INCREASING COOPERATION WITH THE POLICE USING BODY WORN CAMERAS

Barak Ariel [†]

[†]   Lecturer in Experimental Criminology, Institute of Criminology, University of Cambridge, Sidgwick Avenue Cambridge, CB3 9DA UK; ba285@cam.ac.uk; +44(0)1223 767378, and Lecturer, Institute of Criminology, Faculty of Law, Hebrew University, Mount Scopus, Jerusalem, 91905 ISRAEL; barak.ariel@mail.huji.ac.il; +972.2.5882502

ABSTRACT

What can change the willingness of people to report crimes to the police? A six-month study in Denver investigated whether Body-Worn-Cameras (BWCs) can change crime-reporting behavior, with treatment officers wearing BWCs patrolling targeted street segments, while control officers patrolled the no-treatment areas without Body-Worn Videos. Stratified street-segments crime densities were used in this study as the units-of-analysis, in order to measure the treatment effect on the number of 911 calls-for-service in target versus control street-segments. Results from repeated-measures-ANOVAs and subgroup analyses suggest that BWCs lead to greater willingness to report crimes to the police in low-crime level residential street segments, but no discernable differences emerge in "hot spot" street segments. Variations in reporting are interpreted as a sign of improved police-public relations, due to changes of enhanced police accountability, legitimacy or perceived utility caused by the use of BWCs in these neighborhoods. The situational characteristics of the street segments can be used to explain why only low-level street segments are positively affected by BWCs, while in places where BWCs are needed the most – hotspots of crime and disorder – no effect was detected. More research is needed, particularly through randomized controlled trials with surveys of stratified samples of residents, in order to unmask the potential causal link between BWCs and police legitimacy, which is manifested as more cooperation with the police.

KEY WORDS: Body-Worn-Cameras, hotspots, crime-reporting behavior, cooperation, deterrence, legitimacy, utility, procedural justice

INTRODUCTION

On 07/04/2015, a white North Charleston police officer was charged with murdering Walter Scott, an unarmed black man, after shooting Scott eight times. The incident was caught on a bystander's video contradicted several parts of the officer's initial account (New York Times, 07/04/2015). Following this and similar incidents, police departments have been pushing to procure and implement the use of Body-Worn-Cameras (BWCs). At least fifteen States are currently considering "BWCs Bills," requiring police officers to wear these devices on the line of duty. Many believe that BWCs will greatly affect public health, at least through the minimization of use of force– a concern that the Scott incident brings to light more than ever before in late modernity.

Despite great promise, the level of excitement over this hypothesis overwhelmingly exceeds our evidence-base on the efficacy of these devices, on virtually all fronts. With only a few notable exceptions reviewed hereunder, there is in fact a very scant body of published, rigorous, evaluation research on the effects that BWCs might have in policing (Lum et al 2015; White 2014). Using BWCs in police operations can potentially be linked to behavioral changes in police-public encounters, such as reduced reported use of force (e.g., Ariel, Farrar and Sutherland 2014; Jennings, Lynch and Fridell 2015). However, what about other outcomes?

A somewhat neglected avenue of research thus far has been on the possibility that BWCs could effect police legitimacy. One manifestation of enhanced police legitimacy is an elevated likelihood of reporting crimes to the police (Sunshine and Tyler 2003). The more we trust the police – whether because they are viewed as "more just" or because they are perceived to be more effective, or both (Tankebe 2013) – the more we are likely to cooperate, come forward when we have witnessed a crime, and ask for assistance in emergency situations. How would BWCs come into play here? If the public believes that officers that are "being watched" by BWCs are more likely to provide them with what they consider a more professional service than officers who are not under BWCs surveillance, they might be more likely to cooperate. It is logical to assume that an individual will be more willing to come forward and report a crime to the police, if that person experiences an elevated perception of police professionalism. In other words, if BWCs introduce enhanced police accountability, greater transparency, or improved utility, then the public will view the police as more legitimate – and the behavioral display of this legitimacy is increased cooperation.

However, testing the effect of BWCs on reporting behavior presents methodological challenges. Citywide or neighborhood-level assessments, with crude before-after analyses, cannot unmask this effect of BWCs, as alternative explanations to the treatment effect cannot be ruled out (Shadish, Cook and Campbell 2002). Comparison groups are needed. At the same time, a test in which BWCs are randomly allocated to officers (that is, when the officer is the unit of analysis of the test) presents measurement difficulties as well, because crime reporting variations are unlikely to be attributed to particular officers, or even a group of officers. A more useful unit of analysis for testing the treatment effect on crime is a small geographic unit: the street segment level, and specifically 'hotspots'. Street segments present a palatable option for studying the effect of BWCs on crime reporting, for two major reasons, both of which are associated with consistency. First, there is ample evidence which suggests that crime levels – especially the "hottest" hotspots – remain stable in terms of their spatiotemporal location (Weisburd, Bushway, Lum, and Yang 2004), including the overall number of crime incidents they experience over time (Weisburd et al 2006). Therefore, if treatment is allocated to *some* street segments but not others (preferably through a random allocation of units), the untreated segments can serve as viable counterfactuals, to which the treatment effect can be compared. Second, and specifically for the framework of BWCs research, street segments are likely to receive a similar "dosage" of police attention (e.g., similar number of patrol hours and policing style), which is conditional primarily on the number of incidents the streets experiences: "hotter" hotspots receive more response policing, while "cooler" hotspots receive less (Ariel, Weinborn and Sherman 2016). Thus, these features – consistency of both crime patterns and overall police intervention – create comparable units of analysis with less "statistical noise," where the effect of BWCs (or any intervention) can be more accurately estimated.

Using this approach, I tested the effect of BWCs on street segments in Denver, Colorado over a period of six months. Street segments in one police district out of six was assigned BWCs, while all other street segments in other districts served as comparison sites without BWCs. Officers were asked to conduct patrols across the city, as they normally would, and to respond to calls for service. The target areas were established on the basis of street segments methodology (Weisburd, Groff and Yang 2012), and were then stratified according to 'heat' level, that is, the number of incidents in the year prior to the study. The stratification process created homogenous street segments (see Ariel and Farrington 2014), ranging from 1-20 incidents per year in the lower stratum, to 100 or more crimes incidents per year in the upper bound stratum, mirroring variations in other characteristics of street segments as well

(residential, night time economy, etc.). Results were then measured based on changes in "crime density" in the treatment, versus the control district segments – or the number of 911 incidents per 100 meters.

The paper begins with a narrative review of the literature on BWCs, followed by an exploration of the possible effect this technology may have on crime reporting, more broadly. The case for using street segments as the unit of analysis is then laid out, with an emphasis on the stability and stronger tests that can potentially be used in the study of BWCs (and beyond), when using these small-size, yet prevalent units. A description of the methods used in this experiment are depicted next, explaining the treatment and control conditions, the stratification procedure, the outcome measures and the statistical analyses used to estimate the causal link between police patrols with BWCs on crime reporting. After presenting the results of these analyses, the paper concludes with a discussion of the findings and the limitations of the approach.

WHY BODY-WORN VIDEOS WOULD HAVE AN EFFECT ON CRIME REPORTING?

The proliferation of Body-Worn-Cameras (BWCs) in police operations is continuously increasing worldwide. BWCs allow the police to record, both in audio and video, interactions between the police officers and witnesses, victims and offenders. While CCTV and police dashboard cameras have been widely available for some time now, BWCs allow officers to collect and store evidence from the officer's perspective. Their expansion has recently been portrayed dramatically, with massive media hype and a premier endorsement at the Presidential Taskforce on Crime, as a device that can aide in restoring the seemingly lost legitimacy of the police (wh.gov/community-policing).

Despite their popularity, research on BWCs is surprisingly weak. A recent review of the available evidence conducted by Lum et al (2015) has shown that there are, currently, 12 existing empirical studies of BWCs and about 30 ongoing research projects (see also review by White 2014). While there were attempts to implement BWCs in policing nearly a decade ago (Goodall, 2007; Harris, 2010), evidence on their effectiveness has only surfaced in the last couple of years (see Lum et al 2015). Four of the studies employed randomized controlled trials (Ariel, Farrar and Sutherland 2014; Grossmith et al., 2015; Jennings et al., 2015; Owens et al., 2014), and others have used less robust designs. For example, the most recent studies in Orlando, Florida, where the unit of analysis was the individual officer (Jennings et al 2015), volunteers versus non-volunteers (Young and Ready 2015), or some version of cluster random assignment (Grossmith et al., 2015), have all measured various consequences of the use

of BWCs in police routine operations. Lesser designs, without random allocation of units into treatment and control conditions (e.g., Ariel 2016; Ellis et al 2014), have produced mixed results about the effectiveness of BWCs, ranging from supportive evidence, to null, and even backfiring effects.

The Rialto experiment (Ariel, Farrar and Sutherland 2014) was the first randomized controlled trial that looked at the effectiveness of BWCs, and specifically focused on use of force and complaints. Rialto Police Department, a small jurisdiction in California with just over 50 frontline officers, compared nearly 500 police shifts during which all police-public encounters were equally assigned to either treatment or control conditions. During treatment shifts, Rialto officers were asked to videotape all their encounters with members of the public and to store evidence on a secured cloud. In control shifts, the officers were tasked never to carry on them and use the devices. Outcomes were then measured, in terms of officially recorded use of force incidents and complaints lodged against Rialto police officers. Following this 12-month experiment, Ariel, Farrar and Sutherland (2014) reported a relative reduction of roughly 50% in the total number of incidents of use-of-force compared to control conditions, and about a 90% reduction in citizens' complaints, compared to the 12-months prior to the experiment.

Ready and Young (2015) conducted an experiment with the Mesa, Arizona Police Department. The study analyzed nearly 3,700 field reports completed by 100 sworn patrol officers. Random assignment of the officers into treatment and control groups have resulted in several important findings: first, officers that did not wear body worn cameras were more likely to conduct stop and search, and were also more likely to make an arrest. This means that wearing BWCs may cause officers to be more cautious and risk-averse than control conditions. At the same time, treatment officers were more likely to give citations and initiate encounters. This suggests that BWCs may cause officers to be more proactive with this technology, however without increasing their use of invasive strategies that "may threaten the legitimacy of the organization" (Ready and Young 2015:445).

Finally, Jennings, Lynch and Fridell (2015) have also observed the effect of BWCs on policing, but focused particularly on response-to-resistance incidents. In their controlled experiment, they randomly assigned 46 (of 89) officers to wear BWCs, with the remaining 43 officers assigned to a no-BWC condition. The study has shown that BWCs reduced these types of incidents and serious external complaints. The prevalence of response-to-resistance incidents and the prevalence and frequency of serious external complaints were significantly less for officers randomly assigned to wear BWCs (p. 480).

Notably, the existing research focused primarily on the cooling off effect of BWCs in police-public encounters (see Ariel 2016). "Functions of the [BWC] are to record evidence of activities and behaviors relevant to a crime, deter violence or negative behavior against an officer, and improve the accountability of police officers and reduce the number of complaints against officers" (National Institute of Justice 2012). The mechanism that is assumed to causally link BWCs to these outcomes in police daily operations revolves around two complementary theories: self-awareness and deterrence theories (see Ariel 2013; Ariel, Farrar and Sutherland 2014; White 2014). First, several lines of research suggest that people alter their behavior once made aware that they are being observed (Chartrand and Bargh 1999; Dzieweczynski, Eklund and Rowland 2006; Jones and Nisbett 1971). A rich body of evidence on perceived social-surveillance - self-awareness (Wicklund 1975) and socially-desirable-responding (Paulhus 1988) - proposes that people adhere to social norms and change their conduct because of their cognizance that someone else is watching (Munger and Shelby 1989). Knowing, with sufficient certainty, that our behavior is being observed or judged, affects various social cognitive processes: We experience public self-awareness (Gervais Norenzayan 2012; Duval and Wicklund 1972) and feel the need to cooperate with rules (Milinski, Semmann, Krambeck 2002; Wedekind and Braithwaite 2002; Barclay 2004).

In criminology, these cues are more broadly explored under deterrence theory (the second theory linked to the present study). An extensive body of recent rigorous research, across several categories of criminal behavior, has shown that when certainty of apprehension for wrongdoing is "high", socially and morally-unacceptable acts are dramatically less likely to occur (Von Hirsch et al 1999; Pratt et al 2006; Nagin 2013). Particularly in terms of crime and disorder, when the consequences of apprehension can be bleak (imprisonment, fines, etc.), people simply do not want to get caught. In this framework, getting caught doing something morally or socially wrong is often registered as behavior that can potentially lead to negative consequences, which is an outcome that rational individuals tend to avoid (Klepper and Nagin 2006; Nagin 2013). Studies have, nevertheless, uncovered a propensity to avoid negative outcomes, and findings generally agree that individuals react compliantly to even the slightest cues indicating that somebody may be watching (see Bateson, Nettle and Roberts 2006; Burnham and Johnson 2005; Haley and Fessler 2005; Fehr and Schneider 2010).

It is important to note that this mechanism is true for both the offender and the police officer, as both actors are assumed to be rational calculators. Offenders are deterred by physical cues around them that indicate the probability of getting caught breaking the law, while the power-holders are equally

deterred by ecological signals that motivate them to comply with regulations, rules and norms. As such, a BWC can be viewed as an external contingency and will likely deter people against noncompliance (criminal, administrative, or both). Tilley (1993:3-5) rightly pointed out that the camera may "fire a number of mechanisms", but that one prominent preventative mechanism of a cameras is that it "reduces… [noncompliance] by deterring potential offenders who will not wish to risk apprehension and conviction by the evidence captured on videotape or observed by an operator on a screen on which their behavior is shown." Much like live observers, mirrors, or pictures of eyes, cameras can not only make us continuously conscious of the fact that we are being watched, but also drive us into compliance. When we become aware that a video-camera is recording our actions, we are conscious that unacceptable behaviors are likely to be captured on film, and the perceived certainly of punishment is at its highest. "Getting-away" with rule breaking is, consequently, far less conceivable if one is being-videotaped *and one is cognizant that the behavior is, in fact, videotaped* (Drover and Ariel 2015). Cameras can therefore be viewed as "credible threats" (Jervis 1989:3): Because the certainty of apprehension for negative behavior is more forcefully acknowledged when the BWCs are on, it follows that we can directly exert deterrence when the certainty of getting caught for noncompliance is greatly intensified (Ariel, Farrar and Sutherland 2014).

CAN BWC's AFFECT THE REPORTING OF CRIME?

Despite these encouraging findings about the impact of BWCS on the quality of police-public encounters, the BWC studies left many stones unturned. In particular, the body of work did not test the effects of BWCs on police legitimacy, criminal justice outcomes, or other facets. The use of BWCs may also have repercussions for the ways in which officers view themselves and their professional role (Jennings, Fridell and Lynch 2014), or the extent to which they feel micromanaged in an era of digital surveillance (Wain and Ariel 2014). In the present study, however, we look at another outcome of interest which has been overlooked: the effect of BWCs on reported crime more broadly. Namely, does BWCs reduce reporting of crime and disorder?

The potential effect of BWCs on crime reporting patterns is not straightforward. The mechanism deserves more attention that previously alluded to (e.g., Ellis, Jenkins and Smith 2015). We must acknowledge two rival hypotheses: that using BWCs will *reduce* crime through deterrence, and that using BWCs in police operations will *increase the reporting* of crime. On the one hand, deterrence theory requires that BWCs must be perceived as effective tools, increasing sanction threats by the virtue

of bringing offenders to justice more efficiently. If this is the case, then BWCs should have a suppressive effect on crime. On the other hand, wearing BWCs may change how victims and witness of crime perceive the utility, accountability and overall legitimacy of the police. Victims and witnesses might then think that the police are mandated to act, as they are more accountable, and therefore are more likely to report crime and disorder to the police. As a result, BWCs would have a seemingly counter-effective outcome: "more" reported crimes. These two views are elaborated on, more fully, below.

HYPOTHESIS 1: USE OF BWC's IN POLICE OPERATIONS REDUCE CRIME

The first option suggests that BWCs could reduce crime, given the deterrence effect these devices might have on members of the public. For this causal mechanism to work, we must assume that the (offending) public are made aware that cameras being used, and/or crime and criminality are not interpreted by members of the public as a one-off game. For BWCs in police operations to be linked to the decision (not) to commit an offence, we must assume that police-public encounters are constructed by people interdependently, meaning that a series of "teachable moments" (see Tankebe 2013; Tyler 2011) creates a collective memory of how officers interact with members of the public, and vice versa. When police officers implement a new tactic, or utilize a new apparatus with the aim of changing criminal behavior, its "success" relies on the assumption that rational individuals are aware of it. If the police use different tactics against offenders (e.g., "zero tolerance"), or deal with crime differently (e.g., patrol in hotspots of crime and disorder), we assume that people have learned, both directly and vicariously, about these new approaches and, consequently, change their future decisions accordingly (see Cornish and Clarke 2014; Felson, Messner, Hoskin and Deane 2002; Nagin and Paternoster 1993; Pescosolido 1992). This broadly categorizes the mechanism behind general deterrence.

In similar terms, if using BWCs in police operations were to have *any* effect on crime levels, then people must be assumed to *be aware* that officers are using BWCs. If BWCs have a suppressive effect on the decision to commit crime – by virtue of elevated perceptions of risk sanctions, as explained below – then potential offenders must know, either through personal or vicarious experiences, that the likelihood of arrests and successful prosecution have increased. Potential offenders may then alter their choices accordingly. The virtue of recording the interaction increases the sanction threat of police engagement with offenders (Loughran, Paternoster, Piquero and Pogarsky 2011), because while previously the court would have to rely on the testimony of the officer about the suspect's demeanor, the "theater" of the crime scene and what the offender said to the officer, now there is digital evidence

to build the case against the suspect. Juries can be particularly susceptible to the effect of video footage in reaching a verdict (Feigenson 2010), and it is likely that rational offenders are aware of this and would consequently change their decisions.

Research on risk sanctions shows that offenders seem to place weight on their prior subjective probabilities (Nagin, 2013b, p. 94). Within this framework, one prominent factor in the decision to commit crime is heavily influenced by their individually perceived risk of apprehension (Loughran, Paternoster, Piquero, & Pogarsky, 2011; Nagin, 1998, 2013a). Perception of this risk was found to be highly influenced by proximate influences, including objective sanction risks (Apel, 2013). We can assume that among these objective risks, vivid imagery captured on BWCs can be an important cue that inhibits criminal conduct. The perceived certainty of punishment is causally associated with less crime (Bushway & Reuter, 2008; Cullen, Wright, Blevins, Daigle, & Madensen, 2008; Lochner, 2003; Loughran, Piquero, Fagan, & Mulvey, 2012; McCarthy, 2002; Paternoster, 2010; but cf. Berk & MacDonald, 2010; Tonry, 2008). Increasing the likelihood of being caught can therefore be linked to a reduced likelihood of committing an offence. As recently noted by Ariel, Weinborn and Sherman (2016), "The risk of apprehension by sentinels is firstly associated with degree to which power-holders are perceived as capable agents of the law." If an offender holds the view that the officer does not have incriminating evidence against her, she is more likely to assume that the sanction risk is lower – compared to an instance where such evidence is present. We may therefore assume that when offenders think police officers are more effective, they are less likely to commit crimes. A BWC can therefore change the perceived probability of risk sanctions that an offender has. We have some crude evidence to illustrate this phenomenon around domestic violence: One study has shown that *repeat* calls for service for domestic abuse are reduced when police officers wear BWCs (Owens, Mann and Mckenna 2014). However the evidence is quite thin.

HYPOTHESIS 2: USE OF BWC's IN POLICE OPERATIONS INCREASES THE REPORTING OF CRIME

The second option suggests that a causal link between BWCs and reported crime exists in such a way that BWCs affect the crime *reporting behavior* of witnesses and victims. There are two major theories that lay out this hypothesis: police effectiveness or police legitimacy. While the motivations under each model are invariably different, the result in the context of BWCs is actually the same: BWCs

can alter the extent to which victims and offenders feel more confident in calling the police to solve their problems.

First – and somewhat related to the first hypothesis but from the victim's perspective – it is conceivable that people perceive BWCs as potentially enhancing the capacity of the police to do its work. BWCs can be construed as resulting in more convictions, either because the police have superior tools for bringing offenders to justice, or because officers are now more accountable for their actions. this would be case for instance if better evidence is captured on film, when more accurate testimonies are recorded on tape, or when audiovisual information gathered in the scene of the crime can be analyzed at a later stage by additional crime investigation officers. Tankebe (2009) finds support for this rational-based utilitarian model: judgments about effectiveness of police are most strongly associated with residents' willingness to cooperate with the police. If victims and witnesses construe BWCs as a method of increasing the utility of the police to resolve their problems, then learning that police officers are equipped with these devices can consequently result in a greater likelihood of reporting.

On the other hand, there is ample evidence which support a normative, "process model" as a way to account for variations in the public's cooperation with police. Under this prism, the perceived legitimacy of the police plays an important role in "promoting citizens' compliance with the law, cooperation with police, willingness to collaborate with police, help-seeking behavior, and crime-reporting practices (Fagan and Tyler, 2004, 2005; Lind and Tyler, 1988; Murphy, Hinds, and Fleming, 2008; Reisig, Bratton, and Gertz, 2007; Sunshine and Tyler, 2003; Tyler, 1990; Tyler and Fagan, 2008" (Kochel, Parks and Mastrofski 2011:896). The prominent literature on the antecedents of legitimacy has been procedural justice theory, particularly the model developed by Tom Tyler. According to the Tylerian approach, procedural justice has two main dimensions: quality of treatment, which incorporated politeness, respect and dignity, and quality of decision making. The latter reflects opportunities to participate, unbiased decisions, trustworthiness, and error correction (Tyler and Blader 2000; Sunshine and Tyler 2003). In subsequent studies, researchers have emphasized four key components – public voice, neutrality in decision-making, trustworthiness and treatment with dignity and respect – as the mechanism of this procedural justice (Tyler et al 2007; Gau 2011; Wolfe 2011). A number of correlational studies then examined the link between these four components and feelings of obligation, part of which is the willingness to report crimes to the police (see, e.g. Reisig, Gertz and Bratton 2007; Ferdik, Wolfe, and Blasco 2014). For example, Sunshine and Tyler's (2003) study of New York City residents found that legitimacy influenced people's willingness to cooperate with the police. Likewise,

Tankebe (2013) in London, as well as Murphy and Barkworth (2014) and Murphy and Cherney (2012) in Australia, have all reported similar findings about how procedural justice effects on people's willingness to cooperate with police.

What does this line of research on procedural justice and legitimacy have to do with BWCs, and by implication of people's willingness to report to the police? Similar to the police utility's model, police wearing BWCs can increase the willingness of the public to report crimes – however the mediating factor between the police officers wearing BWCs and reporting is not effectiveness (i.e., the former model), but rather the appreciation of the police as being legitimate. Once citizens are made aware that officers are equipped with BWCs, it is likely that the police are perceived as more legitimate, and this perception translates into an enhanced willingness to cooperate with police.

Under both models - enhanced perceived utility or process-based legitimacy (see Bottoms and Tankebe 2012) – the result may be the same, but the motivations are different. Murphy and Barkworth (2014:178) show that the "effect of procedural justice on victims' willingness to report crime to police is context specific. For some victim types, procedural justice is more important, while for other victim types, instrumental factors dominate their decision to report crime." However, the behavioral expression is the same: greater willingness to report crimes to the police.

TESTING THE EFFECT USING STREET SEGMENTS AS THE UNIT OF ANALYSIS

Whichever the direction of the anticipated outcome – fewer crimes or more reported crimes - we must demand that the hypothesis could be falsifiable (Popper 1954). This is not easy, theoretically and methodologically, because making the link between using BWCs and crime, or crime reporting, is difficult to observe directly. In order to measure the *general* deterrence effect of BWCs on offenders, one would require access to a sample of *potential* offenders and then ascertain whether BWCs had altered their risk perceptions and, ultimately, demotivated them from committing crime. Similarly, a survey of victims or witnesses would be required, in order to look at their preferences and decision-making processes and whether BWCs affected their willingness to come forward. The difficulty is that one would require individuals *who did not call the police* as well, in order to avoid the selection bias problem. For both samples, however, one would require comparable groups that were *not* exposed to the intervention (i.e., BWCs) – and these are challenging to locate. Control settings are even more difficult to create when the unit of analysis is the individual officer, because the distribution of victims/witnesses is generally

exogenous to the individual officer, and cannot be controlled for. Instead, the causal link between the use of BWCs and variation in crime can be achieved when the unit of analysis is the place, and more specifically the hotspot (see Weisburd et al. 2012). There are theoretical as well as practical advantages to using the hotspot as the unit of analysis.

If BWCs have a deterrent effect on crime, then the effect must be directly observed. However, measuring *general* deterrence is very difficult, as one may need access to large stratified samples or entire populations. Beyond measurement, and more precisely in the context of a proper test, it is virtually impossible to split the "general" population into treatment and control conditions, and to observe the causal estimate of the general deterrence condition, under controlled settings (Gibbs 1975). A general deterrence effect, where the entire population is affected by some condition, cannot rule out alternative explanations, as such a design lacks a counterfactual comparison group. A general deterrence effect of BWCs must assume that the general population is aware of the ways in which the treatment (BWCs) is administered and to what extent, a finding which we cannot empirically test. Likewise, crude before-after analyses with no control conditions – that is, some measure of overall crime before the use of BWCs and then again in the follow-up period – is characterized by spuriousness (Lewis 1973), and internal validity becomes highly suspect. Instead, the concept of non-specific-deterrence can best be explained as the "spatiotemporal bounded" deterrence effect of BWCs – or what Sherman et al (2014) refer to as the "*local deterrence*" effect of police presence. The local deterrence effect stipulates that rational offenders are likely to calculate the costs associated with committing certain street crimes, *in certain places,* that are more likely to result in apprehension. Within these confined places – such as hotspots or unique addresses – non-specific-deterrence can be directly measured, as the assumption that non-involved individuals are *aware* of the intervention is not only stronger, but falsifiable: the intervention is precise, definable and its non-specific-deterrence effect, or local deterrence effect, can be estimated within confounded and discernable settings.

Thus, while BWCs may exert a general deterrence effect on crime (i.e., on offenders) or might change the reporting behavior of victims and witnesses through enhanced legitimacy, the causal effect on the *overall* population would be difficult, if not impossible, to measure. However, within the confinement of certain places, the test stands some chance. This means that spatiotemporally-bounded places, which are exposed to police patrols by officers wearing BWCs, would be affected, resulting in invariably different crime counts to similar places that are not exposed to police patrols equipped with BWCs.

Enter hotspots: hotspots are palatable because they provide a stable and consistent backdrop, against which the treatment effect can be estimated. The argument here is straightforward: recent criminological research has shown that crime is heavily concentrated in discrete areas, like blocks and street segments (Braga and Weisburd 2010; Weisburd et al 2012; Sherman, Gartin and Buerger 1989). For example, Weisburd et al. (2004) demonstrated that 50% of the crime recorded in Seattle, over a 14 year period (1989-2002), occurred at 4.5% of the city's street segments. This, and other similar discoveries (see Weisburd 2015), has led to the "law of concentration of crime in place" (Weisburd et al 2012:169), which suggests that these crime concentrations, at such microdots on the map, may be universal.

Yet the prominent feature of hotspots, capable of resolving our issue with finding the right unit of analysis in the study of the link between BWCs and crime variations, is the stability of the hotspots. There are three features, in terms of stability, that are relevant here: spatiotemporal constancy, multiple units of observations (which translate into more statistical power), and police-response stability. Places such as city blocks, street segments and polygons have been shown to be relatively stable over time and space (Weisburd et al. 2012). Weisburd et al. (2004) found that the Seattle street segments that recorded the highest amount of criminal activity at the beginning of their longitudinal study were similarly ranked at the end of it. Such micro-places may remain stable because they provide opportunities for criminal activity that other areas may lack (Brantingham and Brantingham 1995): public housing (Eck, 1994; Newman 1972; Roncek, Bell, and Francik 1981), bars, clubs, fast food restaurants, liquor stores (Block and Block 1995; Brantingham and Brantingham, 1995; Roncek and Maier 1991), and bus stops (Golledge and Stimson 1997; Loukaitou-Sideris, 1999). Ariel (2011) show that the same train stations as well as the same *areas* within these stations – platforms, shops, concourse level, etc. – across England and Wales, remain consistently high over time, with no real fluctuations between their rank orders in terms of crime counts. They may also be stable because their layout attracts criminal elements, like large crowds, alcohol, or other nighttime-economy activities. Whatever the reason, we believe that certain places remain constant - and this spatiotemporal stability feature enables us to assume some degree of baseline equilibrium across the units. Consequently, applying an experimental condition in some places, but not in others, can augment the internal validity of the test.

Another feature of places that is important here is that there are plenty of them. In most major or even mid-size cities, there are necessarily dozens, if not hundreds, of street segments – depending on the size of the city and the threshold used to define the heat of the place (e.g., the number of crime counts

per year). This feature has immediate implications for scientific explorations. For instance, Maryland Scale Level 3 studies (Sherman et al 1998; for instance, see Sherman and Rogan 1995) suffer from low statistical power, as there are normally two study sites only. Comparing how BWCs would affect crime in one police district, but not in another district, lacks the necessary conditions under which we could conclude, with sufficient robustness, that alternative explanations can be overlooked. On the other hand, observing an overall treatment effect *across multiple* units of observations – that is, a sufficiently large number of street segments – increases the power of the test.

Weisburd et al (2004) have found that, for Seattle, the "super-hotspots" remain steady over years. Weinborn (2012) has found similar trends in both Peterborough and Birmingham cities in England. Perhaps more interestingly, police "dosage" at these places is relatively constant as well. Using GPS tracking technologies, both Ariel, Weinborn and Sherman (2016) and Mitchell (2015) show that the duration of police presence in hotspots of crime remain the same over time, when comparing hotspots with the same overall number of incidents.

Finally, there is some evidence to suggest that response dosage to places is relatively consistent across hotspots of similar heat (e.g., Ariel, Weinborn and Sherman 2016; Mitchell 2015). Unless special preventative operations are in place which increase the dosage of police presence or action, the "amount" of policing conducted in places is, interestingly, the same. One possible explanation of this "treatment homogeneity" is the move towards a responsive policing style: given budget constraints, police departments rely less on proactive and more on reactive policing: dealing with one call for service to the next, or conducting random preventive patrols in between these calls (see review and critique in Telep and Weisburd 2012). When construed this way, it is not surprising that, over time, the duration of police dosage within specific street segments would remain constant: the pressure to deal "with the next job" is ubiquitous, and so officers stay at the scene for as little time as possible, which seems to be evenly distributed across similar jobs across places.

Taken together, assigning BWCs in some stable places, but not in others, is an innovative approach to measuring the treatment effect on crime *within* street segments, because we have some degree of controlled settings. Both crime levels, and the reaction to crime, are consistent over time at these places (and can be assumed to correlate as well), which we can observe across multiple units of observations. When some street segments receive the treatment, it is exogenous and therefore any crime variations that follow are hypothesized to be a result of the treatment, not other variables. Still, while this approach provides *some* degree of baseline equilibrium between treatment and comparison sites when conducting

quasi or pre-experimental studies, it does *not* replace a true randomized controlled design in which the units (street segments) are randomly allocated into treatment and control conditions. The new approach provides a workable compromise for when researchers cannot conduct random assignment, and therefore any result should, nevertheless, be tempered by the necessary caveats.

THE DENVER BWC STUDY

The hypothesis that wearing BWCs in routine police operations would affect crime or crime reporting in hotspots of crime and disorder was tested in Denver. All response police officers in one district – which happened to be the busiest in terms of crime, night life and human traffic – were assigned BWCs, while the other districts were not given BWCs at all. Officers were tasked to patrol the streets "as they normally would," in both treatment and comparison hotspots, while the only observable difference between the two arms was the BWCs. "Crime density," or the number of crimes reported by members of the public to the police, within 100 meters of street segments, before deploying the BWCs were measured, and then again during the six months of the experiment, across all hotspots in the city (n=27,003 street segments).

METHODS AND DATA

SETTINGS AND PROCEDURE

Denver County, Colorado, is a 155 square miles city of roughly 650,000, 45% non-White residents (10.2% African-Americans). There are disproportionally more persons living below the poverty line in the city as compared to the rest of the state (19% vs. 13%), however, residents of Denver have a similar per-capita income as the rest of the State - about $33,000 P.A (US Census Bureau 2014). In terms of crime, Denver experiences normal crime patterns (e.g., 6 homicides per 100,000). The police department is one the 50th largest police departments, with nearly 1,500 sworn officers working in six geographic districts. Each district is then broken down into many precincts, and each precinct is patrolled by up to two officers in a police car.

BWCs were allocated to all frontline officers in one district (n officers=119) for a period of 6 months during the second half of 2014, but not to any other of the five geographic districts (n

officers=513)[1]. The street segments within this geographic district are therefore the target areas, with street segments within each of the five other districts serving as comparison sites. The process of implementing the BWC in the field followed a tight protocol: The officers were instructed to have the cameras on during every shift, and to video-record all their engagements with members of the public, without discretion on when to turn the cameras on or off. Every enforcement encounter, including every 911 call for service, was required to be recorded. This was a vital aspect of the trial, particularly given issues with implementation experienced in other BWC studies (see Drover and Ariel 2015; Owens et al 2014). Citizens were told they are being recorded during encounters, which again is a vital aspect of the intervention; as recently suggested by Ariel et al (2016), this would be key to the outcomes - citizens need to know for any effect to occur: "[C]ameras should remain on throughout the entire shift and each interaction with citizens should be prefaced by a verbal reminder that the camera is present" (Sutherland and Ariel 2014). We argue that the verbal reminder delivered by the officer wearing the camera provides a mechanism similar to a nudge (Thaler & Sunstein 2008): 'rules of conduct' are in play – common courtesy from officer and citizen for one. Pushing this further, we argue that the verbal prompt is a mechanism that drives mental processing of the situation towards the rational-deliberative mode of thought (Kroneberg, Heintze and Mehlkop 2010), thus enabling the hypothesized deterrent effect of the camera to actually operate on officer and/or citizen." At the end of every shift, officers in the treatment site were required to return to their stations, and to download, tag and register all the incidents during which evidence was captured on tape.

The implementation of the study received a great deal of local media coverage[2], as well publicizing the new policy by the Denver Police Department itself in various outlets, including YouTube[©3]. This is an important element of the analysis, because the second hypothesis would only be applicable if people knew that the police were equipped with BWCs. On the other hand, one primary limitation of the study

---

[1] Taser Inc.[©] and Evidence.com provided Denver officers with BWCs and the back office equipment to handle the digital media. For Denver Police Department, the cameras were affixed to the collar, so they were visible to those people who came into contact with the police.

[2] Watts, L., Herbabdez, B., and Stewart M. (2014). Denver police officers to wear body cameras; chief says it will create transparency with community. *Denver 7 ABC News*, 27 August 2014 (last retrieved on 02/25/2016). http://www.thedenverchannel.com/news/local-news/denver-police-body-camera-experiment-chief-will-report-on-results-wednesday; Jose, C. (2014). Denver Police could expand use of body cameras on officers. Fox 31 Denver (27.08.2014). retrived on 02/http://kdvr.com/2014/08/27/denver-police-could-expand-use-of-body-cameras-on-officers/

[3] Denver Police Department (2014). Body Cameras - Denver Police Department (published on 08/27/2014). https://www.youtube.com/watch?v=jNrb_YK4RAs; 7 News (2014). Denver Police Officers to Wear Body Cameras (publushed on 08/27/2014). https://www.youtube.com/watch?v=RPyi9ARIjK8

is that manipulation checks were not conducted. We assume yet we have no direct evidence on the extent to which the public were in fact aware of the new BWCs policy implemented by Denver Police. It can however be inferred by the media coverage that *some* portions of the public were aware of the new policy, or that some portions of the populations were aware of the devices by word of mouth, by observing and/or interacting with officers wearing BWCs.

DATA SOURCES

For the purposes of this study, access was granted to geospatial police data on all crimes reported to the police (public-initiated "911" reports of crimes) during the six months of the experiment and six months in the preceding year. The data were broken down into six police districts, and within these districts, more granular analysis, at the street segment level, was then conducted. We were particularly interested in "raw" incidents reported to the police, as opposed to the crime database, which is dependent on police action as well as administrative filtering. Many crimes are police-generated (following stop-and-account, special operations and police initiated crackdowns), and observing crime data, rather than raw 911 incident counts, would have made it difficult to differentiate between the impact of BWCs on crime administration and crime *reporting*, per se.

UNIT OF ANALYSIS: STREET SEGMENTS

There is a substantial body of literature on how to map crime and disorder spatially, and there are different methods of specifying small areas of land that tend to have a disproportional propensity to experience crime and disorder. In terms of hotspots, and from a theoretical perspective, it seems that the size of a hotspot does not change the overall pattern of a skewed concentration of events in these "hotspots", in relation to all events in the city (Weinborn 2013; see also Hart and Zandbergen 2014; Eck 2005). Whether GIS systems define a hotspot as a cluster, street segment, or the archaic method of creating arbitrary circles or grids of crime is used, usually less than 5% of the land "produces" at least 50% of crime and disorder. Different methods support the law of concentration of crime in place, which means that crime maps "can most effectively guide police action [as long as the] production of the maps is guided by crime theories (place, victim, street, or neighborhood)" (Weisburd et al 2012).

However, we see the network of streets segments – that is, streets or road sections between intersections – as the most appropriate way to operationalize the coupling between crime and place (see

Weisburd et al 2012; Newton, Partridge and Gill 2014). Among other reasons, street segments are superior to other approaches because, first, as Weisburd et al. (2004) noted, the street segment has long been seen as a key organizing unit within cities and, second, they are palatable to crime control operations. Sentinels can be directed to very precise coordinates, approachable 3-dimensional patrol targets can be designed and the potential for local deterrence to materialize – even in non-grid cities (Ariel and Sherman, 2014) - through a robust approach of "seeing and being seen" (Sherman and Weisburd 1995, however *cf.* Ariel and Partridge 2016). There may also be additional social control benefits, because police officers are more likely to be seen by non-criminal elements in these street segments (Weisburd et al 2006; Ariel et al 2016), which may both reduce fear of crime and increase public confidence through visible policing (but *cf.* Sosinski, Ariel and Shalhoub-Kevorkian, 2014).

In this study, street segments in the entire City were then stratified based on the number of incidents at the pretreatment stage. The stratification procedure was meant to create more homogenous street clusters, on the basis that "hotter" street segments are qualitatively different than "cooler" street segments (e.g., Weisburd, Lawton and Ready 2012). The treatment effect can be detected more freely, as the 'noise' created by the heterogeneity within each stratum is lower than across the strata. The strata were based on the pretreatment incident counts, with six discrete strata: 1-10, 11-20, 21-30, 31-40, 41-50, and 51 or more incidents in the six months pretreatment period.

 MEASURES AND STATISTICAL PROCEDURES

The main outcome of interest was emergency incidents (911 calls for service) in the city's street segments. However, different street segments can be of different lengths, which is why "crime density" was used instead. Crime density is calculated by dividing the number of events by the street segment length (in meters), and then multiplying by 100 meters. The outcome measure describes the number of events per 100 meters.

The first analytical procedure was dedicated to justifying the stratification. One-way analysis of variance (ANOVA) was used to test whether statistically significant differences emerged *between* the six street segment heat strata, in terms of both incident and crime density measures, at the pre-treatment stage. Next, an additional one-way ANOVA was performed *within* each strata, to test whether or not baseline equilibrium was reached.

For main effects, repeated measures of analyses were used to to test whether wearing BWCs, during patrols in the street segments, were associated with any crime density variations, as compared to control conditions. The "police districts" variable (n=6) served as the grouping variable, while crime density served in the pretreatment and post-treatment period as the outcome variable. The test was repeated six times, across the six strata based on heat. Subgroup analyses in the form of Tukey's Honesty Significant Differences (HSD; Tukey 1949) tests, comparing mean crime densities, were then performed, in order to better understand where the discernible difference was concentrated. To clarify, the main effect analyses were conducted within the six heat strata. Finally, the findings were illustrated graphically, using the estimated marginal means that were detected as a consequence of the repeated measures of analysis.

OUTCOMES

DESCRIPTIVE STATISTICS

Table 1 presents pretreatment and post-treatment raw figures on the key features. As shown, the data are broken down into the six geographic districts, with District 6 serving as the target area, while the other districts served as comparison areas. The number of incidents at baseline ranged from 19,811 to 36,790, with the treatment area situated in the upper bound of this range. There were overall 188,137 incidents, scattered around 27,003 street segments in the entire city (about 7 reported incidents per street segment). On the other hand, there were significantly fewer street segments in the treatment area (n=1,327), making up less than 5% of all street segments in the city, while street segment lengths were generally shorter as in the comparison areas. The mean number of incidents per street segment was substantially larger than any other comparison site (M=26.56; SD=58.46), or the overall mean across all districts (M=6.97; SD=23.15). The same pattern emerged for mean crime density, or number of incidents per 100 meters [(M=7.64; SD=22.35) and (M=1.98; SD=9.38), respectively].

Insert Table 1 Here

STRATIFICATION AND BASELINE COMPARABILITY

One-way between subjects ANOVA was conducted to compare the effect of the stratification criterion on the distribution of incidents, in terms of both 911 calls for service counts as well as crime density. Data on all street segments in the city that experienced at least 1 incident during the baseline period, were used (N=17,726). As noted, there were 6 strata, with 1-10, 11-20, 21-30, 31-40, 41-50 and 50 or more incidents reported to the police within six months. The stratification procedure created strata that were significantly different from one another (Table 2), in terms of both outcome variables [(F(5, 17720) = 4157.7, p ≤ .001) and (F(5,17720)) = 1512.2, p ≤ .001), respectively], with a linear increase in the number of events reported to the police as we move from Strata 1 to 6.

While we would expect that inter-strata variance to be significant, the intragroup variance (that is, within the strata), was hypothesized to be as small and not significant. This was achieved for crime density outcomes only (no statistically significant differences emerged at the .05 level), but not for 911 calls for service. One-way ANOVA tests revealed some significant differences, particularly around the lower crime strata (up to 20 incidents per street segment in six months), but not in the higher crime incident strata. [(F(5,13743)=46.46, p ≤ .001); F(5,2096)=6.23, p ≤ .001); F(5,661)=0.841, p ≤.521); (F(5,377)=1.646, p ≤ .147); (F(5,189)=0.56, p ≤ .73); (F(5,624)=1.619, p ≤ .153), respectively]. As listed in Table 3, District 6 – the target area - in general, displayed higher baseline pretreatment reported crime figures in comparison to all other districts in these cooler street segments.

Insert Tables 2 – 3 Here

TREATMENT OUTCOMES

As planned, repeated measures ANOVAs were then used to evaluate whether significant changes pretreatment to post-treatment were experienced in the Treatment district, compared to the five comparison districts, within each heat strata. Table 4 lists the results of the statistical analyses, including the degrees of freedom, mean squares and F-scores. As shown, significant F-scores for the effect of time were found for all strata, suggesting that overall crime density scores were different in the pretreatment compared to post-treatment periods, in strata 1 through 6 [(F(1, 13743)=64.197, p ≤ .001);

(F(1,2096)=4.071, p ≥ .05); (F(1,661)=23.703, p ≤ .001); (F(1,377)=8.57, p ≤ .01); (F(1,189)=15.689, p ≤ .001); F(1,624)=61.303, p ≤ .001), respectively].

Of main interest is the interaction effect between treatment group and crime density, over time. These were found to be significant for Strata 1 through 3 (up to 30 crime per 6 months), but not for Strata 4 through 6 (31 and over crimes in 6 months) [(F(5, 13743)=12.209, p ≤ .001); (F(5,2096)=7.954, p ≤ .001); (F(5,661)=2.685, p ≤ .05); (F(5,377)=1.199, p ≥ .05); (F(5,189)=.481, p ≥ .05); F(5,624)=1.279, p ≥ .05), respectively]. This indicates that a treatment effect at 'cooler' street segments on calls for service, but no effect for 'hotter' street segments.

Insert Tables 4 and 5 Here

Tukey's HSD subgroup analyses were performed on mean crime densities in each stratum, in order to detect between which pairs of comparisons – that is, between treatment hotspots and comparison hotspots – the significant difference is concentrated. As shown in Table 5, virtually all comparisons are significant in Strata 1 and 2 (up to 20 incidents in 6 months) but one pairwise comparison was statistically significant between treatment area and control area number 5 (mean difference 2.78; SE=.913 95% CI .168-5.85). At the same time, one pairwise comparison in Stratum 6 (51 crimes or more in 6 months) yielded significant outcomes at the .05 level, which might be attributed to random variations between the treatment and control conditions.

The outcomes are displayed in Fig. 1 through Fig. 6 below, which present the estimated marginal means of crime densities, for each stratum. Each line in the graphs represents the pre-post comparisons in terms of the interaction effect. These results, which are born out of the repeated measures ANOVA, illustrate that as the degree of "heat" increases – signified by the increasing number of crimes at baseline as we move from strata 1 to 6 – the effect of BWCs becomes less apparent. While crime density scores remains relatively stagnated in comparison sites the lower strata, with little changes over time, crime density scores in the treatment area (marked by the dashed lines) increased. This pattern is noticeable for Stratum 1, 2, 3 and to some extent 4 as well. On the other hand, as we move towards the hotter hotspots, particularly in street segments with 40 or more calls for service in 6 months, the overall pattern

across all areas – both treatment and comparison districts – is consistent, with an overall decrease in crime densities, over time across both arms, thus illustrating the nonsignificant differences in these strata.

Insert Fig. 1 through 6 Here

DISCUSSION AND CONCLUSIONS

BWCs are one of the most sought after technological solutions in policing in the world today. Yet research evidence on their efficacy is severely lacking (White 2014; Lum et al 2015). The present study looked at one aspect: to what extent, if any at all, does using BWCs, in police operations, influence crime figures. Conceptually, BWCs can be hypothesized to *reduce* crime via a general deterrence *message,* insomuch as BWCs give the police enhanced technologies to effectively apprehend and punish offenders. Under this prism, risk sanction threats are elevated, and rational offenders will commit fewer noticeable offenses. Domestic violence, public disorder and traffic violations stand out as likely candidate categories in which we should expect fewer repeat calls for service, but arguably more candidate categories may also be impacted. At the same time, another likely hypothesis is that the use of BWCs effects crime *reporting* behavior. Under this approach, victims and witnesses will report more crimes to the police as a result of the police using BWCs. Police legitimacy, stemming out of either enhanced perceptions of procedural fairness, increased effectiveness, may be the key mechanism that operates here, as BWCs send the public the message that reporting is dealt with professionally, fairly and effectively.

Methodologically, answering *any* question about the citywide, or even statewide, effects of police policies is particularly challenging, as non-spurious relationships between independent and dependent variables are difficult to attain. Internal validity is at risk. Without proper comparison groups, under controlled conditions, the treatment effect is easily masked or becomes susceptible to rival explanations. Only true randomized experiments can comfortably assume that baseline comparability between the intervention arm of the study (whatever the intervention is) and the comparison arm, has been achieved – which is the key scientific framework for observing the causal estimate of the intervention. For this

reason, assuming we can directly observe the effect of BWCs on crime, is challenging. It is not falsifiable if it lacks proper counterfactual conditions.

I tried to overcome the lack of random assignment into treatment and control conditions by utilizing a large number of street segments and observing the treatment effect within these places. Given the large number of units, street segments as the unit of analysis greatly enhance the statistical power of the test. With all things being equal (but *cf.* Weisburd et al 2001 and more recently see Nelson, Wooditch and Dario 2014), larger samples are more stable and they make them more stable. By comparison, a study comparing the effect of using BWCs in one, or even several jurisdictions, compared to another jurisdiction(s) without BWCs, will still have a relatively small number of degrees of freedom and, by implication, the statistical power of the test will be lower. Street segments, on the other hand, exist in the thousands. Indeed, the variability between the street segments can be quite extensive – for instance in terms of the 'heat' of the segment, the types of problems it experiences, etc. – but in fact we can take advantage of this variability. A multitude of techniques can be used to reduce the 'noise' between units of analysis, so that the treatment effect can be singled out more robustly: stratification, matching and other techniques (see Ariel and Farrington 2014) – these are all valid designs that make the use of street segments superior to larger spatiotemporal units of analysis.

In addition, there is a growing body of literature demonstrating that, overall, places remain predictably constant, consistent, and stable – at least the hotspots on the top of the list, in terms of incident counts. Their size changes very little over time, the frequency of calls for service is generally stagnant and their "position" in the rank-order of hotspots in the city is generally unchanged (see Weisburd et al 2004). These stabilities are useful, because they can be utilized to assess the introduction of an intervention to some places but not others, with a stronger contention of baseline equality.

The present methodological approach gets asymptotically closer to a valid test – clearly more than a study with larger geographic units of analyses such as entire neighborhoods or communities, and certainly more than crude before-after analyses without comparison groups, designs which teach us very little about causal effects (Sherman et al 1998). Yet to be sure, this is still not a proper randomized controlled trial. We ought to acknowledge that the risks to the internal validity of the test were not fully removed. It is also the first study of its kind, so we must wait for additional (and more rigorous) causal studies in this area of research. At the same time, the results are promising.
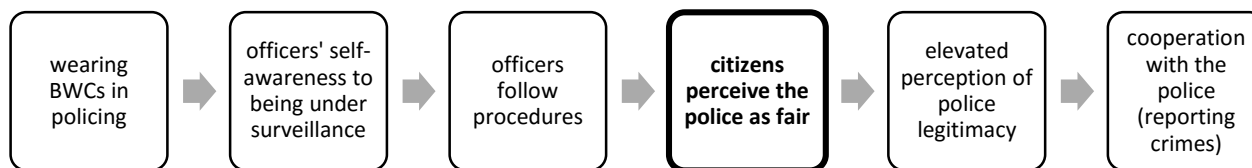
Overall, we find little to no evidence of the deterrence effect of BWCs on general crime patterns. Taken as a whole, the findings show no suppressive effect on crime, when observing overall crime

variations. When breaking the data into heat strata, at no point can we see that crime density before-after differences – the main outcome variable in this study – are lower in the treatment sites than the comparison sites. Thus, if the estimates presented here are valid, they debunk the hypothesis that suggests BWCs can change general sanction risk perceptions.
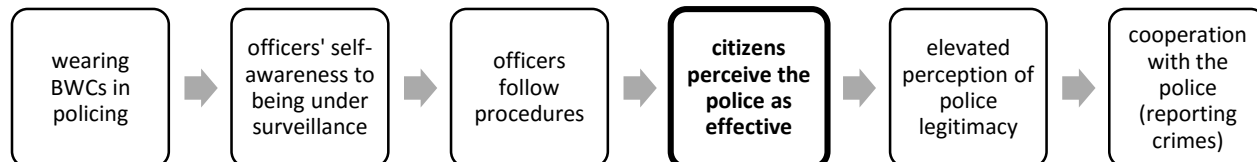
On the other hand, when BWCs were introduced in the target street segments, they seem to overall increase people's willingness to report crimes to the police, compared to street segments without BWCs. If these findings are valid, they provide direct evidence on the effect of BWCs on police legitimacy, of which one manifestations is an elevated level of cooperation with the police. Equipping officers with BWCs enhances the sense of their legitimacy, and as a consequence this promoted citizens to report more incidents reported to the police. This mechanism a closer look.

Legitimacy scholars have thus far provided observational evidence on the willingness to cooperate with the police. These studies defend the view that the inclination to work together with the police 'against' crime is one expression of legitimacy. In particular, perceptions of procedural justice in the implementation of policies served as strong correlates of cooperation with the police (e.g., Huq, Tyler and Schulhofer 2011; Tyler, Schulhofer and Huq 2010). Tankebe (2013) analyzed data from household survey of over five thousand Londoners and his results showed that procedural justice was strongly correlated with public willingness to cooperate with the police, both among victims and non-victims of crime (see also Murphy and Barkworth 2014). Using survey data from a random sample of 1,203 Australians, Murphy and Cherney (2012) reported similar findings about procedural justice effects on people's willingness to cooperate with police.

Enter BWCs: These devices are assumed to cause officers to follow procedures more closely, as they elevate the risk of apprehension for not complying with guidelines. Officers become self-conscious when being observed by BWCs, and as a consequence they are also substantially more likely to follow rules of conduct, than officers who do not wear BWCs. Drawing from the self-awareness literature (Duval and Wicklund 1972), individuals are likely to follow accepted mode of behavior, and in the police profession this awareness translates into fair, respectful and courteous conduct. These key components – public voice, neutrality in decision-making, trustworthiness and treatment with dignity and respect – are the mechanisms of procedural justice (Tyler et al 2007; Gau 2011; Wolfe 2011). The public responds to these officers' behavioral changes, and views the police as more legitimate. The immediate manifestation is an elevated willingness to cooperate with the police. This causal mechanism is depicted in Illustration 1 below:

wearing BWCs in policing	→	officers' self-awareness to being under surveillance	→	officers follow procedures	→	**citizens perceive the police as fair**	→	elevated perception of police legitimacy	→	cooperation with the police (reporting crimes)

However, other dimensions predicate police legitimacy as well, and these dimensions can equally be affected by knowing that police officers wear BWCs. Distributive justice (Bratton, and Gertz 2007) or lawfulness (Tankebe 2013) have been found to matter when it comes to legitimacy. Transparency and accountability are likely to have an effect on legitimacy as well. Bottoms and Tankebe (2012) and Tankebe (2008), who view legitimacy as substantially more complex than merely being comprised of fair procedures, suggest that one critical dimension that plays part here is police effectiveness. It is therefore logical to speculate why citizens would view the police as more *effective* with BWCs – as more fully developed under Hypothesis 2 above – and as a consequence report more crimes to the police. This alternative interpretation is shown in Illustration 2 below:

wearing BWCs in policing	→	officers' self-awareness to being under surveillance	→	officers follow procedures	→	**citizens perceive the police as effective**	→	elevated perception of police legitimacy	→	cooperation with the police (reporting crimes)

This study does not provide direct evidence on neither interpretations, but we can speculate that at least one of these versions is logical. These processes would be better addressed using a citizen survey, which would have shed much light on the interpretation of findings. However, while this study cannot pierce the walls of the "black box" of how callers to the police make decisions to call the police, it can observe that *objectively* higher levels of cooperation with the police were recorded. The data demonstrate behavioral changes in crime reporting patterns that took effect once officers use BWCs – which by itself is novel in the police legitimacy literature, compared to studies that focus on *willingness* to cooperate with the police.

Next, what best explains why is the effect of BWCs on crime-reporting behavior is conditional on street segment 'heat'? Why BWCs would have no impact in the hottest hot spots – where they are most needed - while the hypothesized mechanism works in street segments that need it the least? As shown in Figures 1-6, crime densities in treatment street segments with fewer than 80 incidents per year, there is an increase over time, attributable to the use of BWCs. On the other hand, in hotspots with more than 80 incidents per year, the general trend in the treatment group parallels the trend detected in the comparison group, which is why the differences are not statistically significant. In doing so, the outcomes suggest a more complicated story about the link between BWCs and cooperation with the police than originally anticipated.

We can only speculate at this stage, but the answer would seem to revolve around the socio-ecological context of the street segment (Weisburd 2012), as well as crime patterns within the street segments in which these measures were taken. There is no available sociodemographic database on each street segment in the city, nor an available street-level survey results of perceptions of the residents about the police more broadly. Instead, we turn to images from *Google Street View*, and argue that there are *physical* attributes of the street segments that can be used as proxies to these dimensions. A set of street view images is attached, which show the epicenter of a selection of street segments (see Appendix).

Looking at the "hot hotspots" first (see Appendix), it may in fact be unsurprising that BWCs will have no discernable effect on the decision to cooperate with the police: these are 'busy' places, with primarily a nighttime economy population. These areas are predominately entertainment places for juveniles and young adults, commercial and parks. These are areas that either offenders consciously target, or places occupied by young persons who actively engage in alcohol consumption and potentially antisocial behavior. With this in mind, we know that hotspots are largely areas that are less susceptible to change *anyway*. Detecting a treatment effect in an otherwise persistently 'hot' street segment, is difficult; Even more aggressive interventions such as hotspots policing (Braga et al 2014) has found significant yet weak effects on crime and disorder. Therefore, one plausible explanation for the lack of significant effect in 'hot' crime hotspots is the relatively weak strength of the treatment.

On the other hand, as we move from stratum 6 (high level hotspots) to cooler street segments (stratum 1), we see that the face of the street segment dramatically changes into middle (Fig. 11) and upper social class (Fig. 12) residential neighborhoods. These street segments on the lower bounds of crime are primarily suburban-looking streets with single occupancy detached or semi-detached houses. Residents of these street segments may have more stable views of the police, which in turn have been

improved with the introduction of BWCs. It is likely that some of the heat is generated by crime reports about the same places with similar *local* residents. In turn, crime-reporting behavior seems prone to being affected by the *repeated* interactions between callers and BWCs-wearing-police-officers. Police using BWCs sent repeated players a signal that altered their crime-reporting patterns. Whether by virtue of enhanced perceived accountability, stronger expected utility out of calling the police for problems, or as a result of fairer procedures on the part of the officers, the observed behavioral change is therefore more crime being reporting to the police in these places.

To be clear, these are all projected dynamics that must be explored more granularly in future research. The images presented in the Appendix are suggestive, and there was no aim to systematically assess neither the physical attributes of *all* street segments in our sample. They only serve as proxies to what would affect people's perceptions about the police. In this sense, one future research avenue which seems clear is collecting "local legitimacy" and "local self-efficacy" scores, at the street segment level – which would have broader implications that just BWCs. We did not conduct surveys of residents in these street segments – or with local offenders who may, or may not, have been affected by BWCs – so the link between attitudes towards the police, crime reporting and criminal behavior is unclear at this stage (but *cf*. Weisburd, Hinkle, Famega, and Ready 2011). Therefore, the *reasons* given here for the increases in crime reports in lower bound street segments but not in the hotspots remain contentious.

There are at least two additional limitations to this study. First, there is a strong assumption in this theoretical analysis that people were awareness of the use of BWCs. We assume that the media coverage – both locally in Denver as well as the nationwide attention – that BWCs have received, has come to the attention of the callers about crime and disorder in the target areas. We have no evidence to offer on that causal link, nor any manipulation checks. Future research should pay closer attention to this remedying this shortcoming.

Second, the study has only lasted for six months. This is undoubtedly a short period of time. Studies should always consider longer follow up period, particularly given the possibility of desensitization – namely that both officers and citizens are no longer affected by BWCs. This was not feasible under the present study conditions.

Despite these reservations, the current findings are informative for police practice. The rise in crime reporting in some street segments of Denver should be viewed as a sign of improved police-public relations, potentially as a result of changes of enhanced police accountability, legitimacy or perceived effectiveness – or both. Whichever the theoretical mechanism that explains the change in crime-reporting

behavior, the findings can comfortably be construed as a greater confidence in the police – or else callers would not have dialed 911 at increasing rates when BWCs are used. Residents in some places seem to react positively to them by becoming willing to come forward and report crimes to the police. This is an important policy-oriented discovery.

***

REFERENCES

Ariel, B. (2011). "Hot Dots and Hot Lines: Analysis of Crime in London Underground." Presented at the Annual American Society of Criminology (Washington DC, November 06, 2011).

Ariel, B., and Farrington, D. P. (2014). Randomized block designs. In *Encyclopedia of Criminology and Criminal Justice* (pp. 4273-4283). Springer New York.

Ariel, B., Farrar, W. A., and Sutherland, A. (2014). 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*, 1-27.

Ariel, B., Weinborn, C., and Sherman, L. (2016). Soft policing at hot spots - does it work? A randomized controlled trial" *Journal of Experimental Criminology*

Ariel, B. (2016). "The Effect of Police Body-Worn Videos on Use of Force, Complaints and Arrests in Large Police Departments." *Journal of Criminal Law and Criminology*, 106(1)

Ariel, B. (2013). "Effects of Body-Worn Video on Police- Citizen Encounters: An RCT". Presented at the 5th International Cambridge Conference on Evidence-Based Policing (Cambridge, UK, July 11, 2013)

Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the "tragedy of the commons". *Evolution and Human Behavior*, *254*, 209-220.

Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters*, *2*(3), 412-414.

Block, R. L., and Block, C. R. (1995). Space, place and crime: Hot spot areas and hot places of liquor-related crime. *Crime and place*, *4*(2), 145-184.

Bottoms, A. (2014). Geography of Crime and Disorder. In *Encyclopedia of Criminology and Criminal Justice* (pp. 1943-1956). Springer New York.

Bottoms, A., and Tankebe, J. (2012). Beyond procedural justice: A dialogic approach to legitimacy in criminal justice. *J. Crim. L. and Criminology*, *102*, 119.

Braga, A. A., and Weisburd, D. (2010). *Policing problem places: Crime hot spots and effective prevention*. Oxford University Press.

Braga, A. A., Papachristos, A. V., and Hureau, D. M. (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly*, *31*(4), 633-663.

Brantingham, P., and Brantingham, P. (1995). Criminality of place. *European Journal on Criminal Policy and Research*, *3*(3), 5-26.

Burnham, T. C., and Johnson, D. D. (2005). The biological and evolutionary logic of human cooperation. *Analyse and Kritik*, *27*2, 113-135.

Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*6, 893

Cornish, D. B., and Clarke, R. V. (Eds.). (2014). *The reasoning criminal: Rational choice perspectives on offending*. Transaction Publishers.

Cornish, D. B., and Clarke, R. V. (Eds.). (2014). *The reasoning criminal: Rational choice perspectives on offending*. Transaction Publishers.

Drover, P., and Ariel, B. (2015). Leading an Experiment in Police Body-Worn Video Cameras. *International Criminal Justice Review*, 1057567715574374.

Owens, C., Mann, D. and McKenna, R. (2014). The Essex Body Worn Video Trial: The impact of Body Worn Video on criminal justice outcomes of domestic abuse incidents. College of Policing. UK

Duval, S., and Wicklund, R. A. (1972). A theory of objective self-awareness.

Dzieweczynski, T. L., Eklund, A. C., and Rowland, W. J. (2006). Male 11-ketotestosterone levels change as a result of being watched in Siamese fighting fish. *General and comparative endocrinology*, *147*2, 184-189.

Eck, J. E. (1994). *Drug markets and drug places: A case-control study of the spatial structure of illicit drug dealing* (Doctoral dissertation, University of Maryland, Faculty of the Graduate School).

Eck, J. E. (2005). . Crime Hot Spots: What They Are, Why We Have Them, and How to Map Them. *Mapping crime: Understanding hot spots*, 1.

Ellis, T., Jenkins, C., and Smith, P. (2015). Evaluation of the introduction of personal issue body worn video cameras (Operation Hyperion) on the Isle of Wight: final report to Hampshire Constabulary.

Fehr, E., and Schneider, F. (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity?. *Proceedings of the Royal Society B: Biological Sciences*, *277*1686, 1315-1323.

Feigenson, N. (2010). Visual evidence. *Psychonomic bulletin and review*, *17*(2), 149-154.

Felson, R. B., Messner, S. F., Hoskin, A. W., and Deane, G. (2002). Reasons for reporting and not reporting domestic violence to the police*. *Criminology*, *40*(3), 617-648.

Felson, R. B., Messner, S. F., Hoskin, A. W., and Deane, G. (2002). Reasons for reporting and not reporting domestic violence to the police*. *Criminology*, *40*(3), 617-648.

Gervais, W. M., and Norenzayan, A. (2012). Like a camera in the sky? Thinking about God increases public self-awareness and socially desirable responding. *Journal of Experimental Social Psychology*, *48*1, 298-302.

Gibbs, J. P. (1975). *Crime, punishment, and deterrence* (p. 58). New York: Elsevier.

Golledge, R., and Stimson, R. (1997). Spatial behaviour. *Guilford, London*.

Haley, K. J., and Fessler, D. M. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human behavior*,*26*(3), 245-256.

Hart, T., and Zandbergen, P. (2014). Kernel density estimation and hotspot mapping: Examining the influence of interpolation method, grid cell size, and bandwidth on crime forecasting. *Policing: An International Journal of Police Strategies and Management*, *37*(2), 305-323.

Jennings, W. G., Fridell, L. A., and Lynch, M. D. (2014). Cops and cameras: Officer perceptions of the use of body-worn cameras in law enforcement.*Journal of Criminal Justice*, *42*(6), 549-556.

Jones, E. E., and Nisbett, R. E. (1971). *The actor and the observer: Divergent perceptions of the causes of behavior* pp. 79-94. Morristown, NJ: General Learning Press.

Klepper, S., and Nagin, D. 2006. The deterrent effect of perceived certainty and severity of punishment revisited*. *Criminology*, *274*, 721-746

Kochel, T. R., Parks, R., and Mastrofski, S. D. (2013). Examining police effectiveness as a precursor to legitimacy and cooperation with police. *Justice Quarterly*, *30*(5), 895-925.

Lewis, D. (1973). Causation. *The journal of philosophy*, 556-567.

Loughran, T. A., Paternoster, R., Piquero, A. R., and Pogarsky, G. (2011). On Ambiguity In Perceptions Of Risk: Implications For Criminal Decision Making And Deterrence. *Criminology*, *49*(4), 1029-1061.

Milinski, M., Semmann, D., and Krambeck, H. (2002). Donors to charity gain in both indirect reciprocity and political reputation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*1494, 881-883.

Miller, L., Toliver T., and Police Executive Research Forum (PERF). (2014). Implementing a Body-Worn Camera Program: Recommendations and Lessons Learned. Washington, DC: Office of Community Oriented Policing Services

Munger, K., and Shelby, J. H. (1989). Effects of an observer on hand washing in a public restroom. *Perceptual and Motor Skills*, *69*3, 733-734.

Nagin, D. S. (2013). Deterrence in the twenty-first century. *Crime and Justice*,*42*(1), 199-263.

Nagin, D. S., and Paternoster, R. (1993). Enduring individual differences and rational choice theories of crime. *Law and Society Review*, 467-496.

Nagin, D. S., and Paternoster, R. (1993). Enduring individual differences and rational choice theories of crime. *Law and Society Review*, 467-496.

National Institute of Justice (2012). A Primer on Body-worn Cameras for Law Enforcement. http://www.nij.gov/topics/law-enforcement/technology/Pages/body-worn-cameras.aspx

Nelson, M. S., Wooditch, A., and Dario, L. M. (2014). Sample size, effect size, and statistical power: a replication study of Weisburd's paradox. *Journal of Experimental Criminology*, 1-23.

New York Times (04/07/2015). "South Carolina Officer Is Charged With Murder of Walter Scott." http://www.nytimes.com/2015/04/08/us/south-carolina-officer-is-charged-with-murder-in-black-mans-death.html?_r=0 (Page last visited 04/07/2015)

Newman, O. (1972). *Defensible space* (p. 264). New York: Macmillan.

Newton, A. D., Partridge, H., and Gill, A. (2014). Above and below: measuring crime risk in and around underground mass transit systems. *Crime Science*,*3*(1), 1.

Owens, C., Mann, D., and Mckenna, R. (2014). The Essex Body Worn Video Trial. http://www.BWCsg.com/wp-content/uploads/2013/07/BWC_ReportEssTrial.pdf

Paulhus, D. L. (1988). Balanced inventory of desirable responding BIDR. *Acceptance and Commitment Therapy. Measures Package*, 41.

Pescosolido, B. A. (1992). Beyond rational choice: The social dynamics of how people seek help. *American Journal of Sociology*, 1096-1138.

Popper, K. R. (1954). *Conjectures and refutations* (p. 192). Minumsa.

Pratt, T.C., Cullen, F.T. Blevins, K.R., Daigle, L.E. and Madensen, T.D. (2006). 'The Empirical Status of Deterrence Theory: a Meta-analysis', in F.T. Cullen, J.P. Wright and K.R. Blevins (eds.) *Taking Stock: The Status of Criminological Theory*, New Brunswick, NJ: Transaction Publishers

Roncek, D. W., and Maier, P. A. (1991). Bars, blocks, and crimes revisited: linking the theory of routine activities to the empiricism of "hot spots". *Criminology*, *29*(4), 725-753.

Roncek, D. W., Bell, R., and Francik, J. M. (1981). Housing projects and crime: Testing a proximity hypothesis. *Social Problems*, *29*(2), 151-166.

Shadish, W.R., Cook., T. D., and Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Sherman, L. W., and Rogan, D. P. (1995). Effects of gun seizures on gun violence:"Hot spots" patrol in Kansas City. *Justice Quarterly*, *12*(4), 673-693.

Sherman, L. W., and Weisburd, D. (1995). General deterrent effects of police patrol in crime "hot spots": A randomized, controlled trial. *Justice Quarterly*,*12*(4), 625-648.

Sherman, L. W., Gartin, P. R., and Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place*. *Criminology*, *27*(1), 27-56.

Sherman, L. W., Gottfredson, D. C., MacKenzie, D. L., Eck, J., Reuter, P., and Bushway, S. D. (1998). Preventing Crime: What Works, What Doesn't, What's Promising. Research in Brief. National Institute of Justice.

Sherman, L. W., Williams, S., Ariel, B., Strang, L. R., Wain, N., Slothower, M., and Norton, A. (2014). An Integrated Theory of Hot Spots Patrol Strategy Implementing Prevention by Scaling Up and Feeding Back. *Journal of Contemporary Criminal Justice*, *30*(2), 95-122.

Tankebe, J. (2009). Public cooperation with the police in Ghana: does procedural fairness matter?. *Criminology*, *47*(4), 1265-1293.

Tankebe, J. (2013). Viewing things differently: The dimensions of public perceptions of police legitimacy. *Criminology*, *51*(1), 103-135.

Telep, C. W., and Weisburd, D. (2012). What is known about the effectiveness of police practices in reducing crime and disorder?. *Police Quarterly*, 1098611112447611.

Tilley, N. 1993. *Understanding Car Parks, Crime, and CCTV: Evaluation Lessons from Safer Cities*. London: Home Office Police Department.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 99-114.

Tyler, T. R. (2011). Trust and legitimacy: Policing in the USA and Europe.*European journal of criminology*, *8*(4), 254-266.

Von Hirsch, A., Bottoms, A. E., Burney, E., and Wikstrom, P. O. 1999. *Criminal deterrence and sentence severity: An analysis of recent research*. Hart.

Wain, N., and Ariel, B. (2014). Tracking of Police Patrol. *Policing*, *8*(3), 274-283.

Wedekind, C., and Braithwaite, V. A. (2002). The long-term benefits of human generosity in indirect reciprocity. *Current Biology*, *12*12, 1012-1015.

Weisburd, D. L., Groff, E. R., and Yang, S. M. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford University Press.

Weisburd, D., Bushway, S., Lum, C., and Yang, S. M. (2004). Trajectories of crime at places: a longitudinal study of street segments in the city of Seattle. *Criminology*, *42*(2), 283-322.

Weisburd, D., Hinkle, J. C., Famega, C., and Ready, J. (2011). The possible "backfire" effects of hot spots policing: an experimental assessment of impacts on legitimacy, fear and collective efficacy. *Journal of Experimental Criminology*, *7*(4), 297-320.

Weisburd, D., Lawton, B., and Ready, I. (2012). Staking out the Next Generation of Studies of the Criminology of Place. *The future of criminology*, 236.

Weisburd, D., Lum, C. M., and Petrosino, A. (2001). Does research design affect study outcomes in criminal justice?. *The Annals of the American Academy of Political and Social Science*, *578*(1), 50-70.

Weisburd, D., Wyckoff, L. A., Ready, J., Eck, J. E., Hinkle, J. C., and Gajewski, F. (2006). Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits. *Criminology*, *44*(3), 549-592.

Weisburd, D. (2012). Bringing social context back into the equation. *Criminology & Public Policy*, 11(2), 317-326.

White, D. (2014). Police Officer body-Worn Cameras: Assessing the Evidence. Office of Justice Programs

Wicklund, R. A. (1975). Objective self-awareness. *Advances in experimental social psychology*, *8*, 233-275.

TABLES AND FIGURES

Table 1: Pre and Post Raw Figures: N incidents, N Street Segments and Segment Lengths (in Meters) and Mean Crime Density (per 100 meters)

	District 1		District 2		District 3		District 4		District 5		District 6 ¥		Total	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	**Pre**	**Post**	Pre	Post
N Incidents Pre/Post	36,790	36,418	31,665	33,346	36,504	37,667	28,119	28,883	19,811	20,892	**35,248**	**40,076**	188,137	197,282
N Street Segments	4,272		4,628		8,889		3,946		3,941		**1,327**		27,003*	
Mean Incidents Per Street Segment	8.61	8.52	6.84	7.21	4.11	4.24	7.13	7.32	5.03	5.30	**26.56**	**30.20**	6.97	7.31
	(28.28)^	(30.17)	(18.17)	(20.01)	(13.32)	(14.35)	(18.20)	(19.75)	(19.54)	(20.59)	**(58.46)**	**(66.33)**	(23.15)	(25.39)
Mean Street Segment Length	421.13 (276.96)		422.66 (240.38)		426.51 (310.60)		472.17 (302.78)		475.18 (477.40)		**404.80 (167.54)**		437.70 (319.96)	
Mean Crime Density (All Segments)*	2.74	2.92	1.94	2.08	1.17	1.27	1.72	1.68	1.38	1.30	**7.63**	**6.45**	1.98	2.07
	(12.21)	(17.41)	(6.53)	(9.46)	(6.34)	(10.39)	(6.08)	(5.18)	(9.04)	(8.05)	**(22.35)**	**(23.50)**	(9.38)	(11.75)

*^ figures in parentheses are standard deviations; * including segments with 0 incidents in pre-treatment; ¥ target area*

Table 2: Justifying the Stratification Method: Analyses of Variance between Six Street Segment Strata of "heat" ‡: 911 Calls for Service Counts and Crime Density per 100 meters

		N Street Segments	Means per Segment (S.D.)	*F Scores*
Crime Incident Counts	Stratum 1	13,749	3.64 (2.56)^	
	Stratum 2	2,102	14.44 (2.77)	
	Stratum 3	667	24.91 (2.85)	
	Stratum 4	383	35.15 (2.92)	4157.70***
	Stratum 5	195	45.14 (2.84)	
	Stratum 6	630	109.35 (99.49)	
	Total	*17,726*	*10.61 (27.89)*	
Crime Density (Incident Count per 100 meters)	Stratum 1	13,749	1.02 (1.24)	
	Stratum 2	2,102	3.62 (2.97)	
	Stratum 3	667	6.67 (5.63)	
	Stratum 4	383	9.63 (8.79)	1512.20***
	Stratum 5	195	12.06 (10.08)	
	Stratum 6	630	33.78 (49.08)	
	Total	*17,726*	*3.01 (11.44)*	

^ *Figures in parentheses are standard deviations;* $^{*}p \leq .05;$ $^{**}p \leq .01;$ $^{***}p \leq .001$

‡ *Stratum 1= 1-10 calls for service; Stratum 2= 11-20 calls for service; Stratum 3= 21-30 calls for service; Stratum 4= 31-40 calls for service; Stratum 5 = 41-50 calls for service; Stratum 6 = 51> calls for service, in a six-months pre-treatment period in the year preceding the experiment*

Table 3: Pre-treatment Scores within 6 Heat Strata X 6 Geographic Districts: Analyses of Variance							
Heat Strata		District 1	District 2	District 3	District 4	District 5	District 6[¥]	F Scores
Stratum 1	N	2,355	2,743	3,885	2,222	2,117	427	
	Density	1.25 (1.63)^	1.01 (0.97)	0.87 (1.18)	0.99 (1.13)	0.97 (1.01)	1.57 (1.84)	**46.46*****
	Crimes	4.06 (2.62)	3.74 (2.56)	3.17 (2.37)	3.78 (2.61)	3.42 (2.50)	5.15 (2.92)	**76.22*****
Stratum 2	N	400	430	390	393	230	259	
	Density	3.77 (2.94)	3.84 (2.92)	3.90 (3.70)	3.12 (2.15)	3.01 (3.68)	3.93 (1.83)	**6.23*****
	Crimes	14.36 (2.73)	14.53 (2.88)	14.26 (2.81)	14.40 (2.70)	14.23 (2.64)	14.95 (2.72)	**2.545***
Stratum 3	N	152	118	133	87	50	127	
	Density	6.92 (6.54)	6.37 (3.51)	7.30 (7.31)	6.28 (6.61)	5.69 (5.73)	6.65 (2.30)	**0.841**
	Crimes	24.72 (2.79)	24.65 (2.68)	25.08 (2.96)	25.13 (2.80)	24.96 (2.84)	25.00 (2.99)	**0.548**
Stratum 4	N	61	75	70	58	24	95	
	Density	12.08 (12.30)	8.81 (5.01)	8.28 (7.49)	8.88 (10.29)	11.10 (15.08)	9.78 (5.45)	**1.646**
	Crimes	35.82 (2.87)	34.95 (2.77)	34.90 (3.14)	35.12 (2.87)	35.33 (3.07)	35.04 (2.90)	**0.861**
Stratum 5	N	32	34	45	28	13	43	
	Density	11.65 (6.93)	12.59 (7.78)	12.90 (13.01)	9.29 (5.04)	12.27 (24.02)	12.82 (4.56)	**0.56**
	Crimes	44.50 (2.45)	45.15 (2.38)	45.87 (2.97)	44.11 (2.74)	43.62 (2.66)	46.00 (3.06)	**3.377****
Stratum 6	N	137	76	112	91	53	161	
	Density	37.29 (55.69)	35.65 (32.97)	29.88 (43.73)	22.20 (30.03)	37.41 (65.83	37.98 (54.42)	**1.619**
	Crimes	102.96 (118.28)	106.34 (81.61)	96.31 (50.03)	94.46 (65.06)	125.00 (107.66)	128.52 (123.60)	**2.404***

^ *figures in parentheses are standard deviations;* *$p \leq .05$;* **$p \leq .01$;** ***$p \leq .001$*

[¥] *target area*

Table 4: Repeated Measures of Analysis: Crime Density Outcomes Within Six Heat Strata

Strata X Variables		DF	Mean Squares	F
Stratum 1[‡]	Crime Density	1	883.123	64.197***
	Crime Density * District	5	167.949	**12.209*****
	Error term	*13,743*	*13.756*	
Stratum 2	Crime Density	1	78.465	4.071
	Crime Density * District	5	153.302	**7.954*****
	Error term	*2,096*	*19.273*	
Stratum 3	Crime Density	1	882.628	23.703***
	Crime Density * District	5	99.993	**2.685***
	Error term	*661*	*37.237*	
Stratum 4	Crime Density	1	902.787	8.857**
	Crime Density * District	5	122.177	**1.199**
	Error term	*377*	*101.924*	
Stratum 5	Crime Density	1	1387.461	15.689***
	Crime Density * District	5	42.524	**.481**
	Error term	*189*	*88.433*	
Stratum 6	Crime Density	1	69878.092	61.303***
	Crime Density * District	5	1457.468	**1.279**
	Error term	*624*	*1139.877*	

*$p \leq .05$; **$p \leq .01$; ***$p \leq .001$

[‡] *Stratum 1= 1-10 incidents; Stratum 2= 11-20 incidents; Stratum 3= 21-30 incidents; Stratum 4= 31-40 incidents; Stratum 5 = 41-50 incidents; Stratum 6 = 51> incidents, in a six-months pre-treatment period in the year preceding the experiment*

Table 5: Tukey's *HSD* Subgroup Analyses Comparing Mean Crime Densities - Treatment District vs. Five Comparison District within Six Heat Strata

	Comparison Districts	Mean Difference (I-J)	S.E.	95% Confidence Interval Lower Bound	Upper Bound
Stratum 1	District 1	1.12 *	.150	.696	1.553
	District 2	1.46 ***	.149	1.032	1.881
	District 3	1.67 ***	.146	1.256	2.087
	District 4	1.46 ***	.151	1.033	1.894
	District 5	1.62 ***	.152	1.186	2.051
Stratum 2	District 1	1.47 ***	.299	.618	2.321
	District 2	1.17 **	.294	.331	2.010
	District 3	1.30 ***	.300	.448	2.159
	District 4	1.74 ***	.300	.889	2.598
	District 5	2.41 ***	.339	1.446	3.380
Stratum 3	District 1	1.14	.657	-.737	3.020
	District 2	1.14	.699	-.861	3.134
	District 3	1.02	.678	-.920	2.956
	District 4	1.77	.761	-.404	3.945
	District 5	2.78 *	.913	.168	5.385
Stratum 4	District 1	-.39	1.422	-4.467	3.679
	District 2	2.69	1.339	-1.144	6.526
	District 3	3.67	1.365	-.243	7.578
	District 4	2.31	1.444	-1.828	6.446
	District 5	1.57	1.980	-4.102	7.241
Stratum 5	District 1	1.99	1.761	-3.082	7.059
	District 2	1.15	1.731	-3.836	6.133
	District 3	.074	1.609	-4.558	4.705
	District 4	2.17	1.832	-3.104	7.444
	District 5	1.47	2.387	-5.408	8.340
Stratum 6	District 1	2.71	4.710	-10.751	16.179
	District 2	5.33	5.640	-10.787	21.457
	District 3	13.63	4.986	-.627	27.880
	District 4	17.61 *	5.315	2.415	32.799
	District 5	8.45	6.417	-9.868	26.822

*S.E. = standard errors; * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$*

Fig. 1: Estimated Marginal Means of Crime Densities:
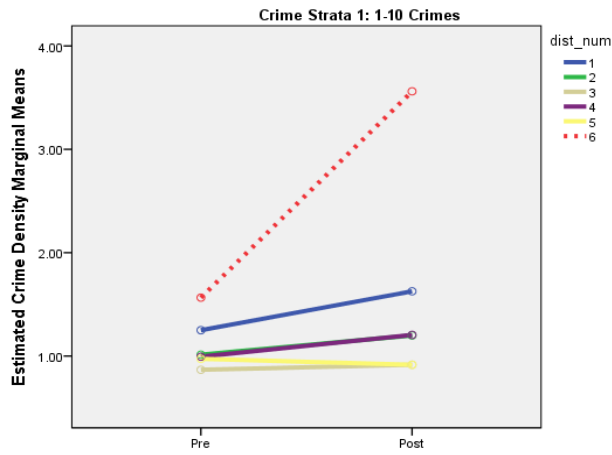**Treatment vs. Control Districts**

Fig. 2: Estimated Marginal Means of Crime Densities:
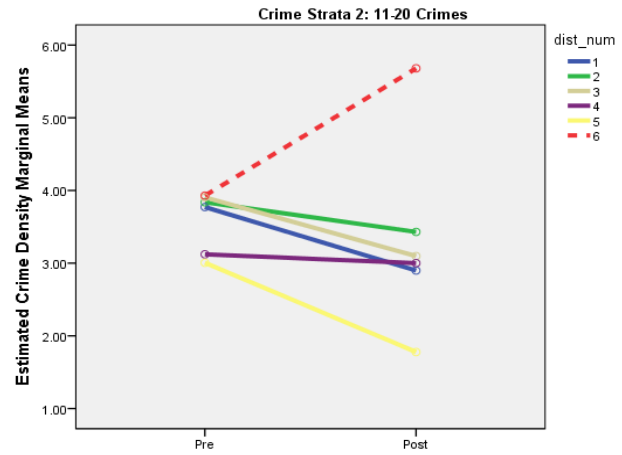**Treatment vs. Control Districts**

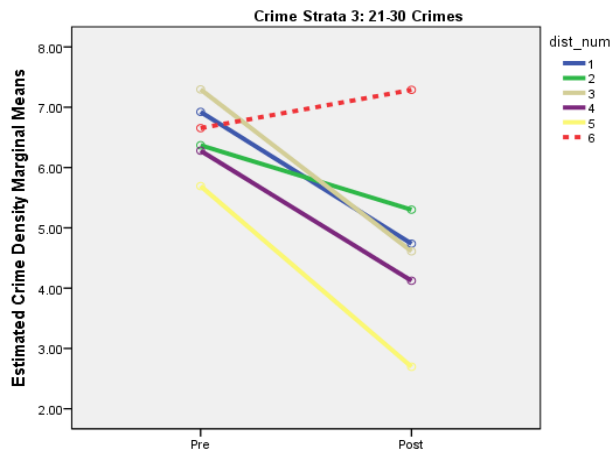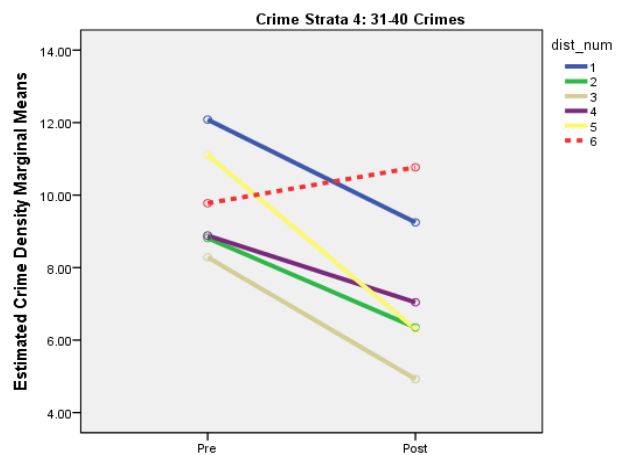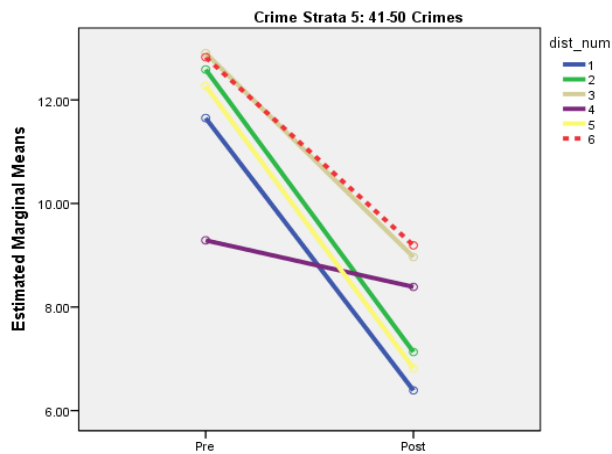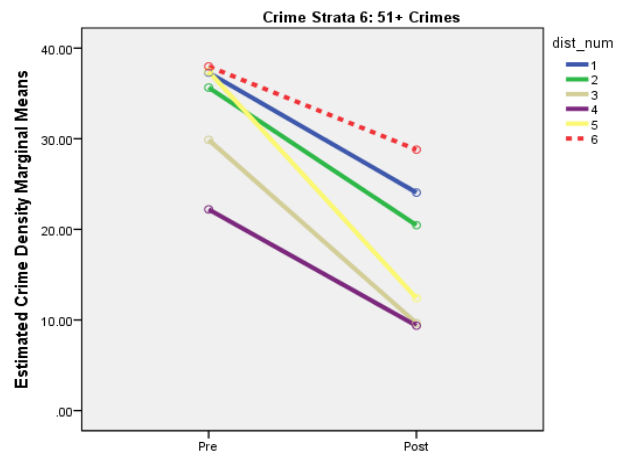**Fig. 3: Estimated Marginal Means of Crime Densities:**
Treatment vs. Control Districts

Fig. 4: Estimated Marginal Means of Crime Densities:
Treatment vs. Control Districts

Fig. 5: Estimated Marginal Means of Crime Densities:
Treatment vs. Control Districts

Fig. 6: Estimated Marginal Means of Crime Densities:
Treatment vs. Control Districts

SUPPLEMENTARY MATERIALS

Google Street View images of street segments

Selection process: random from within middle range of each strata

Views: epicentre of street segment, based on longitudinal and latitude juxtaposition.

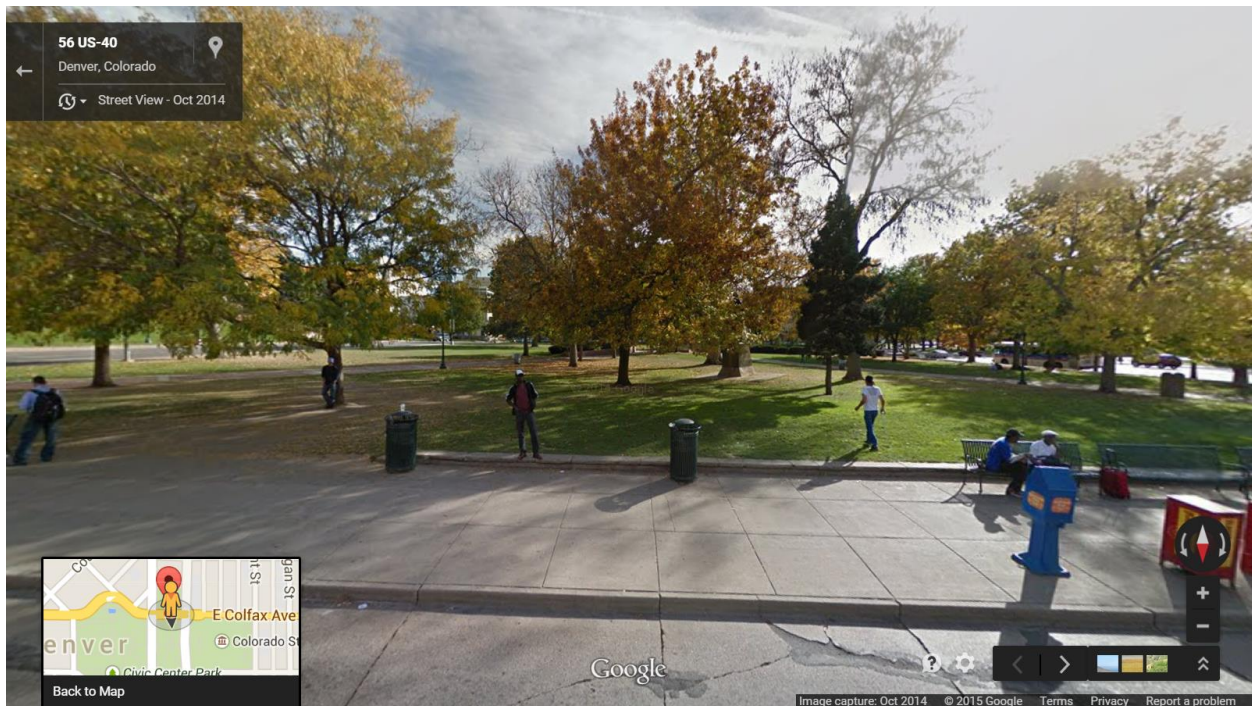**Fig. 7: Stratum 6 (more than 100 incidents per year) - Street Segment Epicentre facing the part (Google Street View)**

Fig. 8: Stratum 5 (80-99 incidents per year) Street Segment Epicentre facing commercial area (Google Street View)

Fig. 9: Stratum 4 (60-79 incidents per year) Street Segment Epicentre facing large apartment complex (Google Street View)
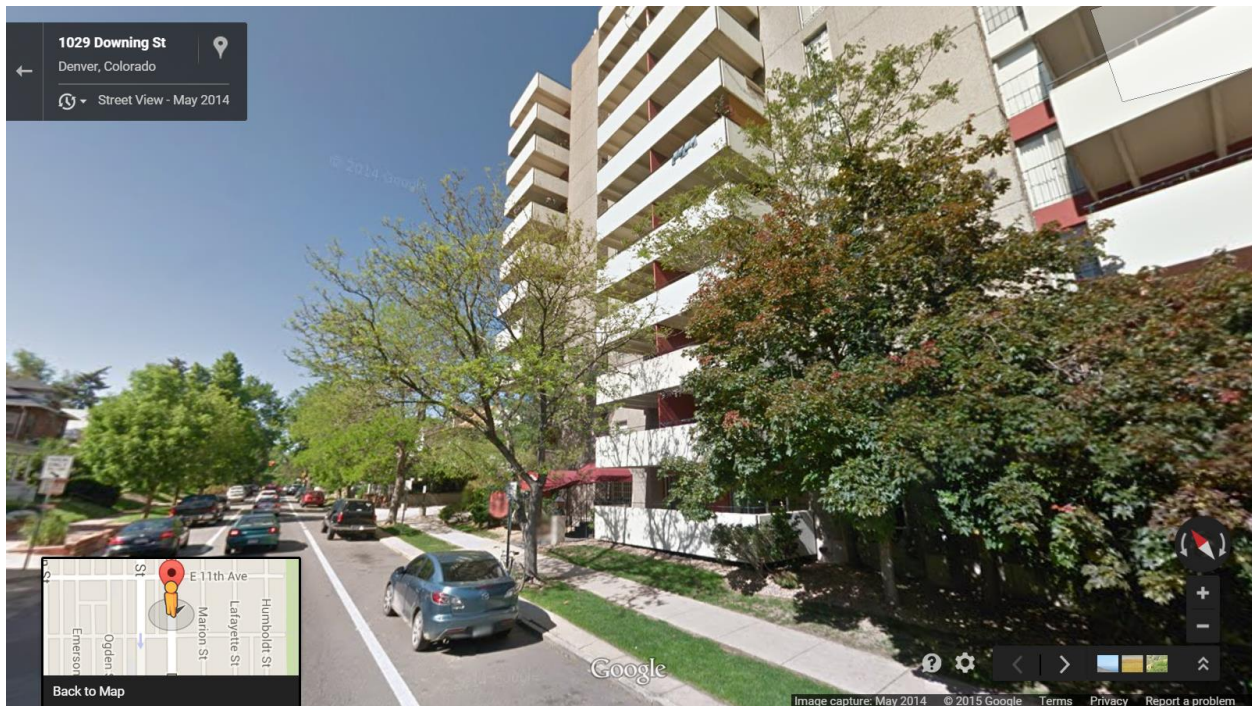
Fig. 10: Stratum 3 (40-59 incidents per year) Street Segment Epicentre facing playground (Google Street View)

Fig. 11: Stratum 2 (20-39 incidents per year) Street Segment Epicentre facing residential area (Google Street View)
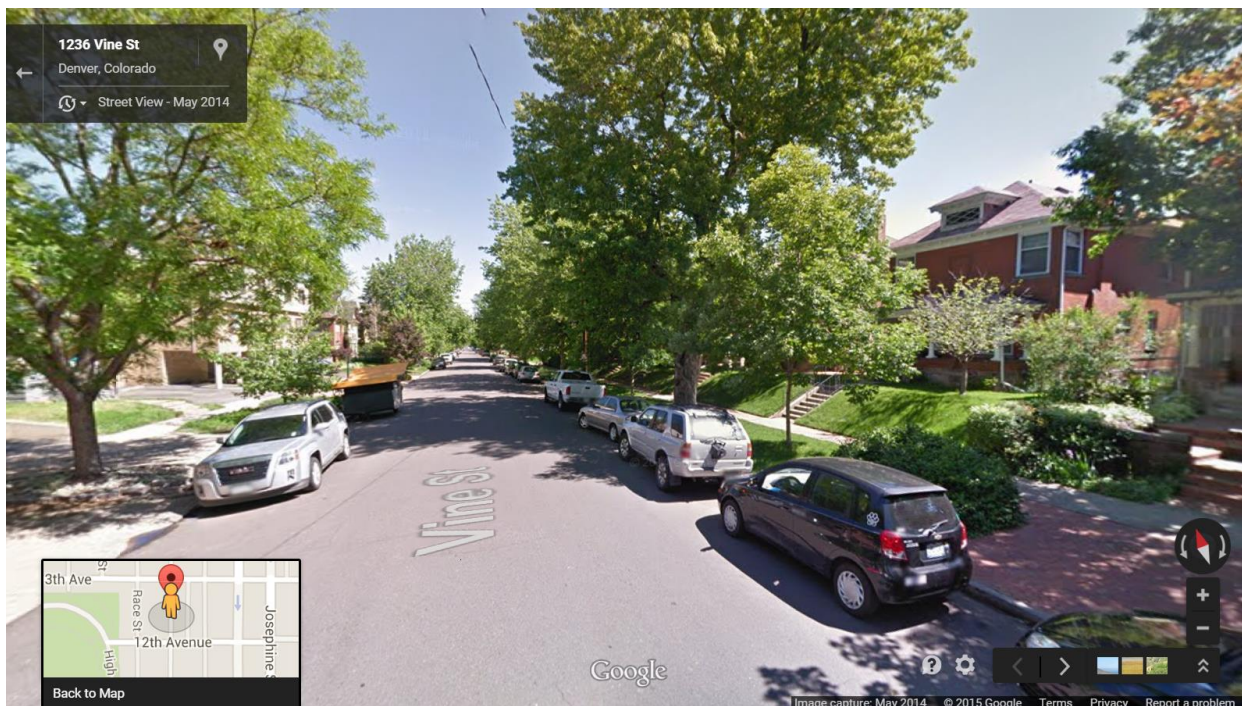
Fig. 12: Stratum 1 (up to 19 incidents per year) Street Segment Epicentre facing residential area (Google Street View)