristics (Weber, 1946). For many citizens around the world, police forces are the primary visible representation of the state's monopoly on violence. Yet, the legitimacy and public confidence in the police is under strain worldwide especially in the wake of allegations of excessive use-of-force.[1] Police body-worn cameras (henceforth, BWC) have been hailed as a technological solution to increase scrutiny and oversight of the police. In this paper, we present experimental evidence showing that police BWCs effectively work to de-escalate police-citizen interactions, and improve the overall accuracy of police reporting.

We implemented a randomized controlled trial in a context that is more representative of the challenges faced by police forces in the Global South: the state of Santa Catarina in Brazil. Our experiment was designed to seed a random allocation of body-worn cameras at the granular police *dispatch* – the relevant unit of analysis in our study.[2] We view such dispatch-level data as the "natural" unit of analysis as it is the level at which citizen and police interactions unfold, and use-of-force and its (de)escalation may occur.

We find that BWC trigger both notable improvements in the accuracy of police reports and, in contrast with much of the existing literature, significant improvements in interactions between citizens and the police and significant declines in the use-of-force. For example, concerning reporting, dispatches treated with a camera present were 9.2% more often referred to the main investigative body, and police reports, on average, included 19.6% more victims. Importantly, 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 arrests –, was reduced by 44.2%.

We further document that the treatment effects are primarily concentrated in

[1]See 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]Throughout the paper, we also interchangeably refer to a dispatch as a police "event".

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, nor 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 in otherwise routine operations.

We next explore why cameras appear to work. In our design, both *whether* and *who* carries a camera is random. This allows us to study whether who carries the camera matters. We find evidence of stronger de-escalation effects and increased compliance with the polices' BWC standard operating procedures if the officer wearing the camera is relatively junior. This suggests that BWCs might work by empowering low-rank officers to monitor their higher-ranked peers, implying that dynamic incentives and career concerns may be important factors driving their effect. This peer monitoring effect goes beyond the usually suggested mechanisms of BWCs reducing negative interactions due to the improved monitoring of both citizen and police behavior that they enable.

Lastly, we also attempt to shed some light on why our results appear to stand in contrast with much of the existing literature, which has mostly found null or very muted effects of BWCs – in particular on use-of-force (see for example the meta-analysis by Lum et al., 2020). Naturally, the differences could simply arise because 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 has been almost exclusively the focus of the existing work).[3] Yet, we show that a more likely explanation for the failure of existing studies to identify effects is due to the research designs and, in particular, the outcome measurement and empirical evaluation strategies these studies adopt. In fact, our research design nests a broad class of commonly used evaluation strategies or outcome measurement approaches that have been employed across experimental BWC studies. This allows us to replicate our own study at coarser levels of analysis or when employing different empirical

[3] Magaloni (2019) marks an exception studying a BWC randomized controlled trial in a neighborhood of Rio de Janeiro, Brazil. Overall, they note very low compliance and little camera footage being produced.

strategies for evaluation. We find indeed that the estimated BWC treatment effects are much more muted or disappear altogether 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, in addition to the noise introduced in outcome measurement when moving from event-level to coarser designs, in combination are the likely culprits.

This paper also contributes to the literature that studies mechanisms that can prevent police misconduct (Shi, 2009; Rozema and Schanzenbach, 2019; Chassang and Padró i Miquel, 2019). Harris et al. (2017) show that acquiring tactical weapons has a positive effect on citizen-officer interactions, reducing both complaints against officers and assaults on officers. Relatedly, Owens et al. (2018) investigates the effects of training on improving citizen-officer interactions. We also contribute to the understanding of police interventions that aim to build trust or improve citizen relations and reduce crime. The meta-study Blair et al. (2020) finds no effects of community-policing intervention across different sites, including the state of Santa Catarina, Brazil. Magaloni et al. (2015) and Ferraz et al. (2016) show that another community policing program in Brazil known as UPP (*Pacifying Police Units*) had a positive effect decreasing violence but in the very specific context of territories dominated by drug trafficking gangs. Blattman et al. (2021) find little evidence that increased police presence and improved services more generally reduce crime in aggregate. Bove and Gavrilova (2017) show that militarized policing can deter street-level crime. Finally, we contribute to a broader debate on the productivity effects of monitoring of actions, employer-employee agency problems, and alignment of employees' incentives to that of the general organization, for which police officers are just one example. See Ornaghi (2019) on civil service reform and Bertrand et al. (2020) and Xu (2018) studying bureaucrats more broadly. Relatedly, Battiston et al. (2021) shows how career incentives play an important role in worker's decision to communicate and their productivity. Our results suggest that career incentives also play a role concerning the use of the body-worn camera and how police officers work.

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. Section 4 concludes and discusses policy implications of the experiment results.

2 Context, Intervention and Data

Brazil is one of the most violent countries in the world – in 2018, the homicide rate was 27.4 homicides per 100 thousand inhabitants compared to 5.0 and 1.2 in the US and the UK, respectively.[4] We implemented the BWC intervention in the state of Santa Catarina, Brazil. Santa Catarina exhibits a homicide rate that is three times higher than the US and 12 times higher than the UK. 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 manning 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 to be easily accessible from the police headquarters in Florianópolis and to represent a variety of settings in terms of socio-demographic characteristics and of baseline violence levels.[5]

Figure 1 provides an overview of the experimental design starting with the project timeline in Panel A. Panel B illustrates the two layers of randomization and how this induces variation at the dispatch level. Out of the roster of sworn police officers per precinct we obtained in July 2018, we randomly selected 1/3 of the officers to be in the treatment group and 2/3 to the control group across 40 stratification blocks.[6] Treated officers would always wear a camera if their 12-hour shift falls on days that – due to our second layer of randomization – were not selected to serve as *blackout days*. In every week during the twelve weeks of the experiment, two days were randomly selected to serve as blackout days with the randomization stratified by day of week providing us both within officer and between shift experimental design. Control officers were mandated not to wear a

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

[5]A map of the experimental locations is provided in Appendix Figure A1, while Appendix Table A1 studies site demographics.

[6]In total we have 150 officers assigned to wear BWCs and 300 control group officers. We stratified by precinct, officer activity, rank, previous internal investigations, and gender.

camera in any shift. Those two layers of randomization induce random allocation of cameras to dispatches, our primary unit of analysis. We consider a dispatch as treated if at least one officer tending to an event was wearing a camera. Since the vast majority of dispatches involve more than one officer, this sparser one-in-three officer-level randomization was chosen such that approximately half of the dispatches and events post-treatment would have a body-worn camera, maximizing power and in sharp contrast with the existing literature which typically assigns cameras to more than 50% of the officers participating – we will elaborate on this in our discussion.[7]

Panel C of Figure 1 displays the number of dispatch events we observe by day over the project period along with a moving average of the number of events and the number of dispatches that had at least one officer present with a camera present. Panel D provides the tabulation of the number of dispatches across the two layers of randomization. Out of the population of events that did not occur on blackout days – 13,264 dispatches – around 58% have had an officer present that was wearing a camera in the respective shift in line with our simulations.

The integrity of the research design was protected due to 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 weapon and equipment. Further, the blackout days were randomly selected at the start of the experiment but only communicated directly to the armory the evening before to avoid potential selection around the blackout days. Further, dispatch operators were blind to whether dispatch units are carrying a BWC. We find no evidence suggesting that there was 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. Prior to the experiment, a series of standard operating procedures needed to be created or updated – particularly concerning data privacy. The research team never had access to any recordings due to individuals' privacy

[7] Appendix Figure A2 shows the result of the simulation which suggests around 50% of the events would count as treated with the 1/3 to 2/3 allocation.

concerns. The police also issued an updated standard operating protocol (*SOP*) that mandated that every operation involving an interaction with a citizen should be recorded, with few exceptions such as sensitive or covert operations. Police officers were also required to inform citizens verbally that "the dispatch was being recorded, according to police protocol" whenever the situation allowed.

Data We primarily draw on three main data sources. The first is *dispatch-level data* which is facilitated by PMSC fully digital data backend called PMSC mobile. The data captures the universe of all dispatch events that were attended by any PMSC officers. These events typically would originate from 911 calls, from self-initiated calls due to routine operations (such as patrolling), or due to scheduled activities (e.g. the execution of court orders). Our main outcome dataset contains a total of 17,661 events that span over the experimental period ranging from September 3rd 2018 to December 10th 2018 (see Panel D in Figure 1). It includes information regarding (i) timing of the event (call, dispatch, arrival at the scene and end of the event); (ii) geographic information (the precise GPS location and full address); (iii) event classification and reporting (dispatch opening and closing classifications, internal prior risk assessment, the facts that were reported during the interaction, and an indicator if the event generated a formal police report); (iv) use-of-force (physical, non-lethal or lethal-force and number of victims, arrests and handcuffs deployed) and (v) the hashed identifiers of officers that attended the event. We can merge this data with the serial number of the camera that has been assigned to treatment officers at the start of each shift.

In order to capture whether events are recorded, we merge the event information along with the officer identifiers to the individual camera log files. These log files along with the recordings cannot be tampered with. The information is extracted from the device in the armory after cameras are handed back at the end of each shift for charging. The log files are subsequently transmitted to the research team. As the log files provide both the serial number of the device and all information on when and for how long the camera was activated, we can match this to dispatch event information to capture whether recordings actually take place. This level of data access and end-to-end measurement of compliance marks a significant improvement vis-a-vis the existing literature.

Further to the dispatch data, we observe a range of officer characteristics such as their job title, rank, gender, the date of admission to the force along with the number of internal investigations that have involved the specific officer.[8] These characteristics were also used to inform the stratification are subsequently used to explore pre-registered heterogeneous effects.

We study two broad sets of outcomes in addition to measuring compliance directly: reporting and police-citizen interactions around a dispatch event. Reporting is measured by (i) if the event generated generates a formal police report, usually forwarded to the Civil Police which is responsible for investigative work preparing the formal judiciary charges and by (ii) if there was any victim, which is a measure of diligence and discretion of police activity. For the interaction margins we focus on (i) a measure of citizen behavior (if there was any filing of contempt, disobedience or resistance charges towards police officers), (ii) if any use-of-force was deployed (either physical, non-lethal or lethal, but excluding handcuffs and arrests) and (iii) if there were any arrests or deployment of handcuffs. We further create an inverse covariance-weighted index combining these three outcomes following Anderson (2008), which we call Negative Interaction Index.

3 Empirical Analysis

3.1 Main Specification and Results

In what follows, we use the following empirical specification

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

where i indicates an event attended by a police dispatch, b is the police precinct, d is the day and w is the week of intervention. The number of officers on the event is n. For our main specifications, we consider that $\text{Treated}_i = 1$ if at least one officer attending a dispatch was assigned to wear a camera. In this sense, our

[8]Such internal investigations could be the triggered by a formal complaints either from within PMSC or from citizens.

main specifications reflect intent-to-treat estimates. We include police precinct by week fixed effects (η_{bw}) and number of officers fixed effects (n_{ibdw}). We also include fixed effects for each of the stratification bins ($\sum_{j=1}^{n} \phi_{o_j(i)}$) along with day-of-the-week fixed effects (τ_d). In our initial specifications, we exclude blackout days and focus exclusively on comparing treated with control events. The disturbance ϵ_{ibdw} is clustered at the police precinct-by-day level.

Results We first document that dispatches treated by cameras produce notably more camera recordings in column (1) of Panel A of Table 1. On average, 23.9% of the treated events were recorded, and zero in the control group. The operating procedures for BWC usage required that cameras should only be activated if there was any interaction with citizens, which does not happen in all dispatches. Therefore, as expected, not all treated dispatches have an associated recording. Importantly, we see no recordings in events attended exclusively by control group officers.

We next study the impact of BWC on police reporting behavior across columns (2) and (3). We find that, on average, dispatches treated with a BWC present are reported 3.0 percentage points more often to the Civil Police, capturing a 9.2% increase. We also find that share of events in which a victim is reported increases significantly in the treatment group by 2.7 percentage points – capturing a 19.6% relative increase. These results suggest that BWCs successfully affect officers' reporting behavior. As we show in Appendix Table A2, we find evidence that officers' description of the type of crime inflicted is also affected by camera usage. Most notably, domestic violence cases are reported 67.5% more often when cameras are present. We interpret these effects as ensuing from the accountability and diligence of the police actions promoted by the camera.

We next study the treatment effects on citizen-officer interactions. The results suggest that BWCs reduce the negative interaction index by 0.37 percentage points, representing a decrease of 44.2%. Further, we find that filing of charges against citizens, the use-of-force by the police, and the use of handcuffs or arrests decreased substantially – respectively, by 28.5%, 61.2%, and 6.2% –, although only the effect on use-of-force is strongly statistically significant at conventional levels. The substantive decline in use-of-force marks a notable contrast with the existing literature that has typically found muted or no effects. We revisit this in the discussion section.

Heterogeneity We explore heterogeneous effects in the subsequent panels of Table 1. In Panel B, we study whether effects are primarily concentrated in events that were classified *ex-ante* as low risk prior to dispatch. This assessment is done prior to every dispatch by considering whether: (i) Are there people with life-threatening injuries? (ii) Is the suspect still on site? (iii) Is the suspect carrying a weapon? and (iv) Is there a risk of turmoil? An event is considered low-risk if the response is negative to all these questions. The results suggest that the effects of BWCs improving citizens and police interactions are fully driven by events that are ex-ante classified as low risk. For those events, the negative interactions index is reduced by 48.0%. No BWC effects are detected among events that are judged to be high-risk ex-ante: the negative interaction index points to a much smaller impact of 9.5% which is not statistically significant. This suggests that BWCs may avoid escalation of situations. In high-risk events which have already escalated prior to dispatch, the presence of a camera itself may not affect the situational dynamic. Taken together, 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.

In Panel C we find evidence that the treatment effects are larger with more cameras on site. This suggests that the extensive and intensive margins of monitoring matter. We find that dispatches with two or more cameras were recorded 7.8 percentage points more often (or a 34.8% increase), the likelihood of a police report increases by 1.4 percentage points (51.6% increase), as well those reports tend to include victims 1.6 percentage points more often (or a 66.3% increase). Those reporting effects are accompanied by a further reduction in the negative interaction index – promoting a further drop of 25.4%. In particular, the use-of-force falls by 80.1%, which however is only significant at the 10% level.

Mechanisms We next show that the characteristics of the officer *who* is wearing the camera appear to matter. In Panel A of Table 2, we explore treatment effects based on the rank of the officer assigned to hold the camera. We classify police officers into a low-ranked "soldier" category and a higher-ranked corporal or above. We see that the BWC treatment effect is only present when an officer with a soldier rank is holding a camera in the dispatch unit. Importantly, compliance with the protocol appears to be notably lower when higher-ranking officers carry the camera: dispatches appear to be recorded 22.8% less often compared to dispatches in which junior officers are assigned to wear the camera.

Early career officers are more likely to show behavioral improvements and protocol compliance when in presence of a camera; this suggests that the reduction of negative encounters between police officers and citizens is led by mostly changes in police behavior rather than citizens changing their conduct when in presence of a camera. These effects are consistent with career concerns being effective mediators of compliance and treatment effects.[9]

Blackout specifications We next investigate whether BWC effects would still be present if officers in treated dispatches were not allowed to wear them. To do so, we leverage on randomized blackout shifts, mimicking a shift-level experimental design that is also commonly used in the literature. The added advantage is that our design gives us treatment variation across shifts *within officers*, as explained in Section 2. We estimate the following split-sample estimation equation:

$$\begin{aligned} y_{ibdw} &= \beta_1 \times \text{Treated}_i \times \text{Treated Shift}_d + \beta_2 \times \text{Treated}_i \times \text{Blackout Shift}_d + \\ &\quad + \eta_{bw} + \tau_d + n_{ibdw} + \sum_{j=1}^{n} \phi_{o_j(i)} + \epsilon_{ibdw} \end{aligned} \tag{2}$$

where $\text{Blackout}_d = 1$ if day d was randomly selected as a blackout day and interact all the fixed effects with the blackout day indicator. Thus β_1 captures the treatment effect within regular days (delivering the same point estimates as in Table 1 Panel A) and β_2 captures the effect of blackout days, which we would expect to be zero, on average, in absence of learning effects.

[9] In Appendix Table A3 we compare the effects when dispatch units were composed of *only* soldiers with at least one officer above the soldier level, irrespective of who, in the dispatch unit, was actually wearing the camera. We show that, except for dispatch recording, the effect is not driven by the dispatch unit rank.

The results are presented in Table 3. Column (1) highlights that, on blackout days, hardly any event gets recorded due to the experimentally induced absence of cameras. We note that all BWC treatment effects on the outcomes measuring citizen and police interactions disappear. Yet changes in the reporting behavior persist. This could indicate that officers that were in the past exposed to the use of cameras behave differently, even in the absence of the camera. This can be consequential for alternative research designs as we discuss below. The learning effects primarily affect the reporting margin but do not appear to have an effect on use-of-force or other citizen-police interaction margins.

Robustness We conduct a few robustness checks that we further discuss in Appendix B. Specifically, results are robust to changes or refinements in the estimation sample (see Appendix Table A4). Further, we do not see changes in the spatial distribution of treatment or control dispatches (see Figure A3). Appendix Table A5 further confirms that treatment status is not meaningfully correlated with location or police response times. Lastly, we rule out that the results on low-risk events mechanically mirror the results on early career officers. This could occur if low-risk events are assigned to junior officers. Appendix Table A6 shows this not to be the case.

3.2 Situating findings vis-a-vis the existing BWC literature

Our results stand in significant contrast with much of the existing literature which has often failed to detect effects of BWCs on use-of-force.[10] Different findings could naturally have arisen due to the different settings in which the experiments were conducted. For instance, our study is the first to evaluate the effects of body-worn cameras in a middle-income, high-crime setting (compared to existing studies which are mostly conducted in the UK or the US). While we cannot rule out that context-specific effects may have played a role in explaining the differences in the estimated BWC effect, we provide evidence consistent with the view that the previous studies were plagued by methodological issues that resulted in a muted evaluation of BWCs effects.

We can make progress in this direction by replicating in our data the evaluation

[10]See Appendix Table A7 for an overview of the literature.

designs used in most past studies. This is feasible as our study directly nests shift- and officer-centric research designs.[11] Doing so uncovers BWC effects that are similar to past studies; this naturally suggests that the BWC literature null can be due to the empirical design rather than the absence of true effects. We argue that such collection of results is consistent with contamination attenuating effects estimated at coarser levels of analysis.

We also investigate how the data aggregation might have affected the BWC effect estimates. Unsurprisingly, we find that more disaggregated data – as used in this study –, provide 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 that make use of aggregate data may suffer from power issues. We finally compare two leading estimation strategies – differences-in-means and differences-in-differences typically used in spatially explicit designs –, and suggest that both can recover relatively similar BWC effects.

We next describe these exercises, and Appendix C provides the econometric details that underpin those analyses.

Unit of randomization and analysis We first contrast the results from our design at the event-level with what we would obtain if we were to reanalyze the data at the officer or shift levels, reproducing the level of variation of several past studies.[12] This allows us to investigate the extent to which the BWC camera effect estimates are sensitive to the experimental design.

The estimate in Panel A of Figure 2, labeled as "event" and in red, shows the point estimate and the 90% confidence interval. This is our estimated event-level BWC effect which delivers a reduction in use-of-force of 61%. We next aggregate the outcome data at the officer-by-day level. The dependent variable is the share of events that had use of force during a given day in which cameras were used for a given officer.

[11] Appendix Table A7 overviews the main features of 31 papers that investigate the effects of body-worn cameras. We classify the studies as shift-centric (7 papers), officer-centric (13) or spatially-explicit designs (11).

[12] Panel A of Table A7 lists seven papers with shift-centric designs. Panel B lists 13 papers that make use of officer-level allocation.

This specification only explored variation between treated and control group officers. The share of events attended by officers assigned to wear a BWC see a decline in use-of-force by 33%. This effect is roughly half the effect that was estimated in the event-level data. The attenuation is not surprising: in our design 1/3 of officers were randomly selected to wear a camera, resulting in around 50% of the events being treated with at least one camera present due to the of dispatches. At the coarser officer-level, since a noticeable share of events attended by control group officers are indirectly treated, this leads us to underestimate the true treatment effect by around 50%. Such contamination-induced attenuation bias may affect existing studies conducted at the officer-level as many contexts involve routinely dispatching officers in teams. What is even more problematic is that almost all existing studies cannot directly test or measure contamination due to a lack of detailed event-level data.[13]

Further, the extent of contamination-induced attenuation bias is likely increasing in the share of officers that wear a camera. Virtually all officer-centric studies opted for a design with 50% of officers assigned to wear a BWC. Assuming a similar dispatch composition as in our context, this implies that 75% of all events are treated with at least one camera (see Appendix Figure A2), undermining power and downward-biasing the treatment effect estimate when studying the officer-level data. As a result, we would expect that to only estimate an effect-size that is 25% of the event-level design at the officer level – implying that we would estimate that use-of-force only declines by 15% (vis-a-vis the actual much larger decline of 60%).

Lastly, we can also speak to the shift-level design due to the blackout versus non-blackout days in our design. Seven studies opted for such a design; the figure indicates the point estimate from one such study (see Ariel et al., 2016b). Similar to that study, we find a statistically insignificant negative effect on use-of-force of about 9.5%. There are two likely explanations for this result. First, learning effects of camera use have the potential to alter officer behavior in control shifts. This is

[13]Of the 12 studies that opted for an officer-based design, we identified whether officers are dispatched in teams for only six studies – out of those, 50% report that officers are dispatched in pairs or more officers. These studies may thus be vulnerable to such attenuation bias.

indeed documented in Table 3 in our own setting. Second, the loss of precision can be associated with power issues from a reduced sample size: it appears that for shift-designs to be able to detect effects, samples need to be much larger and the treatment may need to be much higher powered.

Temporal resolution of outcome measurement We document that accounting for unobserved time-effects may be very important. Holding all other aspects constant, we focus on the officer-level variation studying outcomes either at officer-by-day, officer-by-month, or pooled levels. Specifications at coarser levels may introduce a broad range of biases as it implies that we can not control for the potential confounding effect of time fixed effects which are likely very relevant.

In Panel B of Figure 2 documents what happens to our point estimates with various data aggregations. We move to the officer-by-month data cuts and detect use-of-force estimate of around 30.8%, which is comparable to the estimated effect of 33% in the officer-day level. We only reject the null hypothesis in the officer-by-day data cut, highlighting that controlling for unobservable time factors do not immediately appear to affect the point estimates but may decrease power.

We lastly document our point estimates when conducting inference by computing differences-in-means between treated and control group officers with the data resolution at the pooled officer level. The outcome variable is the share of incidences in which force was used for each officer throughout the experimental period. We find a reduction in use-of-force by around 14.3% and not statistically significant, which is more than a four-fold decrease from the event-level estimates. We illustrate the null-effect point estimate of Yokum et al. (2019), which is an example of a pooled officer-level study, which lies well within the confidence interval of our estimate that is obtained when replicating such a pooled design. The failure to account for potentially confounding time-effects both appears to widen the confidence bands and implies a further attenuation of the estimated effect.

Inference method Lastly, our design also allows us to speak to different causal inference methods. The most common alternative design is difference-in-differences used across eight studies. At the officer-level, this may tackle the attenuation issue that gets introduced by the fact that many officers may attend events together resulting in control group officers being treated indirectly. In Panel C of Figure 2, we document the point estimates that emerge when we use between-officer difference in means estimate and a simple difference-in-difference design. The point estimate of the pooled difference-in-means at the officer level is the same as the third estimate in panel B; moving to the difference-in-differences, we again uncover a negative treatment effect estimate. However, this is still smaller compared to the point estimate at the event-level. Yet, it does suggest that difference-in-differences in general, while still underestimating the treatment effect and being rather imprecise, it comes closer to the event-level results. As indicated, this is not surprising as we implicitly control for some common time effects and also are controlling for some time-invariant confounders, helping precision. We illustrate the point estimate from Braga et al. (2018) as an existing study that opted for such a design.

Our combined conclusion from these exercises is that many existing studies may systematically underestimate the treatment effects. This may explain why the literature has, to date, not systematically shown that BWCs are effective devices in curbing police use-of-force. We now conclude the paper.

4 Conclusion

Police violence is a worldwide concern and there is an urgent need to find ways to increase accountability. In this paper, we investigate the effects of body-worn cameras on police officer reporting behavior, on citizen misbehavior, on use-of-force, and on use of handcuff and arrests. Through a large-scale experiment with an original design, we show evidence that body-worn cameras are effective to reduce use-of-force by police officers. The experiment took place in the state of Santa Catarina, Brazil, and five precincts were part of it with approximately 450 police officers taking part.

The results show that body-worn cameras are effective to improve the nature of police-citizen interaction – much to the contrast of the existing literature. Body worn cameras reduce use-of-force by the police by around 61.2% and improve the reporting accuracy of police officers. Moreover, we show that the decrease in use-of-force takes place in low seriousness events, as judged by a previous measure of risk assessment. The dispatch composition and the characteristics of the officer who is wearing the camera also matter for the treatment effect. Officers early in their careers, with a soldier rank, present higher reductions on the negative interaction index and on the filing of charges against citizens when carrying cameras.

The experiment has important policy implications. First, the results suggest that cameras are effective to curb police violence, which suggests that using them can increase the accountability of police officers. Moreover, the officer who wears the camera is also an important feature for compliance and for the results regarding use-of-force to exist. Officers early in their career and with no previous investigations are more likely to comply with the protocol and show improvements in reporting and interaction margins metrics. Implementing cameras can be an important step towards decreasing excessive use-of-force by the police, but to ensure that the cameras are efficient, it is important to consider the career incentives that exist for police officers that wear cameras. If officers are concerned about career progression, they are more likely to adjust their conduct to the protocol, fearing the possible repercussions. Put together with the results on blackout days, which show that treated officers in blackout days do not show a reduced use-of-force, the effects indicate that wearing a camera is important for inducing behavioral changes even for officers with career incentives, which are driving most of the treatment effect. Therefore, cameras are effective if police officers are concerned about the career implications of misbehaving.

References

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*(484), 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*(6), 744–755.

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*(2), 626–655.

Blair, G., J. Weinstein, F. Christia, E. Arias, E. Badran, R. A. Blair, A. Cheema, A. Farooqui, T. Fetzer, and G. Grossman (2020). Does Community Policing Build Trust in Police and Reduce Crime ? Evidence from Six Coordinated Field Experiments in the Global South. *Working Paper*.

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*(0), 1–30.

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*(3), 1–18.

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*(3).

Chassang, S. and G. Padró i Miquel (2019). Crime, Intimidation, and Whistleblowing: A Theory of Inference from Unverifiable Reports. *Review of Economic Studies 86*(6), 2530–2553.

Ferraz, C., J. Monteiro, and Bruno Ottoni (2016). Monopolizing Violence in Ungoverned Spaces : Evidence from the Pacification of Rio's Favelas. *Preliminary Draft*.

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*(3), 291–313.

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*(3), 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.

Magaloni, B., E. Franco, and Vanessa Melo (2015). Killing in the Slums: an Impact Evaluation of Police Reform in Rio De Janeiro. *Stanford Center for International Development* (556), 1–53.

Ornaghi, A. (2019). Civil Service Reforms : Evidence from U.S. Police Departments. *Working Paper July*, 1–55.

Owens, E., D. Weisburd, K. L. Amendola, and G. P. Alpert (2018). Can You Build a Better Cop? Experimental Evidence on Supervision, Training, and Policing in the Community. *Criminology & Public Policy 17*(1), 41–87.

Rozema, K. and M. Schanzenbach (2019). Good cop, bad cop: Using civilian allegations to predict police misconduct. *American Economic Journal: Microeconomics 11*(2), 225–268.

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

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

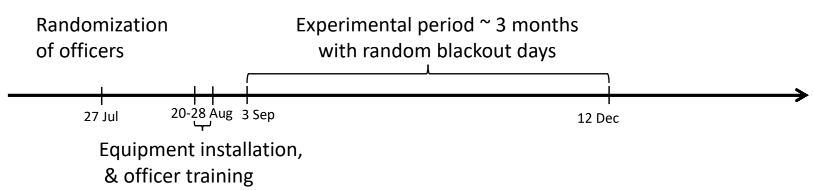
Xu, G. (2018). The costs of patronage: Evidence from the British Empire. *American Economic Review 108*(11), 3170–3198.

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*(21), 10329–10332.

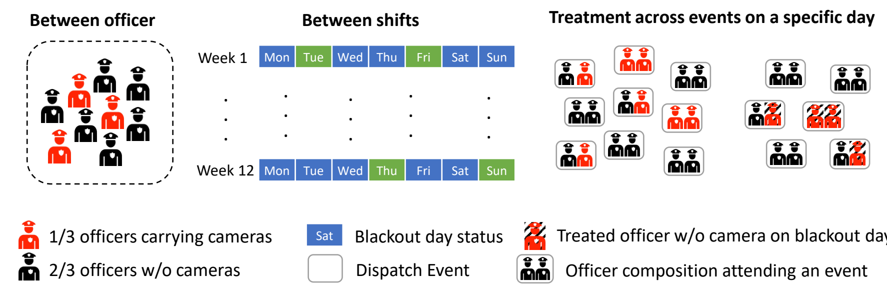
Tables and Figures

Figure 1: Timeline and experimental design

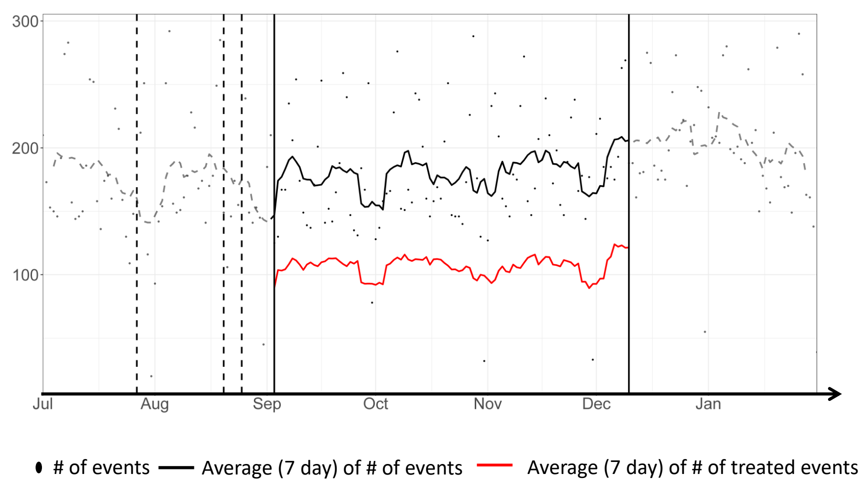
Panel A : Timeline of experiment



Panel B : Experimental design



Panel C : Number of events over time and share of treated events

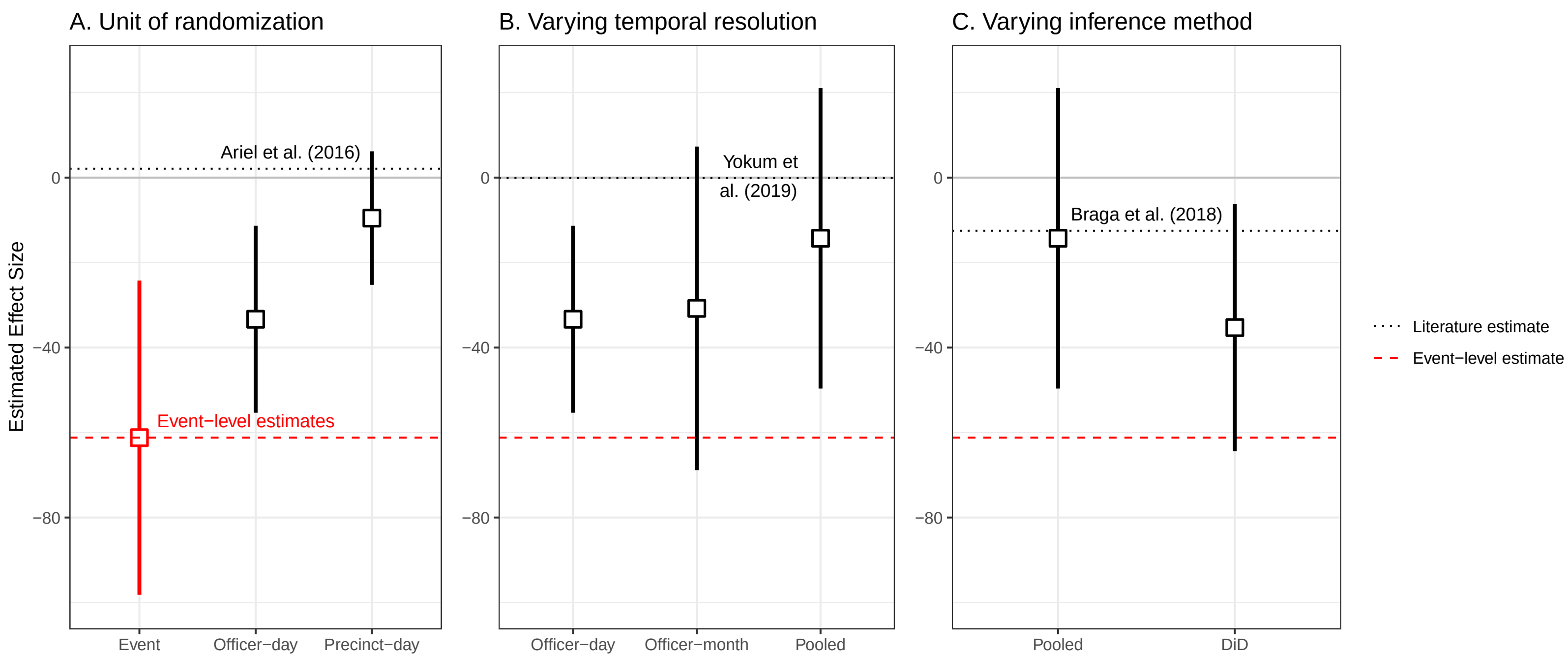


Panel D : Tabulation of # events by implied treatment status

		Days	
		Regular (5/7)	**Blackout (2/7)**
Officers	**T (1/3)**	7803	2646
	C (2/3)	5461	1751

Notes: The figure presents the experimental design. Panel A provides the timeline of the experiment that was conducted in 2018. Panel B illustrates the between- and within officer variation that is randomly induced and how this can map into different treatment status at the individual event level. Panel C plots the time series of the number of events with a police dispatch per day across the experiment along with the seven day moving average of the number of treated- and overall number of events illustrating that, on average, 50% of events haven an officer attending that is assigned to wear a camera. Panel D presents the tabulation of the overall number of experimental events by the treatment status.

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



Notes: The figure presents results on how the estimated treatment-effect sizes vary if we reanalyze the data using different commonly used evaluation strategies. Panel A explores how changing the unit of randomization affects the results. Benchmark results from this paper exploit event-level variation and are presented in red. More commonly used designs exploit only experimental variation between treated- and control group police officers or shifts. Estimates of effect sizes from reference studies in the literature using such designs are annotated as a horizontal dashed line. Panel B explores varying the temporal resolution of the outcome data focusing on the officer-level. Panel C explores different inference approaches commonly used.

Table 1: Effects of body worn cameras on accuracy of police reporting and citizen-police interactions

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Main Effect							
Treat	23.883***	3.018***	2.717***	-0.374***	-0.266	-0.426***	-0.338
	(1.110)	(1.014)	(0.661)	(0.142)	(0.187)	(0.156)	(0.464)
Panel B. Heterogeneity by ex-ante Event Risk Assessment							
T x Low Risk	23.868***	2.818***	2.288***	-0.406***	-0.383*	-0.415***	-0.456
	(1.119)	(1.061)	(0.683)	(0.143)	(0.196)	(0.150)	(0.487)
T x High Risk	24.011***	5.114*	6.903***	-0.080	0.767	-0.503	0.732
	(1.880)	(2.651)	(2.273)	(0.640)	(0.802)	(0.698)	(1.507)
Coef. Difference	-0.142	-2.295	-4.614**	-0.325	-1.150*	0.087	-1.188
	(1.659)	(2.789)	(2.386)	(0.661)	(0.843)	(0.706)	(1.580)
	[0.4659]	[0.2060]	[0.0276]	[0.3116]	[0.0873]	[0.4508]	[0.2268]
Panel C. Treatment Intensity							
T - 1 Camera	22.320***	2.734**	2.397***	-0.331**	-0.210	-0.393***	-0.122
	(1.131)	(1.080)	(0.687)	(0.135)	(0.192)	(0.145)	(0.495)
T - 2 or More Cameras	30.082***	4.145***	3.986***	-0.546**	-0.489	-0.558*	-1.194
	(1.864)	(1.527)	(1.109)	(0.276)	(0.352)	(0.295)	(0.768)
Coef. Difference	-7.762***	-1.411	-1.589*	0.215	0.278	0.165	1.072*
	(1.674)	(1.530)	(1.088)	(0.246)	(0.349)	(0.252)	(0.791)
	[0.0000]	[0.1790]	[0.0733]	[0.1922]	[0.2135]	[0.2571]	[0.0888]
Mean DV Control	0.000	32.815	13.844	0.845	0.934	0.696	5.420
N	13264	13264	13264	13264	13264	13264	13264

Notes: Table presents results on the impact of a body worn camera being present at a police event. Panel A presents the main results capturing the average intent-to-treat effect. Panel B explores heterogeneity by the ex-ante risk level of the events, which characterizes an event as low risk if it has no weapons on the scene, if there are no injuries, if the suspect is not on site and if there is no material risk of general unrest.Panel C investigates treatment intensity heterogeneity, given by the number of officers wearing a camera in events. The dependent variables are "Dispatch recorded" indicating that the dispatch was partially or fully recorded using the body worn camera and hence represents the treatment being delivered. "Police Report" and "Victims in report" capture the extent to which officers formally report events, on which basis the Civil Police would proceed investigations. Interaction Margins comprises: (i) "Negative Interaction Index" is the standardized inverse-covariance weighted average of the three indicators in the group; (ii) "Contempt, Resist and/or Disobey " is an indicator if charges of contempt, disobedience or resistance towards the police were registered; (iii) "use-of-force " is an indicator if there was any deployment of physical, non-lethal (mechanical) or lethal force by the police, not considering use of handcuff or arrest; (iv) "Handcuff and/or Arrest"is an indicator if handcuffs were used or if any arrests made. 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 are excluded from the regression. Standard errors are clustered at the precinct-by-day level. *** p<0.01; ** p<0.05; * p<0.1.

Table 2: Who holds the camera matters: Effects of body worn cameras on accuracy of police reporting and citizen-police interactions

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Rank							
Treated by Officer(s) with Soldier rank	24.743***	2.645**	2.835***	-0.444***	-0.377*	-0.476***	-0.463
	(1.187)	(1.085)	(0.721)	(0.152)	(0.207)	(0.162)	(0.468)
Treated by Officer(s) with higher than Soldier rank	19.112***	3.797*	2.492*	-0.187	0.033	-0.295	-0.036
	(1.987)	(1.953)	(1.377)	(0.235)	(0.337)	(0.254)	(1.002)
Treated by Officers of both types	24.593***	6.392**	1.597	0.144	0.536	-0.055	0.675
	(2.268)	(2.615)	(1.670)	(0.500)	(0.700)	(0.515)	(1.549)
Panel B. Any Previous Disciplinary Procedures							
Treated by Officer(s) with 0 Disciplinary Procedures	25.336***	1.809	1.903**	-0.414**	-0.306	-0.457**	-0.812
	(1.314)	(1.235)	(0.862)	(0.176)	(0.260)	(0.178)	(0.585)
Treated by Officer(s) with at least 1 Disciplinary Procedure	19.842***	4.658***	3.597***	-0.328*	-0.167	-0.421**	0.386
	(1.437)	(1.246)	(0.905)	(0.170)	(0.246)	(0.186)	(0.684)
Treated by Officers of both types	35.007***	2.681	3.649**	-0.341	-0.513	-0.246	-0.895
	(2.554)	(2.091)	(1.540)	(0.402)	(0.511)	(0.428)	(1.103)
Mean DV Control	0.000	32.815	13.844	0.845	0.934	0.696	5.420
N	13264	13264	13264	13264	13264	13264	13264

Notes: Panel A explores rank heterogeneity of who is wearing the camera. Soldier is the lowest rank in the Military Police. Disciplinary procedures are measured as if the officer has any *Inquérito Policial Militar*, *Procedimento Administrativo Disciplinar* and *Sindicância* Administrative investigations results from formal complaints over officer behavior, generated either internally or externally. Dependent variables defined as in Table 1. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Shifts without camera are excluded from the regression. Standard errors are clustered at the precinct-by-day level. *** p<0.01; ** p<0.05; * p<0.1.

Table 3: Exploiting within-shift variation: Effects of body worn cameras on accuracy of police reporting and citizen-police interactions

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated in Treated Shifts	23.883***	3.018***	2.717***	-0.374***	-0.266	-0.426***	-0.338
	(1.112)	(1.015)	(0.662)	(0.142)	(0.187)	(0.156)	(0.465)
Treated in Control Shifts	3.780***	3.541**	0.717	0.041	0.001	0.058	0.146
	(0.886)	(1.684)	(1.252)	(0.185)	(0.304)	(0.209)	(0.695)
Mean Dep. Var Control in Treated Shifts	0.000	32.815	13.844	0.845	0.934	0.696	5.420
Mean Dep. Var Control in Control Shifts	0.000	34.266	17.190	0.681	0.800	0.514	5.425
N	17661	17661	17661	17661	17661	17661	17661

Notes: Table documents within-shift comparison between events that had a police officer assigned to wear cameras and control ones across randomly assigned shifts in which treatment officers were handled cameras (Treated Shifts) and the ones in which they were not (Blackout or Control Shifts). Dependent variables defined as in Table 1. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. All fixed effects are interacted with a Treatment Shift indicator for within-shift estimation. Standard errors are clustered at the precinct-by-day level. *** p<0.01; ** p<0.05; * p<0.1.

Online Appendix

“De-escalation technology: the impact of body-worn cameras on citizen-police interactions”

For Online Publication

Barbosa, Fetzer, Soto and Souza

September 23, 2021

A Implementation Details

The project timeline considered that officers selected to wear the camera would have to go through special training and have their uniforms adapted to hold the camera com fortably and effectively. For these reasons, the randomization allocation was assigned in 27/07/2018, about a month before the experiment started, but was not communicated to the officers until closer to the experiment date. All tests with the cameras and docking stations were conducted during this period to ensure that the information necessary for the experiment was correct and to minimize technical issues during the experiment period.

Shift-level treatment allocation was randomized before the start of the experiment, but we only communicated to the police precincts in 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 to begin with. Importantly, the blackout applies to officers starting their shifts. That is, officers already out for their regular patrol activities at midnight of the start of the blackout would continue to use their cameras until the end of their shift; conversely, any shift that starts during blackout that spans after its end 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 would happen as follows. At the start of their shifts, treated officers would obtain their camera, along with other equipment, from the armory section of the police precincts – from where they obtained their gun, radio, and other equipment of regular and irregular use. The armory sections are usually very secluded and considered to be of high-security environment – due to the nature of the material that is stored therein – and only a few high-ranked 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 in the armory rooms. This ensured that not only the equipment was maintained, regularly inspected, and kept to a good working order throughout the experiment and 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 pulled to the central HQ on demand due to bandwidth issues. The research team established routines to consolidate the camera automatic logs in a central database. In this way, it is 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 cameras to the armory section, docked in the station, and readied for the next use. This recycling process usually took between 4 and 6 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 all the officers that would start their shifts in the blackout day would receive from the armors all the equipment but the cameras. Figure 1 Panel B provides a clear visual exposition of our experimental design: 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 (red-colored officers), this dispatch (as well as the event they tended to) is classified as a treatment one. Thus, the average treatment effect of BWC implementation over police activity and police-citizen interactions is identified by comparing events attended by dispatches with at least one officer assigned to wear a camera with events with none.

As for blackout shifts, Figure 1 Panel B shows that in the green days (blackout), all treated (red) officers would not be 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 is allowed to wear cameras, allowing us to identify if the effects would persist were the treatment technology not present. Importantly, the dispatch operators were blind to whether dispatch units were manned with officers wearing a body-worn camera. This prevented the endogenous allocation of dispatch calls to be recorded (or, conversely, to avoid recoding). Panel C of Figure 1 presents the overall time series of events that occur in the study period and around it. The solid lines refer to the seven day moving average of the number of events (black) and the number of events attended by at least one officer wearing a camera (red), highlighting that as per our initial design this leaves around 57% of the events attended by some officers that are assigned to wear a camera. Panel D of Figure 1 provides the exact tabulation of the numbers of events that occur by their respective treatment status.

B Robustness Checks

To address potential concerns that would threaten the validity of our findings, we add some robustness checks. In Appendix Figure A3, we show a heat map with the geographic location of dispatches by treatment status. Results show that the spatial distribution of treatment and control dispatches is virtually the same, which suggests that there is no selection in space, reinforcing that treated and control dispatches attend events in similar areas of the city. One could argue that police officers with cameras would change their patrolling behavior and law-enforcement activity in order to avoid interacting with citizens. Therefore the observed decrease in the interaction margins would be attributable to a change in the type of events they attend to rather than cameras improving dispatch officers' behavior when present. Appendix Figure A3 shows we have no reason to believe that this is driving the estimated treatment effects. We further confirm this in Appendix Table A5 where we run the time-to-response, latitude, and longitude on the treatment status. We find that the treatment and control respond in similar times and are, on average, in the same geographical coordinates.[1]

Appendix Table A4 shows that the main results are robust to changes in the estimation sample. Panel A reproduces the main effects. In Panel B, we exclude dispatches that involved any officer that did not participate in the experiment randomization.[2] The results are mostly unchanged although we see a stronger

[1]The result is however significant for the longitude outcome. The point coefficient is .002 of a degree, which is approximate to 222 meters distance. Although statistically significant, in practice the average difference in longitude is negligible.

[2]Some officers didn't participated in the experiment because they were transferred to the precinct after the randomization, or because they are originally from other precincts, or they are

reduction in the filing of charges, and a smaller effect on reduction of use-of-force. Panel C looks at dispatches with two officers, which is the modal dispatch size. The results show that when we restrict the sample to these events, the effect on use-of-force becomes statistically insignificant, even though it remains negative and sizable in magnitude. The effects on the negative interaction index and on citizen bad behavior remain strong. At last, Panel D excludes dispatches with more than 4 police officers and again the results remain virtually the same. Overall, our results remain qualitatively unchanged in this exercise.

Finally, to support the career concerns hypothesis, we run a sequence of tests. Our results from Table 2 Panel A show that the effects are only present when an officer with soldier rank is holding the camera. Also, Table 1 Panel B shows the effects are primarily concentrated in ex-ante lower risk events. One could argue that high-rank officers are called more to higher-risk events with a lower potential of de-escalation of conflicts. In that case, our results wouldn't be driven by career concerns of younger officers, but rather the allocation of events to officers according to their rank. In Appendix Table A6, we show that the presence of higher-ranked officers in an event is not correlated with many of its characteristics. In particular, high-risk events are not correlated with the presence of a high-rank officer. Neither is time-to-dispatch, the negative interaction index, and if the dispatch was recorded or not. These suggest that the rank of the officers who attend to a given event is uncorrelated with the event characteristics. Appendix Table A3 also supports the hypothesis that events with only soldiers and events with high-ranked officers are similar in most margins, with the exception of recording the dispatch and citizen bad behavior.

C BWC effects and the literature null

Experiments on the effects of BWC on use-of-force do not consistently show that the cameras effectively work to decrease excessive use-of-force, and mixed evidence across studies as shown by Lum et al. (2019). Appendix Table A7 lists

administrative or IT officers which are only allocated to dispatches occasionally. In fact, 84% of events only have the involvement of officers that were part of the randomization roster.

the main BWC papers in the literature.[3] 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 is detected. As it can be seen, the literature is not conclusive on the BWC camera effects on use-of-force. We argue below that most papers were plagued with methodological issues that attenuated the camera effects.

We start in Panel A with the studies that allocated cameras on the basis of shifts. Those papers, whether providing experimental evidence or not, allocate cameras to treatment and control shifts. We argue that this design is potentially problematic as a single given officer may be allocated to both a treatment and a control shift. 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 use shift analysis, five have use of force as an outcome and only one finds statistically significant results (at the 5% level) that suggest that BWC affect use-of-force. Ariel et al. (2015) conducted the first experiment on BWC and it is by far the most cited paper in the literature. The shifts were randomized to be conducted with and without cameras, and the results suggest that BWC reduce use-of-force by the police. However, these effects are 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 BWC on use-of-force. Ariel et al. (2016a) suggest that one potential explanation for muted results comes from compliance with the protocol. They show that use-of-force rates were higher in sites where the compliance with the protocol was lower, and vice-versa. Magaloni (2019) does not find any effects of BWC in 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 find that BWC were effective to reduce 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 are always in the same assigned group throughout the duration of the experiment. This design also presents its challenges. First,

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

contamination is a substantial concern: among the officer-centric papers we could identify, half had routinely more than one 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 treat half of the police officers, which results in a much higher share of treated events – if an event is 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 imply on around 75% of treated events (see Figure A2), implying in a considerably smaller control group and potentially undermining power. A corroborating evidence from Braga et al. (2020), who use officer-level randomization combined with spatial selection of districts, indeed shows evidence of large contamination from treated officers to control officers. Finally, some papers included in the experiment officers that volunteered to wear a BWC (Jennings et al., 2015; Ready and Young, 2015; White et al., 2017; Headley et al., 2017; Braga et al., 2017, 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 BWC on police operations.

Finally, in Panel C we list papers that make use of spatially explicit empirical designs. Out of 11 studies, only three look at use-of-force as an outcome and only Kim (2021) find evidence of the impact of BWC. They use a different empirical strategy and take advantage of the variation in the timing of the adoption across US agencies to assess the effects of BWC on a national level. 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.

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

This experiment and the existing literature The level of detail of our dataset and the two-layer design of our experiment allow us to replicate shift and officer-centric designs in our data. In fact, when we mimic the analysis of other papers our results become closer to theirs as we show in Figure 2. This is suggestive evidence that the muted effects of BWC on the use-of-force in the literature are likely due to issues in the experimental setup, e.g. not accounting for contamination or undermined power.

In Panel A of Figure 2, the point estimate displayed in red shows the 61.2% percent reduction in use-of-force, which is computed from the nominal effect size of Table 1, along with the 95% confidence interval.[4] The event-level estimate is also signaled with the red dashed horizontal line across all panels, for ease of comparison with other designs.

We explore in Panel A how changing the randomization unit can affect the BWC effect estimates. We start with replicating studies that randomize officers into treatment and control. The outcome variable y_{id} is the share of incidents that the officer i used force during day d. We then explore the experimental variation in officer allocation to the treatment and control groups in the following specification:

$$y_{id} \quad = \quad \beta_{\text{officer-day}} \times \text{Treated}_i + \eta_{bw} + \tau_d + \sum_{j=1}^{n} \phi_{o_j(i)} + \epsilon_{id}. \tag{3}$$

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 fixed effects $\phi_{o_j(i)}$ for each of stratification bins. The disturbance ϵ_{id} is clustered at the police precinct-by-day level. $\text{Treated}_i = 1$ is if the officer was assigned to wear a camera and we are interested in the estimated β_{officer}. Even though this is not at the event level, this specification aggregates the data at a somewhat granular level. Even so, we see that the effect sizes are reduced from 61.2% to 33.3%. The attenuation of the results is consistent with spillovers effects, since the analysis at the officer level does not account for the fact that control officer will mechanically tend to dispatches with

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

treated officers.

The final estimate in Panel A explores the shift randomization. In this case, we collapse the data at the day-by-precinct level, and we explore the randomization between blackout and non-blackout days. This is close to the experimental design of shift-centric papers because a day is approximately composed of two 12-hour consecutive police shifts. In the following specification, the outcome variable y_{db} is the share of events in which force was used at police precinct b during day d.

$$y_{db} \quad = \quad \beta_{\text{precinct-day}} \times \text{Treated}_d + \eta_{bw} + \tau_d + \epsilon_{db}. \tag{4}$$

The fixed effects we control for are police precinct by week and day-of-the-week. The error term ϵ_{db} is clustered at the police precinct-by-week level. The effect sizes calculated with this model are around 9.5%, a substantial attenuation from the 61.2% reduction in use-of-force that was originally estimated from the event-level specification, and not significant statistically. This effect size is in fact comparable with studies that originally make use of variation at the officer level. For example, Ariel et al. (2016b) also uses data at the precinct-shift level and explores shift randomization, and is within our confidence interval. Appendix Table A8 shows the results estimated with our data present in Panel A – not only for use-of-force, but also for all the main outcomes considered throughout this paper.

Panel B of Figure 2 explores how the results change when we vary the temporal resolution of the data. Aggregation implies that you can control less for omitted effects, which could be controlled for using more granular fixed effects (such as day of the week for example). The first effect size is replicated from Equation (3), the most granular result we consider with this unit of randomization. The second model aggregates the data to officer-month level. In this case, the outcome variable is the count of use-of-force incidents by police officer i during month m. The estimates are obtained from the specification below:

$$y_{im} \quad = \quad \beta_{\text{officer-month}} \times \text{Treated}_i + \eta_{bm} + \sum_{j=1}^{n} \phi_{o_j(i)} + \epsilon_{im} \tag{5}$$

where officer is i and month m. We include police precinct-by-month and stratification bins fixed effects. ϵ_{im} is clustered at the police precinct-by-month level. Although the effect size does not change considerably, the precision decreases and the result becomes statistically insignificant, potentially because the coarser data reduced the sample size, and do not allow for the inclusion of granular fixed-effects as in Equation (3).

We can further aggregate the data for each officer and consider all the experiment period. In this case, the outcome variable is the number of instances of use-of-force by officer i during all the experimental period. This essentially only explores the cross-sectional variation, and we refer to this estimate as the "pooled" specification. The estimating equation is:

$$y_i = \beta_{\text{pooled}} \times \text{Treated}_i + \eta_b + \sum_{j=1}^{n} \phi_{o_j(i)} + \epsilon_i \tag{6}$$

where we include precinct and stratification bin fixed effects. The error term is clustered at the officer level. 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 uses data at the officer level pooled during the experiment period. The effect size they find for use-of-force is virtually zero, in both magnitude and statistical significance. Our pooled result is comparable to theirs and, again, their point estimates fall within our confidence intervals. In Appendix Table A9 shows the regression results for all the outcomes, and confirm that the attenuation bias occurs across all outcomes.

Panel C explores how the methods most commonly used by the literature – difference-in-means and differences-in-differences (DiD) –, would affect the estimates. The first follows the model just described in equation 6 and the DiD follows a model such as:

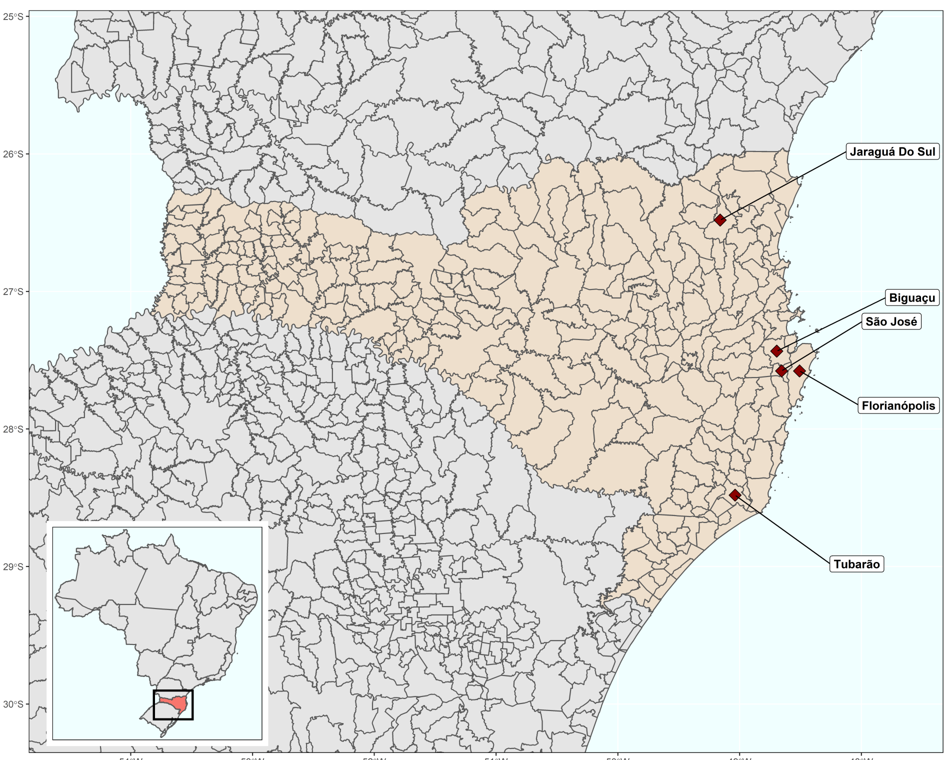
$$y_{it} = \beta_{\text{did}} \times \text{Treated}_i \times \text{Post}_t + \gamma_i + \epsilon_{it} \tag{7}$$

where t indicates a three-month time period and Post_t is equal to one in the period post-implementation. We include police officer fixed effects and cluster the standard errors at this same level. Unlike the difference-in-means, our effect sizes using the difference-in-differences are statistically significant and expressive in magnitude, even though the data is aggregated at such a coarse level. This can indicate that the within-officer analysis, which considers baseline information for each officer, might be effective to control for unobserved effects that can lower precision and introduce noise to the analysis. Again, we single out Braga et al. (2018) as a paper with similar inference method. Their results are within our confidence interval. In Appendix Table A10 we can see how changing the inference method matters and the pattern is similar across most outcomes.

Overall, our results suggest that the mixed evidence existing in the previous literature on BWC and use-of-force is suggestive that the experimental design (and the ensuing methodological challenges) can be a major reason for why muted effects of BWC were detected in previous work. This does not rule out that contextual factors played a role in explaining the difference in the results; in fact, it was expected that camera effects could substantially differ across study sites. Through this analysis, we however do not observe any evidence to this effect. We instead observe that mimicking in our data the research design in other papers estimates similar effect sizes. We conclude that research designs are more likely to be the main reason for differences in the policy evaluations across studies.

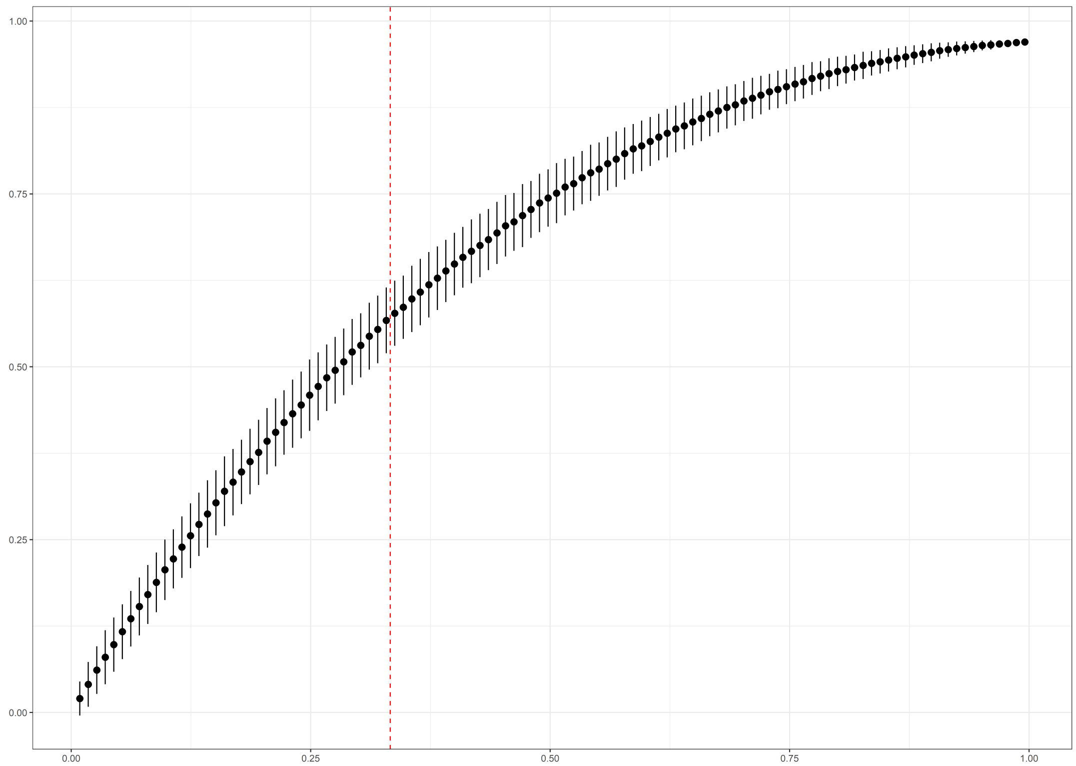
Appendix Tables and Figures

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



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

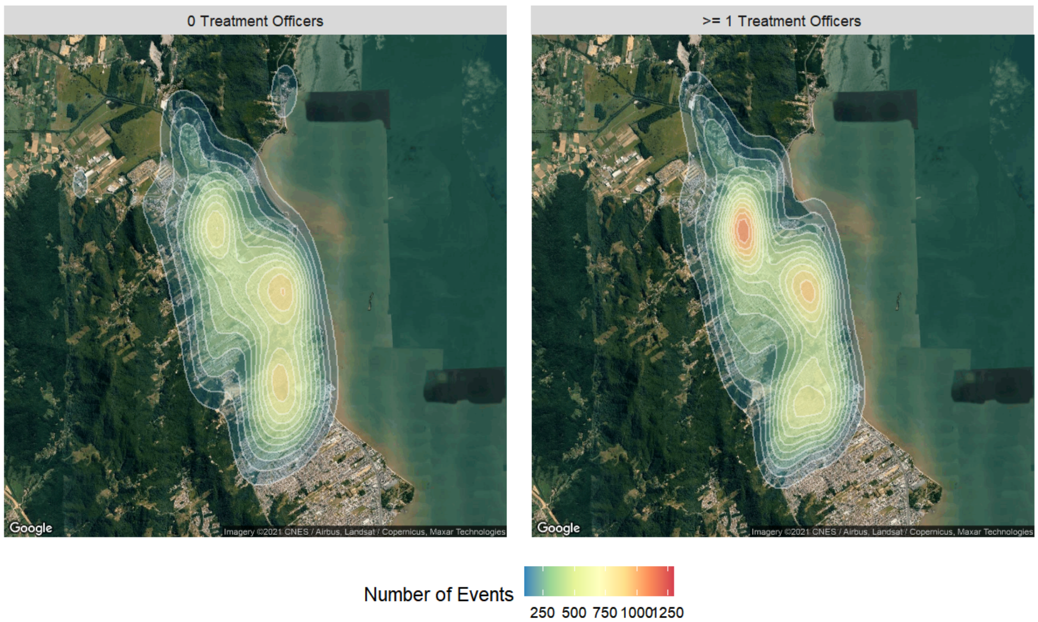
Figure A2: Induced treatment allocation at the event-level from officer-level camera randomization



Notes: Figure displays simulation results that were used to identify the optimal share of treated officers to ensure 50% of the dispatch events would be treated by virtue of having at least one camera attending the dispatch. The horizontal axis captures the share of officers assigned to receive treatment while the vertical axis plots the share of events that are treated by at least one camera at the dispatch. The vertical dashed line indicates the experimental design chosen whereby 1/3 of the officers are assigned to wear a camera while 2/3 are control group officers that never wear a camera themselves.

Figure A3: Spatial Distribution of Treatment and Control Dispatches

Biguaçu



Florianópolis

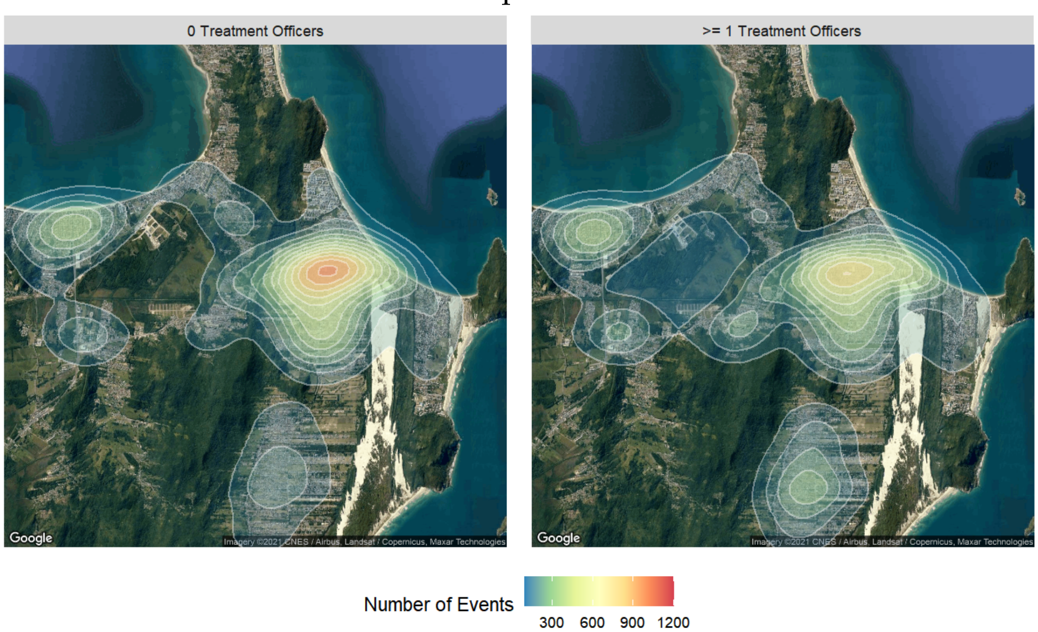
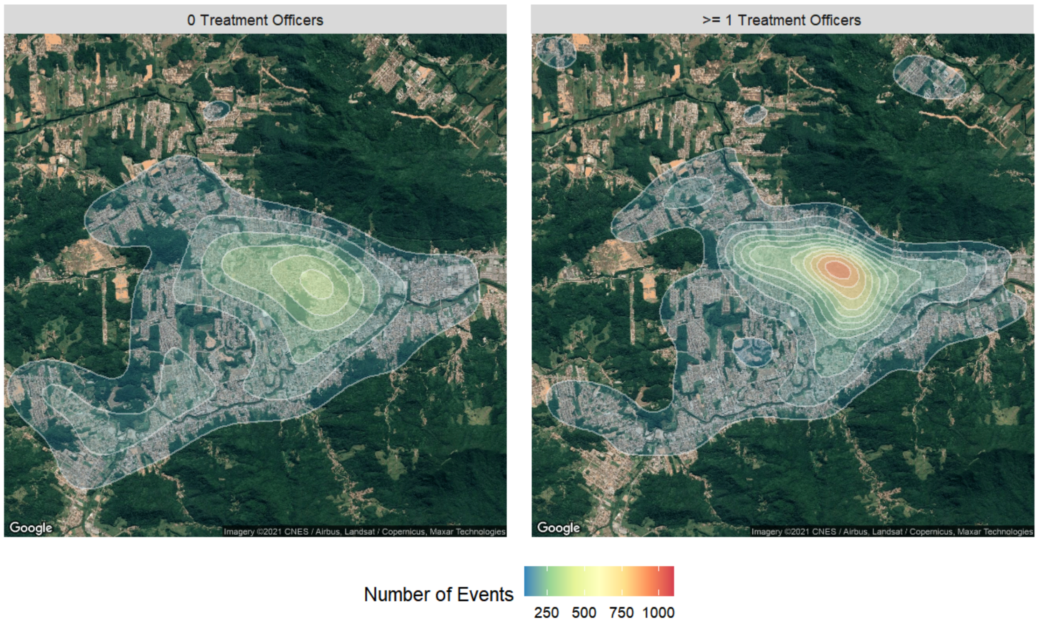


Figure A3 (cont): Spatial Distribution of Treatment and Control Dispatches

Jaraguá do Sul



São José

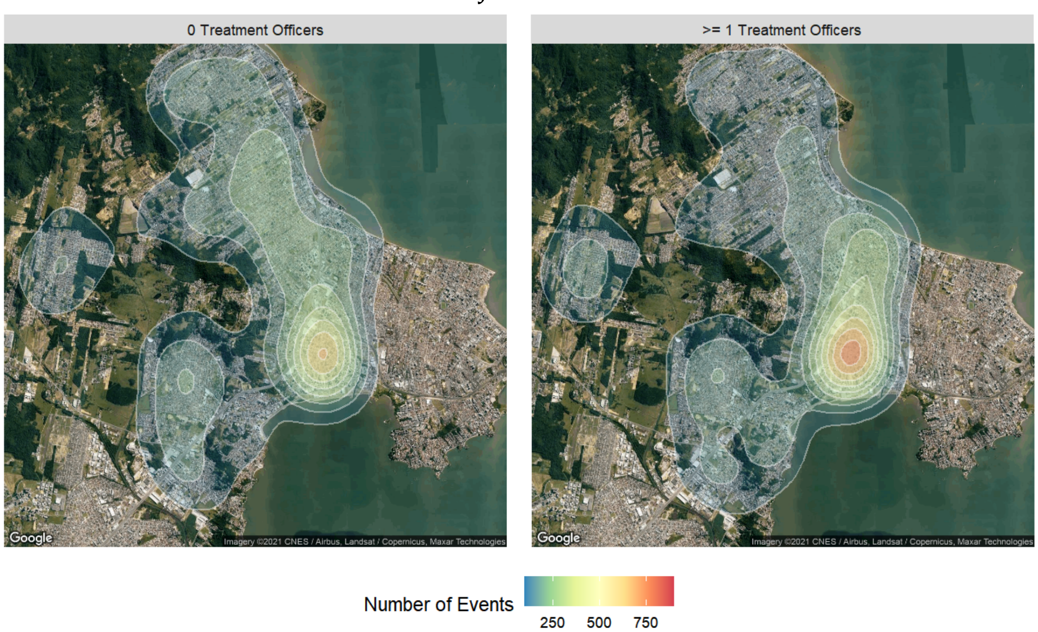
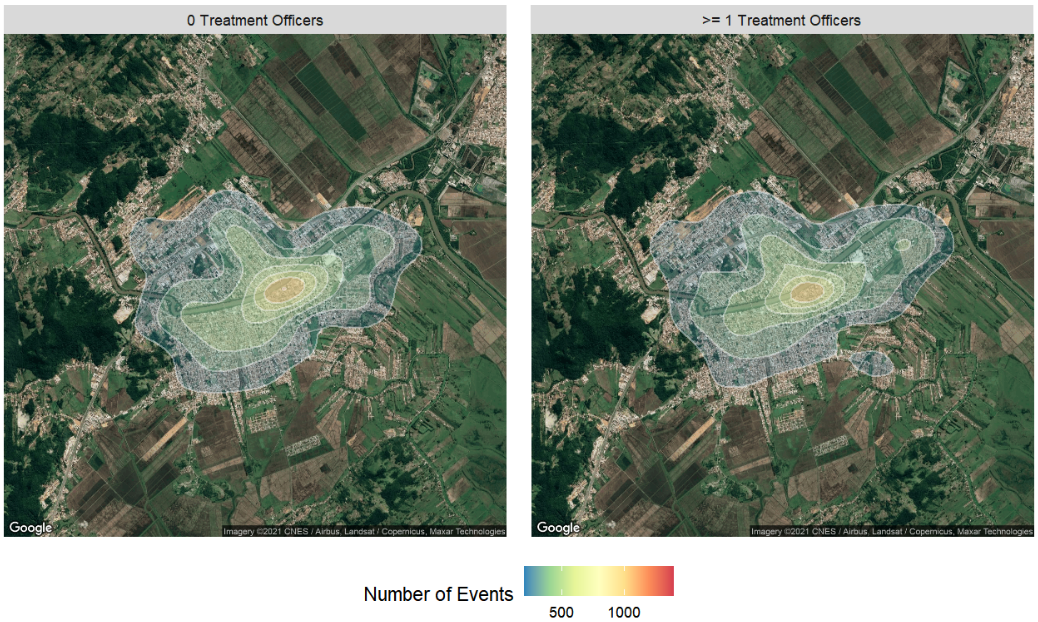


Figure A3 (cont): Spatial Distribution of Treatment and Control Dispatches

Tubarão



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

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: Socio-demographic characteristics and baseline violence across the five study sites and the average in Santa Catarina state. Sociodemographic data from 2010 IBGE Census, Homicide Rate from the 2016 IPEA Atlas da Violência and use-of-force and Crime Events Incidence from author's calculations using PMSC data from March to July, 14th 2018. Income in Brazilian Reais per month. Standard errors in parenthesis.

Table A2: Endogenous Fact Reporting

	Noise	Verbal	Robbery	Assault	Threat	Domestic Violence
	(1)	(2)	(3)	(4)	(5)	(6)
Treat	0.286	0.051	0.862*	0.695*	0.130	1.112***
	(0.581)	(0.496)	(0.454)	(0.389)	(0.328)	(0.284)
Mean DV Control	8.570	6.611	4.761	3.626	3.333	1.648
N	13,264	13,264	13,264	13,264	13,264	13,264

Notes: Intention-to-treat specifications. Unit of observation is a police event. All dependent variables are indicators if a given criminal typology was reported at the end of the event and 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 are excluded from the regression. Standard errors are clustered at the precinct-by-day level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table A3: Dispatch Composition

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Rank							
T x All Soldiers	26.193***	2.834**	2.703***	-0.378**	-0.500**	-0.323*	-0.272
	(1.275)	(1.178)	(0.803)	(0.157)	(0.222)	(0.166)	(0.515)
T x At least 1 Above Soldier Rank	18.409***	3.541**	2.799**	-0.363	0.287	-0.669**	-0.473
	(1.466)	(1.730)	(1.193)	(0.293)	(0.347)	(0.336)	(0.936)
Coef. Difference	7.784***	-0.707	-0.097	-0.016	-0.786**	0.346	0.201
	(1.623)	(2.019)	(1.451)	(0.329)	(0.415)	(0.372)	(1.058)
	[0.0000]	[0.3633]	[0.4735]	[0.4812]	[0.0300]	[0.1764]	[0.4248]
Panel B. Any Previous Disciplinary Procedures							
T x None	24.423***	0.818	3.227***	-0.403**	-0.465	-0.390**	0.339
	(1.591)	(1.592)	(1.178)	(0.181)	(0.297)	(0.176)	(0.765)
T x At least 1	23.786***	4.293***	2.499***	-0.358*	-0.185	-0.436**	-0.566
	(1.260)	(1.221)	(0.858)	(0.196)	(0.243)	(0.216)	(0.562)
Coef. Difference	0.637	-3.475**	0.728	-0.044	-0.280	0.046	0.905
	(1.747)	(1.920)	(1.537)	(0.270)	(0.386)	(0.279)	(0.925)
	[0.3579]	[0.0363]	[0.3184]	[0.4348]	[0.2344]	[0.4352]	[0.1647]
Mean DV Control	0.000	32.815	13.844	0.845	0.934	0.696	5.420
N	13,264	13,264	13,264	13,264	13,264	13,264	13,264

Notes: Intention-to-treat specifications. Unit of observation is a police event. Dependent variables defined as in Table 1 and heterogeneities are defined as in Table 2. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Shifts without camera are excluded from the regression. Standard errors are clustered at the precinct-by-day level. *** p<0.01; ** p<0.05; * p<0.1.

Table A4: Sample Robustness

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Main Specification							
Treat	23.883***	3.018***	2.717***	-0.374***	-0.266	-0.426***	-0.338
	(1.110)	(1.014)	(0.661)	(0.142)	(0.187)	(0.156)	(0.464)
Mean DV Control	0.000	32.815	13.844	0.845	0.934	0.696	5.420
N	13264	13264	13264	13264	13264	13264	13264
Panel B. Only Experimental Officers							
Treat	24.522***	3.583***	2.489***	-0.318**	-0.364*	-0.303*	-0.012
	(1.148)	(1.086)	(0.734)	(0.158)	(0.204)	(0.168)	(0.489)
Mean DV Control	0.000	31.417	13.149	0.815	0.946	0.655	4.973
N	11,200	11,200	11,200	11,200	11,200	11,200	11,200
Panel C. Two Officers - Modal Dispatch Size							
Treat	23.927***	3.959***	3.278***	-0.270**	-0.374**	-0.224	-0.149
	(1.186)	(1.118)	(0.741)	(0.126)	(0.175)	(0.140)	(0.438)
Mean DV Control	0.000	30.643	13.205	0.595	0.812	0.417	3.795
N	9,922	9,922	9,922	9,922	9,922	9,922	9,922
Panel D. At Most Four Officers							
Treat	23.894***	3.107***	2.880***	-0.328**	-0.288*	-0.344**	-0.451
	(1.109)	(1.025)	(0.669)	(0.129)	(0.170)	(0.141)	(0.460)
Mean DV Control	0.000	32.389	13.612	0.750	0.858	0.597	5.146
N	12,541	12,541	12,541	12,541	12,541	12,541	12,541

Notes: Intention-to-treat specifications. Unit of observation is a police event. Dependent variables defined as in Table 1. Specifications include police precinct-by-week, day of the week, number of officers and stratification bins fixed effects. Shifts without camera are excluded from the regression. Standard errors are clustered at the precinct-by-day level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table A5: Testing for endogenous allocation of events

	Time to Dispatch	Latitude	Longitude
	(1)	(2)	(3)
Treat	-1.543	0.001	0.002***
	(1.287)	(0.000)	(0.001)
Mean DV Control	10.579	-27.468	-48.787
N	13,264	13,264	13,264

Notes: Intention-to-treat specifications. Unit of observation is a police event. Sample includes all events in the experimental period and excluded shifts without cameras. Dependent variables are: (i) Time to Dispatch, which measures the length of the interval between communication and dispatch arrival in minutes and (ii) Latitude and (iii) Longitude, both measured in degrees. 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-day level. *** $p<0.01$; ** $p<0.05$; * $p<0.1$.

Table A6: Correlation between event characteristics and officer rank

	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Time to Dispatch	High-risk
	(1)	(2)	(3)	(4)	(5)	(6)
Any High Rank Officer	-0.264	5.995***	3.402***	0.137	-1.583	0.010
	(1.425)	(1.543)	(1.250)	(0.263)	(1.364)	(0.010)
Mean DV Control	0.000	32.815	13.844	0.845	10.579	0.103
N	13,264	13,264	13,264	13,264	13,264	13,264

Notes: The table shows the correlation between the presence of a high-rank officer in an event and the reporting margins and event characteristics. Data and dependent variables as defined in Table 1, time to dispatch measures length of the interval between communication and dispatch arrival in minutes and high-risk is the ex-ante risk assessment indicator used in Table 1. 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-day level. *** p<0.01; ** p<0.05; * p<0.1.

Table A7: Characteristics of Notable BWC Studies in the Literature

Paper	Year (1)	# Citations (2)	Country (3)	(Quasi) Experiment (4)	Unit of (quasi) randomization (5)	N (6)	T (7)	C (8)	Share of Treated Units (9)	Avg # of officers per dispatch (10)	Analyisis Unit (11)	Time period (12)	UoF as outcome? (13)	Effects on UoF (14)	Inference method (15)
Panel A: Shift centric studies															
Ariel et al. (2015)	2014	633	US	Yes	Shift	988	489	499	0.495	1	Shift	None	Yes	Decrease	Poisson regression
Ariel et al. (2016a)	2016	178	US	Yes	Shift	4,915	2,447	2,468	0.498	-	Shift	None	Yes	Null	Means test
Ariel et al. (2016b)	2016	221	US	Yes	Shift	4,915	2,447	2,468	0.498	-	Shift	None	Yes	Null	Means test
Ariel et al. (2017)	2017	187	US	Yes	Shift	3,882	1,908	1,974	0.491	-	Shift	None	No	-	Means test
Henstock and Ariel (2017)	2017	78	UK	Yes	Shift	430	215	215	0.500	1	Shift	None	Yes	Decrease	Odds-Ratio
Ariel et al. (2018)	2018	48	US	Yes	Shift	4,915	2,447	2,468	0.498	-	Shift	None	No	-	Odds-Ratio
Magaloni (2019)	2019	-	BR	Yes	Unit-shift	21,472	16,390	18,642	0.468	1+	Officer-shift	Shift	Yes	Null	OLS
Panel B: Officer centric studies															
Jennings et al. (2015)	2015	243	US	Yes*	Officer	89	46	43	0.517	1+	Officer	None	Yes	Decrease (check)	% change
Ready and Young (2015)	2015	239	US	Yes*	Officer	3,698	50	50	0.500	1+	Contact report	None	No	-	HGLM
White et al. (2017)	2017	84	US	Yes*	Officer	298	82	67	0.550	-	Officer	Pre-post	Yes	Null	DiD
Jennings et al. (2017)	2017	62	US	Yes	Officer	120	60	60	0.500	-	Officer	Pre-post	Yes	Decrease	% change **
Headley et al. (2017)	2017	89	US	Yes*	Officer	103	26	25	0.510	-	Officer	Pre-post	Yes	Null	% change
Braga et al. (2017)	2017	62	US	Yes*	Officer	832	218	198	0.524	1	Officer	Pre-post	Yes	Decrease	DiD
Braga et al. (2018)	2018	107	US	Yes*	Officer	832	218	198	0.524	1	Officer	Pre-post	Yes	Decrease	DiD
Peterson et al. (2018)	2018	-	US	Yes	Officer	504	252	252	0.500	-	Officer	Pre-post	Yes	Null	DiD
Wallace et al. (2018)	2018	64	US	Yes	Officer	228,220	82	67	0.550	1+	Call-officer	Call	No	-	DiD
Yokum et al. (2019)	2019	32	US	Yes	Officer	1,922	1,189	1,035	0.535	-	Officer	None	Yes	Null	OLS between officer
Koslicki et al. (2020)	2020	10	US	No	-	-	-	-	-	-	Officer * Month	-	Yes	Null	Time series analysis
Braga et al. (2020)	2020	9	US	Yes	Officer + District	562	140	141	0.498	1	Officer	Pre-post	Yes	Decrease	DiD
Braga et al. (2020)	2020	-	US	Yes	Officer + Precinct	7,778	1,991	1,898	0.512	-	Officer	Pre-post	No	-	DiD
Panel C: Spatially explicit designs															
Katz et al. (2014)	2014	170	US	No	Area	2	1	1	0.500	-	Area	None	No	-	Means test
Grossmith et al. (2015)	2015	107	UK	Yes	Team	2,060	814	1246	0.395	1+	Officer	None	No	-	Means test
Morrow et al. (2016)	2016	104	US	No	Area	4	1	1	0.500	-	Area	Pre-post	No	-	Means test **
Ariel (2016a)	2016	79	US	No	District	17,726	1	5	0.167	2	Street segment	Pre-post	No	-	Means test
Ariel (2016b)	2016	92	US	No	Area	924,457	1	5	0.167	-	Call	Pre-post	Yes	Null	Odds-Ratio
Hedberg et al. (2017)	2017	160	US	No	-	44,380	22,660	22,720	0.499	1+	Incident	None	No	-	GLM
Mitchell et al. (2018)	2018	10	UY	No	Region	38	5	14	0.263	-	Region	Pre-post	No	-	Means test
Owens and Finn (2018)	2017	32	UK	Yes	Team	-	814	1,246	0.395	1+	-	-	No	-	-
Bennett et al. (2019)	2019	-	US	No	Squad areas	2	1	1	0.500	-	Squad * Week	Week	Yes	Null	Diff. in trends test
Stolzenberg et al. (2019)	2019	6	US	No	-	1	-	-	-	-	Monthly rates	Monthly	No	-	None **
Kim (2021)	2021	-	US	Yes	Agencies	6,929	-	-	-	-	Agency * Month	-	Yes	Decrease	DiD

Notes: 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 case a randomization unit is indicated with * next to a Yes it means that the officers included are partially self-selected into the experiment implying that caution needs to be put on detected effects as these could be quite specific LATE estimates. The table focuses on the respective randomization design, the outcome measurement approach, method of inference and whether effects on use-of-force (UoF) have been identified. Inference methods chosen often do not follow more rigorous program evaluation techniques, and the studies 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: Alternative designs - Varying randomization Unit

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Event level - Main specification							
Treat	23.883***	3.018***	2.717***	-0.374***	-0.266	-0.426***	-0.338
	(1.110)	(1.014)	(0.661)	(0.142)	(0.187)	(0.156)	(0.464)
Mean DV Control	0.000	32.815	13.844	0.845	0.934	0.696	5.420
N	13264	13264	13264	13264	13264	13264	13264
Panel B. Officer-day level - Between Officer							
Treat	0.128***	0.007	0.010*	-0.004***	-0.004*	-0.005***	-0.007
	(0.007)	(0.007)	(0.006)	(0.002)	(0.002)	(0.002)	(0.005)
Mean DV Control	0.125	0.411	0.190	0.020	0.021	0.015	0.107
N	10,763	10,763	10,763	10,763	10,763	10,763	10,763
Panel C. Precinct-day level - Between Shifts							
Treated Shift	0.117***	-0.017	-0.018**	-0.001	-0.001	-0.006	-0.006
	(0.011)	(0.011)	(0.008)	(0.001)	(0.002)	(0.006)	(0.006)
Mean DV Control Shift	0.023	0.354	0.173	0.009	0.010	0.063	0.061
N	489	489	489	489	489	489	489

Notes: Intention-to-treat specifications. Panel A replicates our main results of Table 1. Panel B delivers between officer estimates with officer-day unit of observation. Sample removes shifts without cameras and the specification includes precinct-by-week, day of the week and stratification bin fixed effects. Panel C shows the estimates of the between shift analysis, with precinct-day unit of observation. Specification include precinct-by-week and day of the week fixed effects. In Panel B and C, all dependent variables measure the share of events by officer/shift that had an occurrence of the outcomes as defined in Table 1. Standard errors are clustered at the precinct-by-day level. * p<0.01; ** p<0.05; * p<0.1.

Table A9: Alternative Designs - Varying Temporal Resolution

		Reporting Behaviour		Interaction Margins			
	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Between Officer - Pooled							
Treat	0.100***	0.005	0.010	-0.003	-0.006*	-0.002	-0.012
	(0.014)	(0.016)	(0.012)	(0.002)	(0.004)	(0.003)	(0.010)
Mean DV Control	0.123	0.433	0.182	0.013	0.021	0.014	0.106
N	450	450	450	450	450	450	450
Panel B. Between Officer - Month							
Treat	0.116***	0.009	0.007	-0.005*	-0.007	-0.004	-0.014
	(0.019)	(0.016)	(0.007)	(0.003)	(0.005)	(0.003)	(0.011)
Mean DV Control	0.127	0.431	0.182	0.015	0.023	0.013	0.100
N	1,350	1,350	1,350	1,350	1,350	1,350	1,350
Panel C. Between Officer - Day							
Treat	0.128***	0.007	0.010*	-0.004***	-0.004*	-0.005***	-0.007
	(0.007)	(0.007)	(0.006)	(0.002)	(0.002)	(0.002)	(0.005)
Mean DV Control	0.125	0.411	0.190	0.020	0.021	0.015	0.107
N	10,763	10,763	10,763	10,763	10,763	10,763	10,763

Notes: Intention-to-treat specifications. All estimates deliver between officer estimates including the experimental period only, excluding shifts without cameras. Unit of observation is at the police officer-shift level. All dependent variables measure the share of events by unit of analysis that had an occurrence of the outcomes as defined in Table 1. Panel A pools the data from September to December, 2018 and estimates the ITT with stratification bin, precinct fixed effects. Panel B uses officer-month unit of observation and estimates the ITT with stratification bin and precinct-month fixed effects. Panel C uses data at the officer-day level and includes precinct-week, stratification bins and day of the week fixed effects. Standard errors are clustered at the officer, precinct-by-month, and precinct-by-day level respectively. * p<0.01; ** p<0.05; * p<0.1.

Table A10: Alternative Designs - Varying Inference Method

	Dispatch Recorded	Police Report	Victims in report	Negative Interaction Index	Contempt, Resistance and/or Disobedience	Use-of-force	Handcuff and/or Arrest
		Reporting Behaviour		Interaction Margins			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Difference in means							
Treat	0.100***	0.005	0.010	-0.003	-0.006*	-0.002	-0.012
	(0.014)	(0.016)	(0.012)	(0.002)	(0.004)	(0.003)	(0.010)
Mean DV Control	0.123	0.433	0.182	0.013	0.021	0.014	0.106
N	450	450	450	450	450	450	450
Panel B. Difference in Differences							
Treat x Post	0.218***	0.001	0.019**	-0.003	-0.001	-0.006*	-0.035***
	(0.015)	(0.014)	(0.010)	(0.003)	(0.003)	(0.003)	(0.008)
Mean DV Control	0.000	0.436	0.183	0.015	0.021	0.017	0.129
N	900	900	900	900	900	900	900

Notes: Intention-to-treat specifications. Unit of observation is at the police officer level. All dependent variables measure the share of events by officer that had an occurrence of the outcomes as defined in Table 1. Panel A delivers between officer estimates including only the experimental period and treated shifts, and estimating the ITT with stratification bin and precinct fixed effects. Panel B delivers within officer estimates including the experimental period and three months before the experiment (June to December, 2018) and estimating the ITT with officer fixed effects. Standard errors are clustered at the officer level. * p<0.01; ** p<0.05; * p<0.1..

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*(4), 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*(3), 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*(3), 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*(6), 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*(2), 293–316.

Bennett, R. R., B. Bartholomew, and H. Champagne (2019). Fairfax County Police Department's Body-worn Camera Pilot Project: an evaluation.

Braga, A., J. R. Coldren, W. Sousa, D. Rodriguez, and O. Alper (2017). The benefits of body-worn cameras: New findings from a randomized controlled trial at the Las Vegas Metropolitan Police Department. *Office of Justice Program's' National*

Criminal Justice Reference Service, 1–79.

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*(4), 851–876.

Braga, A. A., C. Chandler, J. Eberhardt, D. Long, J. MacDonald, J. McCabe, J. Perlov, and J. Yates (2020). The Deployment of Body Worn Cameras on NYPD Officers. *Technical Report*, 66.

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*(3).

Grossmith, L., C. Owens, W. Finn, D. Mann, T. Davies, and L. Baika (2015). Police, camera, evidence: London's cluster randomised controlled trial of body worn video. *College of Policing* (November), 1–50.

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*(October), 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*(6), 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*(11), 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*(6), 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* (December), 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*(5), 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*(3), 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*(1), 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.

Mitchell, R. J., B. Ariel, M. E. Firpo, R. Fraiman, F. del Castillo, J. M. Hyatt, C. Weinborn, and H. Brants 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*(4), 510–524.

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*(3), 303–325.

Owens, C. and W. Finn (2018). Body-worn video through the lens of a cluster randomized controlled trial in London: Implications for future research. *Policing (Oxford) 12*(1), 77–82.

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.

Stolzenberg, L., S. J. D'Alessio, and J. L. Flexon (2019). *Eyes on the Street: Police Use of Body-Worn Cameras in Miami-Dade County*. Number January. Weston Publishing, LLC.

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*(3), 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 Jr, M. C., E. A. Rasich, S. Egrari, S. Egrari, and N. Weil (2021). BWC in Policing: Benefits and Costs. *NBER Working Paper No. 28622*.

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*(21), 10329–10332.