PLACE-BASED INTERVENTIONS AT SCALE: THE DIRECT AND SPILLOVER EFFECTS OF POLICING AND CITY SERVICES ON CRIME

Christopher Blattman
University of Chicago

Donald P. Green
Columbia University

Daniel Ortega
Development Bank of Latin America (CAF) and IESA

Santiago Tobón
Universidad EAFIT

Abstract
In 2016, the city of Bogotá doubled police patrols and intensified city services on high-crime streets. They did so based on a criminological consensus that such place-based programs not only decrease crime, but also have beneficial spillovers for nearby streets. To test this, we worked with Bogotá to experiment on an unprecedented scale. We randomly assigned 1,919 streets to either 8 months of doubled police patrols, greater municipal services, both, or neither. Such scale brings econometric challenges. Spatial spillovers in dense networks introduce bias and complicate variance estimation through "fuzzy clustering". But a design-based approach and randomization inference produce valid hypothesis tests in such settings. In contrast to the consensus, we find intensifying state presence did not generate substantively meaningful or statistically significant reductions in crime. Our estimates suggest modest direct effects but with crime displaced nearby, especially property crimes. Confidence intervals suggest we can rule out total reductions in crime of more than 2% from the two interventions. (JEL: K42, O17, E26, J48, C93)

The editor in charge of this paper was Imran Rasul.

Acknowledgments: The authors thank the collaboration of the National Police of Colombia and the Mayor's Office of Bogotá, especially Bogotá's 2016–2018 Secretary of Security, Daniel Mejía. Innovations for Poverty Action in Colombia and the Center for the Study of Security and Drugs at Universidad de los Andes coordinated research activities. Survey data were collected by Sistemas Especializados de Información. For research assistance, we thank Juan Carlos Angulo, Peter Deffebach, Marta Carnelli, Daniela Collazos, Eduardo Garcia, Sebastián Hernández, Sofia Jaramillo, Richard Peck, Patryk Perkowski, Oscar Pocasangre, María A. Rodríguez-Uribe, Zach Tausanovitch, and Pablo Villar. For comments, we thank Thomas Abt, Roseanna Ander, Anthony Braga, Adriana Camacho, Aaron Chalfin, Steve Durlauf, Marcela Eslava, Claudio Ferraz, Larry Katz, David Lam, Leopoldo Fergusson, Nicolás Grau, Sara Heller, Daniel Mejía, Ben Olken, Jan Pierskalla, Tristan Reed, Jacob N. Shapiro, Rodrigo Soares, Juan F. Vargas, David Weisburd, Dean Yang, and many conference and seminar participants. Data and analysis were funded by the J-PAL Governance Initiative, 3ie, the Development Bank of Latin America (CAF), Fundación Probogotá, Organización Ardila Lülle via Universidad de los Andes CESED, COLCIENCIAS in the Government of Colombia, and the J. William Fulbright Program.

E-mail: blattman@uchicago.edu (Blattman); dpg2110@columbia.edu (Green); dortega@caf.com (Ortega); stobonz@eafit.edu.co (Tobón)

Journal of the European Economic Association 2021 19(4):2022–2051 DOI: 10.1093/jeea/jvab002

Teaching Slides
A set of Teaching Slides to accompany this article are available online as Supplementary Data.

1. Introduction

Police patrols and city services are the most elementary tools of crime control. More than 90% of US police agencies use some form of "hot spots policing" that concentrates police on the highest crime streets, for instance. Another common tactic is to improve lighting or clean up disordered places. Such "place-based" interventions focus on the locales where crime occurs rather than the people responsible.[1]

The current consensus is that place-based policies reduce total crime. Studies show that more intense or better quality policing reduces crime in high-crime areas, as does tackling social disorder.[2] Moreover, two systematic reviews argue that place-based policing not only deters crime, but that there are more instances of positive spillovers to nearby streets than negative ones. Thus, they conclude that place-based policies diffuse their benefits nearby, at least on average.[3] As a consequence, more and more countries are adopting place-based tactics, especially in Latin America (Abt and Winship 2016).

Whether place-based interventions deter or displace crime is important, not only to evaluate the policy's overall effectiveness, but because the answer has implications for our understanding of criminal incentives and behavior. In economic models, criminals weigh the returns from committing crimes against the risk of capture and expected sanctions (e.g. Becker 1968; Ehrlich 1973; Chalfin and McCrary 2017). Place-based interventions should increase people's perceived risk of detection and capture in that place, and thus reduce the likelihood they commit a crime there.[4] One might expect

1. On policing, see Weisburd and Telep (2016) and Police Executive Research Forum, (2008). On disorder, see Braga, Welsh, and Schnell (2015). On place-based theory, see Weisburd, Groff, and Yang (2012).

2. Chalfin and McCrary (2017) find that more police are usually associated with falling crime city-wide. A systematic review of hot spots policing by Braga et al. (2019) identified 65 eligible studies (including 27 experiments). Among 78 tests of the core hypothesis, 62 report improvements in crime. Exceptions include ongoing experimental evaluations in Medellin (Collazos et al. 2020) and Trinidad and Tobago (Sherman et al. 2014). For tackling social disorder, evaluations of municipal services are relatively rare. Braga, Welsh, and Schnell (2015) review interventions designed to tackle social and physical disorder, but the majority tend to be a policing strategy rather than attempts at urban renewal. There is some evidence that street lighting reduces crime (Farrington and Welsh 2008). Cassidy et al. (2014) review five studies suggesting there is weak evidence that urban renewal reduces youth violence.

3. See Braga et al. (2019) and Weisburd and Telep (2016). Two natural experiments that put round-the-clock police in areas of London and Buenos Aires also found no evidence of spatial displacement (Draca, Machin, and Witt 2011; Di Tella and Schargrodsky 2004).

4. Police presence disrupts this crime or raises the risk of capture, while city services light dark areas or increase the number of people on the street. State presence may also signal order, telling criminals to stay away and citizens that the state is present—a version of the famous "broken windows" hypothesis (Wilson and Kelling 1982; Apel 2013). Note that "Broken windows policing" is sometimes used to describe

that they choose to commit crimes in nearby areas where the crime is less likely to be detected. If not (or if there are actually positive local spillovers, as some suggest) it suggests at least one of the following hypotheses are true: that criminal rents are highly concentrated and unequally distributed within cities; that some offenders are resistant to moving crime locations; that the supply of crime is elastic to the actual or perceived risk of apprehension in a small number of areas; or that most crimes are the product of momentary emotions or opportunities.[5]

One reason for caution is the statistical power to detect spillovers. The median hot spots policing study has fewer than 30 hot spots per treatment arm. The evidence on the direction of spillovers is mixed, with studies pointing both ways.[6] While slightly more studies point to positive spillovers than negative ones, the sample sizes are such that large adverse spillovers are likely within the confidence interval.[7] Thus, we believe it is too soon to draw firm conclusions about benefits diffusing.

Most of all, the literature will benefit from larger-scale interventions. With this in mind, we worked with the city of Bogotá to design a city-wide, multi-arm security experiment. Bogotá is a large, thriving, relatively rich and developed city, in a middle-income country, with a professionalized and well-regarded police force. Out of 136,984 street segments in the city, the government identified 1,919 with moderate to high levels of crime—what we will call "hot spots".[8] Within this experimental sample, we worked with the city to randomize two interventions: increasing police patrol time; and improving the delivery of city services such as street lighting and cleanup.

We find no evidence that re-allocating state attention to these hot spots reduced crime overall. Our estimates suggest that intensifying state presence modestly reduced crime on directly-treated street segments. But we also estimate that crime increase on nearby streets, especially property crime, seeming to outweigh the direct effect. Using a 95% confidence interval to place an upper bound, we can rule out reductions in aggregate crime of more than 2–3% within the 79,767 streets directly treated or nearby. These results are striking given the intensity of the intervention and the policy consensus.

Specifically, in January 2016, a new city government in Bogotá intensified police patrol time on 756 of the 1,919 hot spots for a period of 8 months. Hot spots already

intensive, zero tolerance policing. But more visible state presence and physical order should send similar signals.

5. For discussions of these alternatives, see Clarke and Weisburd (1994); Weisburd et al. (2006); Chiba and Leong (2014).

6. One very wide review of 102 place-based interventions found indications of positive spillovers in a quarter and adverse spillovers in a quarter (Guerette and Bowers 2009). Among 40 tests for the crime displacement hypothesis included in the systematic review by Braga et al. (2019), 29 point to benefits diffusing to nearby areas, and the remaining 11 to adverse spillovers. Most studies, however, are imprecise an cannot rule out the possibility of spillovers moving in the opposite direction.

7. As we elaborate below and in an Online Appendix, some of the most positive results come from studies with fewer than ten hot spots or clusters. These studies estimate extremely narrow standard errors—so narrow, it seems unlikely they are accurately estimated (such as not accounting for clustering). Typically there is insufficient information in the papers to judge. Systematic reviews take each paper's estimates at face value, and so do not adjust for these unusually narow intervals in very small sample sizes.

8. A segment is a length of street between two intersections, a common unit of police attention (Weisburd, Groff, and Yang 2012).

received 6% of patrollers' time per day, and the intervention nearly doubled this (from 92 minutes of patrol time per day to 167, an increase of 77 minutes). Then, in March, the city chose 201 of the hot spots for more intensive street light repair and cleaning.

Both interventions reallocated existing city resources. No new police or city contractors were added. Importantly, however, they designed the intervention so that control and spillover streets would not see a fall in state presence. This is because, out of 130 street segments per police quadrant, no more than 2 streets were treated intensively—something we can confirm with geo-referenced data on the location of patrols every 30 seconds.

We used a factorial design and randomized hot spots to one of four arms: intensive policing; more municipal services; both; or neither. To identify direct effects, we compare average crime levels in treated hot spots to control hot spots. To estimate spillover effects, we draw 250-meter rings around treated and control hot spots and compare crime levels on these nearby potential spillover streets. Bogotá has geo-referenced crime data on all segments in the city, and 77,848 of these segments lie within the 1,919 250-meter rings.

In the presence of spillovers, such scale and density in an experiment brings econometric challenges. Obviously, there is potential for interference between experimental units when control hot spots fall within a treated hot spot's spillover ring. In addition, potential spillover streets may fall within the ring of both a control and treatment segment. We took a design-based approach to minimize both kinds of overlap, first randomizing police quadrants to treatment or control status, then randomizing hot spots within treated quadrants. Still, with such a large experiment in a fixed space, many rings nevertheless overlap.

This density and overlap is an identification concern for two reasons. First, this overlapping may be correlated with unobservable characteristics that drive crime, especially because the overlapping is greatest in the densest and highest-crime parts of the city. We show how this potential confounding is adequately controlled for with inverse probability weights (IPWs). IPWs transparently and systematically re-weight the locations so that there is no correlation between spillover assignment and the locations observed or unobserved attributes.

Second, we show that the close proximity of hot spots also leads to hard-to-model patterns of "fuzzy clustering" (Abadie et al. 2016). In most randomizations, segments close to hot spots tend to be assigned to the same spillover status. Failing to account for this clustering will bias estimated treatment effects and understate standard errors. Without a fixed geographic unit of clustering, however, we cannot use standard clustering correction procedures. This is a common but under-explored problem with experiments in social or spatial networks. We show how randomization inference (RI) can estimate exact p-values in these settings.

With this design-based approach and econometric corrections, we find that intensifying state presence only modestly and imprecisely reduced crime on treated streets. Crime fell roughly 0.13 and 0.15 standard deviations per hot spot from intensive policing or municipal services, respectively. Neither decrease is statistically significant. In terms of actual crimes directly deterred, over the 8 months of the interventions, these estimates imply a decline of 100 reported crimes in total.

Meanwhile, we see some evidence of adverse spillovers. On a street-by-street basis, these are small in magnitude, but with tens of thousands of nearby streets, small effects add up. We estimate that treatment increased the total number of crimes reported in nearby streets by about 800, or 2%. A 90% confidence interval includes zero, and so this should not be taken as strong evidence of adverse spillovers. As mentioned above, however, the 95% confidence interval exclude reductions in crime of 2–3% or greater, from what amounted to a major reorganization of the police force and city service provision.

These results are robust to a number of alternative measures and models. We draw similar conclusions if we model spillovers with more structure, such as continuous decay functions, or use alternative spillover ranges. Results are relatively stable over the course of the intervention. Also, we are aware that reported crimes are incomplete and reporting could be correlated with treatment. Thus we also conducted a survey of about 24,000 citizens, providing measures of unreported crimes, security perceptions, and attitudes toward the state. We estimate similar direct treatment effects with these data.

Furthermore, some of the "hot spots" the city designated also have relatively modest levels of crime. If we restrict our attention to the highest-crime subsample—true hot spots—we can simulate a smaller and more targeted intervention. The same patterns hold, as the effect of each intervention is roughly proportional to the levels of crime.

Finally, in all of the above, it is property crime, as opposed to violent crime, that appears to be displaced. There is weak evidence of a fall in homicides as a result of the intervention. We should take this subgroup analysis with caution, but the difference between aggregate effects on property and violent crime is statistically significant at conventional levels. And while speculative, displacement of property crime and deterrence of violent crime is consistent with standard economic models of crime: increasing the risk of detection stops criminals from committing motivated crimes in that specific place, but most likely the crime is not deterred and rather committed elsewhere. But expressive violence, once avoided, may be less likely to sustain its motive and be displaced.[9]

A handful of other large-scale studies, powered to detect spillovers, find patterns similar to Bogota. A large-scale trial of intensive policing in another Colombian city, Medellin, draws similar conclusions—small direct effects and no evidence of beneficial spillovers, with wide confidence intervals for aggregate effects including the possibility of adverse spillovers (Collazos et al. 2020). In Mexico, Dell (2015) finds that drug trafficking, a crime with extremely strong and sustained motives, displaces to nearby municipalities in response to increased enforcement. Drunk driving is another criminal behavior that, once underway, may be hard to deter. An experiment with drunk driving checkpoints in India shows that drunks just take other routes (Banerjee et al. 2017).

9. In a recent randomized trial in Bogotá, Nussio and Norza Cespedes (2018) find that information campaigns on the number of arrests at a specific location (i.e. objective information on the probability of apprehension) decrease reports of motivated crimes but not of expressive violence (at treated places).

All of these studies are outside the United States, and so it is possible that something about US cities or policing is different, leading to positive instead of adverse spillovers. We discuss reasons why the spillovers in these smaller-sample US studies are imprecise and varied, and why it may be premature to conclude that place-based interventions do not displace crime.

As more interventions go to scale, we illustrate how design-based approaches, weighting, and RI can reduce bias and improve standard errors. Economists tend to impose a fair degree of structure on spillovers. Where the nature of spillovers is unknown, however, a more flexible design-based approach is more appropriate (Gerber and Green 2012; Aronow and Samii 2013; Vazquez-Bare 2017).

These identification problems and solutions are applicable to a variety of issues beyond crime. Many urban programs are both place-based and vulnerable to spillovers. This includes efforts to improve traffic flow, beautify blighted streets and properties, foster community mobilization, and rezone land use. The same challenges could arise with experiments in social and family networks (Abadie et al. 2016; Vazquez-Bare 2017). Experiments in dense interrelated networks present a textbook case of where design-based approaches and RI needs to enter the econometric program evaluation toolkit.

This paper is organized in six sections including this introduction. Section 2 describes the interventions, Section 3 presents the experimental sample and design, Section 4 presents the data, Section 5 shows the results, and we discuss and conclude in Section 6.

2. Setting and Interventions

2.1. Bogotá

Bogotá is a middle-income city of roughly 8 million people, where income per capita is about a third of the United States. The nature of Bogotá's crime varies, from pickpocketing and cell phone theft in busy commercial areas, to burglary of businesses and homes, to drug sales and any resulting violence. Most violent crimes appear to be expressive rather than instrumental, as the Mayor's office estimates that 81% of homicides in 2015 resulted from fights, whereas 12% were contract killings and 5% were violent robberies. Homicide levels are comparable to a large US city, at 15.6 per 100,000 people in 2016. Finally, most offenders are individual young people. There are some semi-organized youth gangs, and some organized crime, but they do not seem to be responsible for the vast majority of the street crime or violence.

Like many cities, crime in Bogotá is also concentrated. From 2012 to 2015, just 2% of the city's 136,984 street segments accounted for all murders as well as a quarter of all other reported crimes. These higher-crime streets are widely dispersed. They include wealthy areas where criminals come to mug pedestrians, burgle homes, or steal expensive cars, as well as more barren industrial areas with little traffic, where it is easier to sell drugs or steal. We review crime data in more detail below.

Bogotá has moderate to low levels of police compared to large US and Latin American cities. Bogotá has about 18,000 police officers in operational activities, including about 6,200 patrol agents. We estimate about 239 police per 10,000 people, compared to 350 in Colombia nationally, 413 in New York, 444 in Chicago, 611 in Washington, or 257 in Los Angeles.[10]

Police patrols are generally well-regarded. We discreetly observed police patrols and qualitatively interviewed residents on 100 of the treated streets, as described below. The police is not without problems, but patrol officers are generally regarded as competent and non-corrupt. If anything, residents complained that officers were not present often enough.

2.2. Interventions

In January 2016, a new mayor came to power. The first item of his election platform was to identify the 750 highest-crime streets in the city and, within the first 100 days, target them with greater city services, especially police patrols and municipal services. We can view both interventions as an intensification of normal services. Since no new funds or personnel were added, this was a randomized reallocation of services.[11]

2.2.1. Intensive Policing. Intensive policing began in February 2016 and ended in October—a total of 8 months.[12] It generally meant almost doubling normal police patrol time on treated streets. Commanders told patrols to visit treated hot spots about six additional times per day for about 15 minutes each time, mostly during the day unless near a bar. Commanders also instructed patrols to continue their normal duties in treated segments: running criminal record checks; stopping, questioning, and frisking suspicious people; door-to-door visits to the community; conducting arrests or drug seizures; and so forth.[13]

10. Data for Colombia were reported by the Secretariat of Security of Bogota, data for the United States are from the Department of Justice Statistics, and other data are from the United Nations Office on Drugs and Crime.

11. When it comes to the police, the Mayor's office can influence tactics, force allocations, and equipment, but has little say in total force size. City police forces in Colombia are a branch of the National Police and report to the Minister of Defense. But the city has the power of the purse, as it pays for police equipment. The Colombian Constitution also calls on police to comply with Mayors' requests and policies. Changes in force levels are much more expensive, however, and the national government rejected the Mayor's request to increase the number of police. Thus, the Mayor's office focused on increasing police efficiency.

12. The government, however, did not publicize the eligible high-crime streets, the existence of an experimental design, or which specific streets were being targeted. The Mayor's office initially planned to run this intensive policing intervention for at least 4–6 months. They extended the intervention in part to permit the research team enough time to fund and conduct a survey of citizens.

13. The only exception was in three streets known as "The Bronx". Early in our intervention period, the police and city invaded and cleared the three streets. This was a much more intensive, one-time intervention. Two of the three streets happened to be assigned to treatment and one had been assigned to the control group. Police cleared the streets and the city demolished the buildings. In this extreme case, it is obvious that more policing can reduce crime.

As we will see below, in Section 5, our main measure of policing is average patrol minutes per day on each segment, measured by global positioning system (GPS) data. We estimate control hot spots received 92 minutes of patrolling time per day, on average. Treated hot spots received an extra 77 minutes, an 84% increase (a little lower than the targeted 90 minute increase). By comparison, non-hot-spot street segments received an average of 33 minutes of patrolling time per day.

The quadrant (*cuadrante*) is the basic patrolling unit, equivalent to a police beat in the United States. Bogotá has 19 urban police stations. Stations are divided into *Comando de Atención Inmediata* (CAIs), a small local police base that coordinates patrol agents and takes civilian calls. Each CAI has about ten quadrants. There are 1,051 quadrants in the city, each with 130 street segments on average.

Quadrants have six assigned patrol officers. They patrol in pairs, on motorbike and foot, in three shifts of 8 hours each. In practice, patrols are expected to move about throughout their shift, by motorbike. They may patrol a street on motorbike or dismount to speak to shopkeepers, passersby, and suspicious people. Patrols carry a hand-held computer that allows them to check a person's identification number for outstanding warrants. Patrols have daily quotas. They are expected to regularly stop and frisk suspicious people, and will seize illegal weapons (usually knives) and contraband. Patrols tend to focus on young men.

The hand-held computer also contains a GPS chip that records the patrol's location roughly every 30 seconds (when operational). The city piloted and introduced the system in late 2015, under the previous mayor. The new system lets station commanders view patrol positions in real time and get regular performance statistics. Thus, the study period is a period of increased monitoring and measurement of patrol activity.[14]

2.2.2. Municipal Services. Services include trash collection, tree pruning, graffiti clean-up, and streetlight maintenance. The agencies report to the Mayor's office, but the Mayor's power is limited by contracts and difficulties in monitoring and enforcing instructions. One city office coordinates street light maintenance and a second office is in charge of all clean-up activities. Both offices contract private companies to service the streets. Contractors were expected to perform their usual duties, but the Mayor's office gave contractors lists of segments where they were asked to assess issues and deliver the appropriate services. The municipal services intervention began in April and they were instructed to continue until the end of the intensive policing intervention—a total of 6 months.

14. Naturally, the devices that track patrol locations every 30 seconds periodically malfunction, and occasionally the system has an outage. Thus any estimate of minutes is probably an underestimate, one that is unlikely to be correlated with treatment. Furthermore, note that before the intervention, 1–2 weeks of GPS data suggested that experimental sample of streets received at least 38 minutes of patrol time per day. It is doubtful that actual time rose from 38 to 92 minutes. Rather, the 38 minutes is probably an understatement of average patrolling time per street, as there were fewer patrols with GPS devices patrolling city streets. The police did not have data on pre-intervention patrol times, since the GPS devices were piloted November 2015 through January 2016. See Online Appendix A.1.

2.2.3. Impact on State Presence on Control and Spillover Streets. Our best assessment is that the increase in patrol time and city services on treated hot spots did not take a material amount of time away from control hot spots or normal streets, for two reasons.

First, treated hot spots are less than 1% of all 136,984 city streets, so increased attention to treated streets likely has a minor effect on attention to control and non-experimental streets. In the case of policing, for instance, there are 130 segments in the average quadrant, and so the 77-minute increase in patrol time on just 2 segments means roughly 1 minute less time for all other segments, on average. Indeed, the city government mandated that no more than two segments per quadrant could be assigned to treatment exactly for this reason—to avoid major reductions in police time on other streets.

Second, the introduction of the patrol geo-location was part of a policy to increase the efficiency and time on the street of patrols in general. Our best assessment is that all segments (including control hot spots) received at least 10–20% more patrol time than in the pre-intervention period. Online Appendix A.1 displays patrolling minutes before and during the intervention. Pre-treatment data are spotty and incomplete, as the city was piloting the GPS system for the first time. Thus, we do not have reliable estimates of the pre-treatment patrol levels. Nonetheless, all indications point to a substantial increase in patrols on every segment, with patrol times simply rising more substantially in treated hot spots.

In any event, our experiment can be viewed as identifying the effect of the differential in average patrol time or municipal services, irrespective of the source.

2.2.4. Comparison to Other Policing Studies. The Bogotá intervention is broadly similar in style and intensity to several US hot spots policing interventions. The most comparable interventions intensify patrol time but maintain normal duties, such as: a Minneapolis study that raised patrol time to 3 hours per day on 55 hot spots (Sherman and Weisburd 1995); a Jacksonville study where officers surveilled 78 hot spots for an additional 1–2 hours per day (Taylor, Koper, and Woods 2011); and a Sacramento study that added 15-minute police patrols (Telep, Mitchell, and Weisburd 2014).[15] Furthermore, the 8-month period is close to the average duration of other hot spots policing studies.[16] This is also broadly comparable to a Medellín hot spots policing intervention that preceded the Bogotá study, where 384 hot spots were treated with 50–70 more minutes patrols over 6 months in 2015 (Collazos et al. 2020).[17]

15. One difference is that our study focuses on direct effects and spillovers during the course of the intervention, and we look at spillovers over a larger-than-usual range of 250 meters. To the extent that spillovers are highest during the active phase of the interventions, or displace over larger spatial areas, there may be mechanical reasons for us to observe higher rates of displacement than studies that focus on a 2-block radius or post-intervention spillovers. Alternatively, past spillover bounds may have been too small and underestimated spillovers.

16. Braga et al. (2019) document the time duration of all 65 hot spots policing interventions in their systematic review. Excluding our own study, and 4 outliers (3 of them focusing on a few intervention days and 1 on a 9-year observational analysis), the remaining 60 studies have an average duration of 9.5 months.

17. This Medellín study does not observe direct treatment effects on both property and violent crimes, although they do find evidence of a decrease in a particular form of crime: car thefts. They also find a

The Bogotá intervention is distinct in targeting and duration from several "limiting cases" of extremely high or low changes in police attention. Two natural experiments evaluated round-the-clock police in strategic areas of London and Buenos Aires (Draca, Machin, and Witt 2011; Di Tella and Schargrodsky 2004). Blanes i Vidal and Mastrobuoni (2017), in another natural experiment, evaluate the effect of an added 10 minutes of patrolling per day in the 200 meter areas around the site of the prior week's burglaries. Otherwise, policing studies tend to examine a change in policing style rather than intensity, and so are not comparable.[18]

3. Experimental Sample and Design

3.1. Selecting the Experimental Sample

Figure 1 maps Bogotá's 136,984 street segments and indicates the 1,919 segments in our experimental sample. To create this sample, the city started with the 2% highest-crime segments, using an index of reported crimes from geo-coded official statistics, between January 2012 and September 2015.[19] The city then asked each station's commanders and staff to use their knowledge (such as crime calls, or observed street disorder) to verify the segments, because (i) most petty crime is unreported, and (ii) crimes could be geo-located to the wrong street. The police eliminated about a third of these segments, adding others in their stead, leaving 1,919 segments that account for 21% of the city's reported crimes.[20] To examine spatial spillovers, we also analyze the 77,848 street segments that lie within 250 meters of the 1,919 segments in the experimental sample. At this scale, we were ex ante powered to detect direct effects of 0.15 standard deviations and spillovers as small as 0.02 standard deviations (with 80% power, see Online Appendix B).

decrease in car thefts in places nearby targeted segments. The context has some differences as well. For instance, while Medellín has about 60% more police per capita than Bogotá, the city also has highly organized criminal gang structures throughout the city, and police in these low- and middle-income neighborhoods may not be effective in deterring gang-associated crime because of the local power and influence of these groups.

18. Some hot spots interventions take a "zero tolerance" approach, enforcing the most minor infractions; others focus on "problem-oriented policing", where officers try to proactively address problems identified jointly with communities; and still others place license plate readers on street corners, or crack down on drug corners and houses. There are also studies of the quality of police response, such as Blanes i Vidal and Kirchmaier (2018).

19. We constructed a geo-fence of 40 m around each segment and assigned a reported crime to that segment whenever it fell within its geo-fence. Online Appendix A.1 reports further details. A calculation error meant that 608 segments outside the top 2% were included in this initial sample. These were generally high crime segments, as 90% of those streets were above the 95th percentile of baseline crime, and all were above the 75th percentile. In retrospect, this error proved useful since it gave us more variation in baseline crime levels, which we use to study treatment heterogeneity.

20. Homicides are recorded by police. For any other crime to be included in the database, victims had to travel to 1 of 19 police stations, file a formal report, and include relevant details such as location. Our endline survey (discussed below) suggests that official statistics record only about a fifth of all crimes.

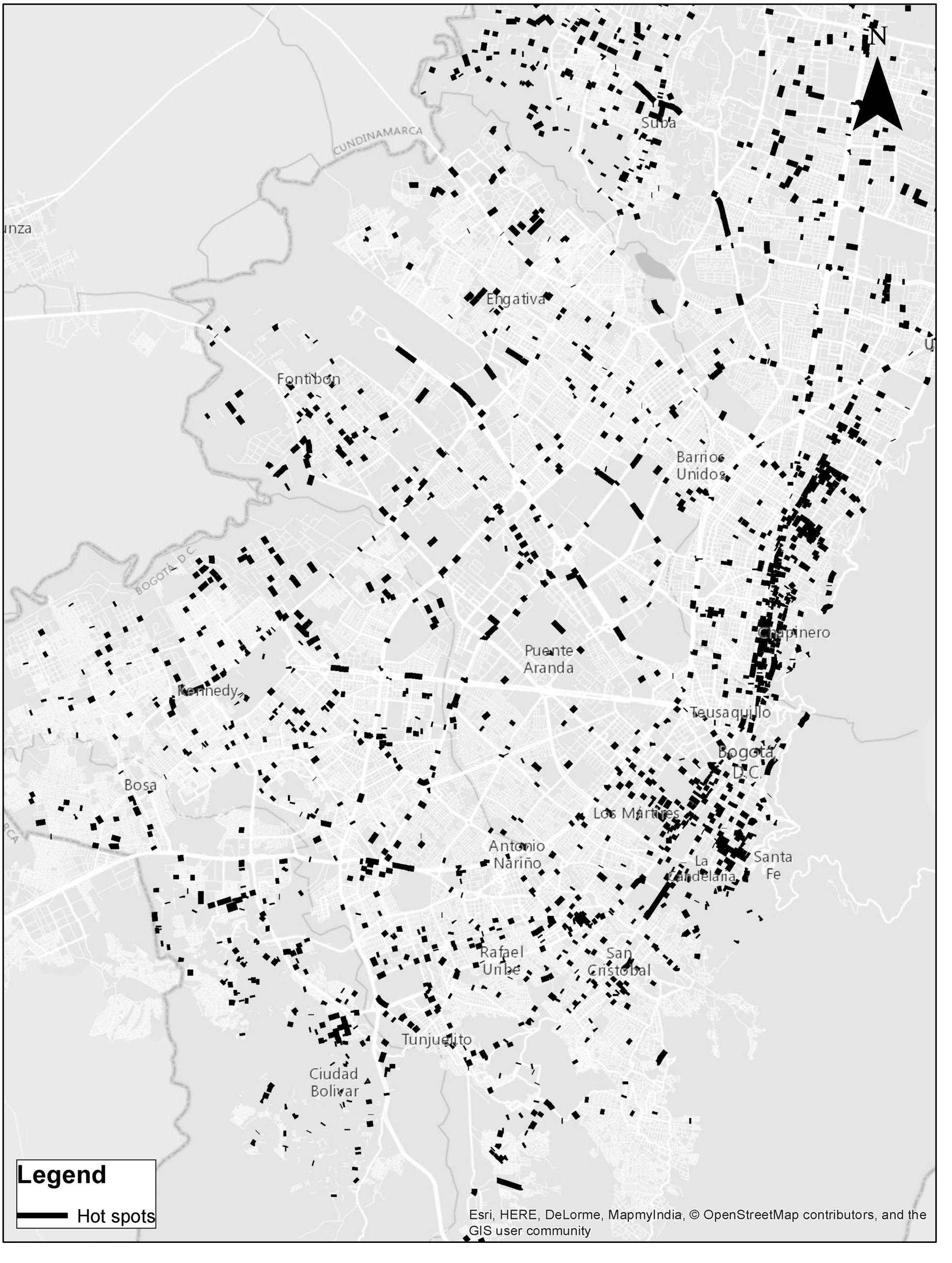

FIGURE 1. Map of experimental sample. Experimental street segments, in black, are the 1,919 streets included in our experimental sample.

On average, the 1,919 hot spots have about 5 times as many reported crimes as those in the non-experimental sample. Only some of these segments, however, correspond to what the US literature calls a hot spot, in the sense of extreme levels of crime. An advantage of our larger and more varied experimental sample, however, is that we can estimate the effect of increased state presence on crime in a mostly normal set of streets. Moreover, we will still be able to simulate a more targeted hot spots intervention by looking at impacts on the subsample of highest crime streets—our prespecified heterogeneity analysis.

3.2. *Experimental Design and Estimation*

We used a factorial design and randomized the 1,919 hot spots to 1 of 4 experimental conditions: intensive policing; more municipal services; both; or neither. To identify direct treatment effects, we compare average crime levels in treated hot spots to control hot spots. To estimate spillover effects, we draw 250-meter rings around treated and control hot spots and compare crime levels on these nearby potential spillover streets. Our analysis follows a pre-analysis plan registered prior to the intervention, with minor adjustments.[21]

We estimate treatment effects with the weighted least squares regression:

$$Y_{sqp} = \beta_1 P_{sqp} + \beta_2 M_{sqp} + \lambda_1 S^P_{sqp} + \lambda_2 S^M_{sqp} + \gamma_p + \delta E_{sqp} + \Theta X_{sqp} + \varepsilon_{sqp} \tag{1}$$

where Y is the outcome in segment s, quadrant q, and police station p; P is an indicator for assignment to intensive policing; M is an indicator for assignment to municipal services; S^P and S^M are indicators for segments within 250 m of a treated segment; γ is a vector of police station fixed effects (our randomization strata); and X is a matrix of pre-specified baseline control variables. The sample includes the 1,919 hot spots

21. The plan is available at available at https://www.socialscienceregistry.org/trials/1156. This regression departs from the pre-analysis plan in three minor ways. First, the plan indicated that we would focus on pairwise comparisons of each intervention *separately*, dropping from the regression any segments with a zero probability of assignment to any of the conditions. That approach generates similar results but, in retrospect, is problematic. Most importantly, we realized that ignoring municipal services treatment and doing a pairwise comparison of intensive policing treatment and control streets produces biased results, since assignment to municipal services is slightly imbalanced across intensive policing experimental conditions (see Table 1). Hence, we estimate the effects of both interventions jointly.
Second, based on feedback and referee requests, our main analysis focuses on the full sample of direct and spillover effects, using administrative crime data on nearly 80,000 streets that were directly treated or within 250 meters. This analysis has the greatest statistical power. Our pre-analysis plan instead specified a broader security index for the dependent variable, including both survey and administrative data, on a much smaller sample. Particularly given the modest direct effects, and the increased importance of estimating spillovers, the large-sample results with more precise spillovers are more relevant. We report the smaller-sample security index results as a secondary analysis in the paper. It does not materially change our conclusions.
Third, we report the interaction between the two treatments as a secondary analysis, rather than a primary analysis, given that there are just 75 street segments that received both treatments (fewer than we expected prior to implementation). We estimate interaction effects by including $\beta_3(P \times M)_{sqp} + \lambda_3(S^P \times S^M)_{sqp}$ in equation (1).

plus the 77,848 spillover segments within 250 m of one of the hot spots. The indicator variable E_{sqp} takes the value of one for the experimental sample of 1,919.[22] Weights are the IPWs of assignment to each experimental condition, as discussed further below.

This regression effectively compares means across the different experimental conditions, conditional on whether segments are in the hot spot pool or the non-experimental pool: treatment versus control hot spots, and segments ringing treated hotspots to streets ringing control ones (whether experimental or not). Below we discuss how we address the issue of overlapping rings. In addition, some crime is undoubtedly displaced in non-Euclidean ways (e.g. to distant segments where crime benefits are high and detection risk is low).[23] We will fail to estimate such spillovers.

3.2.1. Design-based Approach. In the presence of spillovers, there are two issues with the above approach. One is that we did not know the structure of spillovers ex ante, including the appropriate ring radius. The second is that, in the presence of spillovers, there is potential for interference between experimental units when control hot spots fall within a treated hot spot's spillover ring. In addition, potential spillover streets may fall within the ring of both a control and treatment segment.

We designed randomization and estimation procedures to minimize both issues. First, we used a flexible approach to choose the appropriate spillover radius.[24] We allowed for the possibility of no spillovers, as well as more distant spillovers into an additional 250–500-meter ring, and prespecified a decision rule for determining the appropriate spillover ring or rings. Online Appendix A.2 provides details on how the 250-meter ring was determined. For robustness, we also estimate spillovers using a continuous rate of decay, but we prefer the rings estimate because it is non-parametric, because it also allows us to say something about aggregate effects of the intervention, and because it allows us to use IPWs to correct for the ex ante identification concern discussed below.

Second, we used a two-stage randomization to minimize the overlap in rings around treatment and control hot spots. For intensive policing, we first blocked our sample by the 19 police stations, then randomized segments to intensive policing in two stages: first assigning quadrants to treatment or control, then assigning segments within treatment quadrants. We assigned no more than two segments per quadrant to intensive policing. This procedure assigned 756 segments to intensive policing and 1,163 to control. For the municipal services intervention, we identified the subset of the

22. This pooled sample constrains the estimated λ coefficients to be the same for all spillover segments, regardless of whether they are a control hot spot or a spillover street in the nonexperimental sample.

23. Ferraz, Monteiro, and Ottoni (2016) find evidence of non-spatial spillovers in Rio de Janeiro's favela pacification. Such non-spatial spillovers would lead us to overstate direct effects and understate spillovers.

24. An alternative would be to try to model spillovers structurally, as Adda, McConnell, and Rasul (2014) do in their study of the depenalization of small quantities of cannabis possession in a London neighborhood. In our instance, the nature, direction, and range of spillovers is unknown, and the range of potential crimes and nuisances affected by the treatment are many and diverse. Hence we took a more flexible and non-structural approach.

1,919 hot spots with visible disorder.[25] We blocked on police station and the previous intensive policing assignment, and assigned 201 segments (14% of eligible segments) to municipal services.[26]

Despite the above design, the scale and density of the experiment cannot eliminate the overlap in rings. This leads to two potential identification concerns and econometric corrections for each.

3.2.2. Confounding with Spillover Assignment, and the IPW Correction. The first concern is that the street segments within overlapping rings tend to be in higher-crime or denser parts of the city, and so are unlike other streets. In most randomizations, these streets will be assigned to the spillover condition in most randomizations. This creates potential confounding if there is a correlation between unobserved characteristics of this type of streets and our outcome (crime). This could mechanically lead us to conclude that there are adverse spillovers. Controlling for baseline characteristics and crime histories reduces this bias but does not eliminate it.

We can control for this potential confounding with IPWs, whereby we weight segments by the inverse of the probability of their assignment to that condition.[27] In effect, unusual segments with a low probability of assignment to a pure control condition receive more weight when they are assigned to the control condition. These weights are a transparent and systematic way to transform the assignment process such that all segments have the same probability of being exposed to spillovers. Online Appendix A.3 elaborates.[28]

3.2.3. Fuzzy Clustering and RI. The fact that segments in denser and higher crime parts of the city are assigned to the spillover condition in most randomizations also creates patterns of "fuzzy clustering" (Abadie et al. 2016). These clusters are difficult to model because they have to do with distance from other experimental segments rather

25. We sent enumerators to take five photographs and rate segments for the presence of disorder. They looked for graffiti, garbage, and run-down buildings. A limitation is that we measured disorder after 2 months of policing treatment. We had no reason to expect the treatment to affect physical disorder, and there is no statistically significant difference between experimental and non-experimental segments. Of the 1,534 segments they were able to safely visit, 70% had at least one maintenance issue. We made these, plus the 385 segments they could not visit safely, eligible for municipal services assignment.

26. These 201 were the first "batch" to be treated. We also randomized a second batch of 214 segments for later treatment should the city decide to expand services. Two months into treatment of the first batch, however, our analysis of compliance records and visual inspection of segments suggested that continued municipal services were needed to maintain order in the first batch, and so the city did not give contractors the list of segments in the second batch. Thus, the second batch remains in our control group.

27. Each segment's ex ante probability of exposure spillovers can be estimated with high precision by simulating the randomization procedure a large number of times. Such IPWs have a long history in survey sampling and have become common in the analysis of randomized trials with varying probabilities of assignment (Horvitz and Thompson 1952; Gerber and Green 2012). Online Appendix A.3 describes and maps IPWs in our sample.

28. As we will see, once we add baseline controls, the IPW correction makes little difference to our estimates because the regression controls for strata that predict the weights.

than an observed characteristic such as a quadrant. Thus there is no geographic unit with which to calculate clustered standard errors. This fuzzy clustering not only leads to incorrect standard errors, Online Appendix A.4 also shows that it is a source of bias in the estimate itself. This is because cluster size (which is greatest in the densest and highest-crime parts of the city) is correlated with potential outcomes. IPWs will not correct this specific bias, and of course would not adjust standard errors appropriately.

Randomization inference gives precise p-values based on the empirical distribution of all estimated treatment effects that could arise under our design and data under the null hypothesis of no effect for any unit. Randomization inference reassigns treatment randomly thousands of times, each time estimating the treatment effect that could have arisen by chance. What RI allows us to do is to assign a p-value for a given treatment effect by observing where that treatment effect falls in the distribution of all possible estimated effects from the 10,000 randomizations we simulate under the assumption of no effects. We use these RI p-values in place of the conventional standard errors-based p-values whenever we estimate treatment effects in the presence of spillovers. Additionally, the simulations used in the RI procedure provide an estimate of the bias insofar as they reveal the average estimated coefficient when there is no true effect. All of our tables report bias-corrected treatment effects.

3.3. Summary Statistics and Randomization Balance

Table 1 reports summary statistics and balance tests for the experimental sample of hot spots. In October 2016, the police updated all 2012–2016 crime data with more accurate GPS coordinates and additional crime categories, and we report both the original and updated data.[29] The hot spots had up to 82 crimes reported in the previous 4 years using the original data, and up to 461 with the updated data. In both cases, however, the average is much more modest: about five reported crimes per segment, or just one per year.[30] Thus our sample has a variety of moderate to high crime streets. We discuss this variation in more detail in Online Appendix A.5.

Random assignment produced the expected degree of balance along covariates. For the most part, background attributes appear balanced across experimental conditions. There are some minor differences between treatment and control segments (for instance, treated segments are slightly less likely to be in industrial zones or middle income status, and treated quadrants have slightly fewer experimental segments), but overall the imbalance is consistent with chance and is robust to alternative balance tests.[31]

29. Some crimes moved to nearby segments, and the correlation between the old and new data is 0.35 at the segment level and 0.86 at the quadrant level. These corrections were unrelated to this study.

30. Quadrants with at least one segment had an average of 3.5 reported crimes per segment across the whole quadrant, while the average quadrant in the whole city reported 1.5 crimes.

31. To see whether covariate imbalance lies within the expected range, we test the null hypothesis that the covariates do not jointly predict experimental assignment. We use multinomial logistic regression with RI to model the four-category experimental assignments for segments in the experimental sample (treatment, <250 m, 250–500 m and >500 m), or the three-category assignments for streets in the non-experimental sample (<250 m, 250–500 m and >500 m). To obtain exact p-values, we use RI. Using simulated random

TABLE 1. Descriptive statistics for the experimental sample and tests of balance (N = 1,919).

	Summary statistics			Balance test			
				Intensive policing		Municipal services	
Variable	Mean (1)	Std. dev. (2)	Max. (3)	Coeff (4)	p-value (5)	Coeff (6)	p-value (7)
Crimes reported per segment, 2012–2015 (original)	4.53	5.72	82	−0.18	0.68	−0.14	0.89
# of violent crimes	1.88	2.94	56	−0.20	0.37	−0.06	0.89
# of property crimes	2.66	3.97	50	0.03	0.96	−0.08	0.91
Crimes reported per segment, 2012–2015 (updated)	5.18	18.24	461	−0.35	0.86	−0.18	0.91
# of violent crimes	1.40	5.38	78	0.33	0.28	0.33	0.23
# of property crimes	3.78	14.09	407	−0.69	0.61	−0.50	0.81
Patrol minutes per day (11/2015–01/2016)	38.03	70.27	1029	−2.61	0.69	3.02	0.56
Rating of baseline disorder (0–205)	1.18	0.74	5	−0.02	0.48	0.08	0.11
Meters from police station or CAI	551.37	351.46	2805	−25.70	0.24	−1.80	0.79
Zoned for industry/commerce	0.38	0.49	1	−0.10	0.01	0.04	0.23
Zoned for service sector	0.13	0.34	1	0.02	0.31	0.03	0.24
High-income street segment	0.07	0.25	1	0.00	0.91	0.01	0.49
Medium-income street segment	0.55	0.50	1	−0.05	0.11	0.01	0.83
Segments in quadrant	127.21	86.99	672	2.59	0.72	−3.95	0.47
Hot spots in quadrant	3.67	2.68	14	−0.05	0.39	−0.24	0.24

Notes: Columns (1)–(3) display summary statistics for our experimental sample of 1,919 segments, weighted by the probability of being in the observed experimental condition. Columns (4)–(7) report balance tests for treated versus all control units using weighted least squares.

4. Data

When looking at the full sample of hot spots and nearby non-hotspots, our sole available outcome is an index of officially-reported reported crimes. Police shared data on reported crimes and operations 2012–2017, by day, geolocated to each of the 136,984 streets.[32] We use an additive index with equal weights for all crimes. We can subdivide this index into property and violent crimes, although our main measure pools all crimes into one index.

Of course, most crimes and nuisances go unreported. Therefore, we conducted a survey of 24,000 residents of Bogota. In addition to giving us a broader measure of street-level security, survey data allow us to test whether treatment increased crime reports. In October 2016, we surveyed a convenience sample of 10 people per street segment on 2,399 segments—the 1,919 hot spots, plus a representative sample of 480 segments within 250 meters of a hot spots. As a result, in addition to looking at reported crime impacts alone, we can estimate direct effects of treatment on a broader "insecurity index". This index is a z-score that we produce after by adding up, with equal weights, two kinds of measure.

1. *Perceived risk of crime and violence on the segment.* Our survey asked respondents to rate perceived risk on that segment on a four-point scale from "very unsafe" to "very safe" in five situations: general risk during the day; a young woman walking alone after dark; a young man walking alone after dark; talking on a smartphone. We construct an index of perceived risk by taking the average across all respondents in the segment and producing a z-score.
2. *Crime incidence on the segment.* We construct a standardized index of crime that equally weights: (i) a z-score we produce with opinions of the incidence of crime on that segment, as well as personal victimization since the beginning of the year;[33] and, (ii) a z-score we produce with officially-reported crime on the

assignments, we obtain a reference distribution of log-likelihood statistics under the null hypothesis; we then calculate the p-value by locating the actual log-likelihood value within this reference distribution. The p-value is non-significant, as expected, for both the experimental and non-experimental samples: $p = 0.681$ for segments and $p = 0.531$ for non-experimental segments. We draw similar conclusions from tests of treated vs control units >250 m away and between control units <250 m and >250 away.

32. Prior to the intervention, we received the 2012–2015 data on the city's priority crimes: homicides, assaults, robberies, and car and motorbike theft. A total of 77% of the crimes had exact coordinates and the rest had the address, which we geolocated ourselves, with about 71% success (or 93% of all reported crimes). We also received all data on arrests; gun, drugs, and merchandise seizures; and stolen cars and motorbikes recovered. In October 2016, the police provided updated data that corrected for geolocation problems (thus retrospectively changing pre-intervention data). With the new information we also received data on reported cases of burglary, shoplifting, sexual assaults, family violence, threats, extortion, and kidnapping. Some US studies use emergency call data. Initially these were not available, and our pre-analysis plan excluded them. Later, partially complete data became unexpectedly available, and our main results are robust to their inclusion (not shown).

33. The survey measured perceived incidence and personal victimization by walking respondents through a list of 11 criminal activities. After finding out whether any of these activities happened on the street since the beginning of the year, we asked respondents about each crime to establish perceived frequency

segment since the beginning of the intervention. For the final index, we add up both components and produce a z-score.

These survey data illustrate some of the advantages of including the police in street selection. It means that our experimental sample includes some segments with low levels of reported crime, but moderate to high unreported crime and nuisances. According to the survey responses, for instance, 3 in 10 of the people stopped on each of the experimental segments reported a personal experience of crime on that segment in the previous 8 months. This is a relatively high rate of victimization. Perceived risk is 10 of 60 on average, stretching as high as 20 or 30 in the highest-crime streets.[34]

We also learn some important features of crime underreporting, most importantly that official crime reporting does not appear to be correlated with treatment (a common identification concern in policing studies). Our survey asked whether or not people had experienced eleven kinds of crime since the beginning of the year, whether they had attempted to report it, and if they were successful. Online Appendix A.5 reports results. On average, 27% reported a crime, and an additional 9% of people say they attempted to report the crime but were unsuccessful. High-value property crimes like vehicle theft were reported about half the time, robberies and assaults about a quarter of the time. We see no effect of treatment on crime reporting.[35] This suggests that administrative data are suitable for outcome assessment even while the treatment is being delivered.

Finally, we drew on four other sources of data.

1. *Survey data on legitimacy.* The survey also asked residents about their trust in and satisfaction with the police and the Mayor's office. We discuss these secondary outcomes in Online Appendix C.1.
2. *Administrative data on police, municipal services, and streets.* The police shared GPS patrol locations for all 136,984 streets, 2015–2017.[36] For streets

(ranging from "everyday" to "never" on a 0–6 scale), and whether it happened to the respondent him or herself on that segment. We show results for the two individual components in order to give a sense of the absolute impacts and differences between survey and administrative data.

34. Some top 2% streets also have a small number of officially-reported crimes in the pre-intervention period. Also, a small number of the non-2% streets have a sizable number of crimes. Why is this so? To understand the lower-crime "top 2%" streets, recall that in 2016 the police issued a more complete and correct version of their 2012–2015 geo-located crime data. Some "top 2%" segments had some crimes reclassified away from them but remain in the sample nonetheless. Other reasons include the fact that less serious crimes were given less weight in the sample selection, so that a street with one murder was more likely to enter the experimental sample than one with several muggings. Finally, a miscalculation in the sample selection admitted a small number of moderate crime streets into the "top 2%" sample. Furthermore, none of the experimental streets should be considered "low-crime", simply lower reported crime.

35. The survey asked respondents their likelihood of reporting a future crime to the police, on a scale of 0–3. The average response in control segments was 2.0, with a treatment effect [standard error] of 0.016 [0.029] from policing and 0.035 [0.032] from municipal services.

36. Not all hand-held computers were functional at all times, and at times over 2016, the system went offline for a few days to a few weeks, and so we use data only during those periods when the system was generally operational in a given police station—on average 33 of the 37 weeks of the intervention.

assigned to municipal services, agencies shared their diagnosis and compliance. The city also shared administrative data on the baseline variables reported in Table 1.

3. *Survey of street disorder.* To measure levels of street disorder before and after treatment, we sent enumerators to take photographs and rate the presence of graffiti, garbage, and boarded-up buildings on a 0–5 scale.[37]
4. *Qualitative interviews.* We began with informal qualitative interviews with dozens of police officers and citizens about their experiences with the intervention and police tactics in general. We also hired observers to discreetly visit 100 streets in the experimental sample for a day and passively observe police behavior. They also interviewed citizens in each segment about police behavior and attitudes.

5. Results

5.1. Program Implementation and Compliance

We begin with first-stage program impacts, reported in Table 2.

Police Patrol Time and Actions. Police complied with their new orders for the full 8 months. As discussed in Section 2, we have the geolocation of police every 30 seconds most days on most segments. We estimate that treated hot spots received an extra 77 minutes of patrol, an 84% increase over the 92 minutes they patrolled control hot spots (and more than 5 time the patrolling received by non-hot spots). We see no effect of increased policing on arrests or police actions such as drug seizures. This implies any direct effect of the policing comes from deterring or displacing criminals rather than incapacitating them. Incapacitation, of course, would reduce the chance that crimes are displaced.

Municipal Services. The evidence on service delivery compliance is more mixed. After assigning 201 segments to municipal services, city agencies diagnosed each one in March. They identified 123 segments needing clean-up services, and 47 needing lighting improvements. They were instructed to perform these services for 6 months, but qualitatively they seem to have worked most intensively for 3 months, June through August. Tree pruning and graffiti cleaning were one-time treatments; rubbish collection

37. We visited 1,534 of a total of 1,919 scheduled streets in March (3 months before the municipal services intervention began) in order to narrow down the number of eligible experimental segments. We did not collect data in the remaining 385 streets because of security concerns from the enumerators. (Note that there was no association between intensive policing treatment and these security concerns.) As we discuss in Section 3.2.1, 1,459 were eligible for the municipal services interventions and 414 of them were assigned to treatment. Those streets were split into two batches of 201 and 213 streets, respectively, in order to randomize timing, but only the first batch was effectively treated. Then, in order to assess the levels of compliance, we sent enumerators to the 414 streets in the first and second batches in June (1–2 weeks after municipal services started to be delivered) and December (2 months after the end of the intervention). Again, because of security concerns of the enumerators, we visited 409 in June and 410 in December.

TABLE 2. "First-stage" effects of treatment on measures of compliance and effectiveness.

Dependent variable	Control mean (1)	ITT and standard error of assignment to: Intensive policing (2)	(3)	Municipal services (4)	(5)	N (6)
A. Intensive policing measures:						
Daily average patrolling time	92.001	76.571	[4.424]***	−3.333	[4.371]	1,919
Arrests per segment	0.333	−0.053	[.082]	0.026	[.102]	1,919
Drug seizures per segment	0.041	−0.002	[.020]	0.029	[.024]	1,919
Gun seizures per segment	0.009	0.006	[.008]	0.007	[.013]	1,919
B. Municipal services measures						
Eligible for lights intervention	0.349	−0.007	[.048]	−0.139	[.048]***	415
Received	0.000	−0.010	[.020]	0.199	[.026]***	1,919
Eligible for garbage intervention	0.000	0.011	[.025]	0.627	[.032]***	1,919
Received	0.000	0.015	[.026]	0.382	[.033]***	1,919
3 months post-treatment (December)						
Graffiti on segment	0.624	0.019	[.053]	0.059	[.047]	409
Garbage on segment	0.245	0.021	[.051]	0.002	[.043]	409
Visibly broken street light on block	0.029	0.022	[.016]	−0.015	[.017]	399

Notes: This table reports intent to treat (ITT) estimates of the effects of the two interventions, via a weighted least squares regression of each outcome on treatment indicators, police station (block) fixed effects, and baseline covariates (see equation (1), where we have constrained the coefficient on the interaction term to be zero and ignored spillovers). The regression ignores spillover effects. Daily average patrolling times exclude quadrant-days without data. Standard errors are clustered using the following rules: (i) for all treated segments except with cluster size 2, each segment is a cluster; (ii) for all other untreated segments, each segment gets its own cluster identifier; (iii) for entirely untreated quadrants, they form a cluster; and (iv) for quadrants with exactly two units assigned to treatment, those units form a cluster. The proportion of people reporting increased state presence comes from our citizen survey, the enumerator assessments were collected by the research team, and the remainder of the outcomes come from police administrative data. * significant at the 10% level, ** significant at the 5% level, and *** significant at the 1% level.

was expected to be semi-regular. Based on city data, 74 of the 123 streets (60%) were cleaned up, and in 41 of the 47 streets (87%) they repaired broken lights and replaced poor lights with better ones. No graffiti was cleaned-up.

We also visited segments in daytime in June and December 2016 to photograph and rate the streets. The before and after photos generally display relatively tidy streets and before-after differences are imperceptible. It is possible that lights repairs were more evident, but it was unsafe to visit segments at night. We see no effect of treatment in Table 2. One possibility is that the extensive margin is the wrong margin to evaluate, and another is that the disorder in cleaned up segments could have re-accumulated over days or weeks.

5.2. *Direct and Spillover Effects on Officially-Reported Crime*

Table 3 reports estimates of direct and spillover effects, as well as a rough measure of aggregate effects.[38] On directly treated hot spots, both intensive policing and municipal services reduce officially-reported crimes slightly, but the coefficients are not statistically significant. Control hot spots report 0.743 crimes on average over 8 months

38. Online Appendix C.2 presents results with an interaction term between the two treatments.

TABLE 3. Estimated direct, spillover, and aggregate impacts of the interventions, accounting for spillovers within <250 m, pooling the experimental and non-experimental samples.

		Dependent variable: Reported crime per segment			
Impacts of treatment	Control mean (1)	Coeff. (2)	RI p-value (3)	# segments (4)	Total = (2) × (4) (5)
A. Direct treatment effect					
Intensive policing	0.743	−0.098	*0.386*	756	−74.4
Municipal services		−0.133	*0.185*	201	−26.8
Subtotal					−101.3
B. Spillover effect					
Intensive policing	0.283	0.017	*0.112*	52095	871.8
Municipal services		0.002	*0.645*	21286	42.4
Subtotal					914.12
Net increase in crime					812.9
				95% CI	(−648, 2192)
				90% CI	(−317 , 1986)

Notes: Column 1 presents the average number of crimes reported to the police in control segments. Column 2 displays the bias-adjusted treatment effect while column 3 displays RI p-values. Column 4 displays the number of units in each group. Column 5 displays the product of the bias-adjusted treatment effect and the number of units in each group. The confidence interval on the bottom of the table is constructed using RI.

of the intervention. Intensive policing reduced this by −0.098, a 13% improvement.[39] Municipal services reduced this outcome by −0.133 crimes, an 18% improvement.[40] While these percentage changes are substantial, if we multiply each estimated treatment effect by the number of treated streets, however, the total amount of crime directly deterred is small. Reallocating police and municipal services to higher-crime streets directly deterred just 101.6 crimes over 8 months.

Panel B of Table 3 estimates spillovers onto segments within 250 meters of a treated hot spot. For intensive policing, all spillover coefficients are positive (i.e. spillover segments experience more crimes when they are near a treated segment). They are not statistically significant, but the p-value on the adverse spillovers from intensive policing is 0.112. There are so many segments (from the non-hot spot sample in particular) that these small spillover coefficients add up to an increase in 914 reported crimes.

39. We can see that these results are not driven by a decrease in patrolling time in control streets by estimating the marginal effect of one additional hour of patrolling time. The marginal effect of an additional hour of patrols is a decrease of about 0.1 crimes. This is similar to the average effect of −0.098, as the average treatment street received 76 minutes of extra patrolling time. Online Appendix C.3 has these IV estimates.

40. Our qualitative work and compliance data hinted that the lighting intervention may have been more compliant, effective, and persistent than the street clean-up. But the data do not support this conclusion. Both lighting and cleanup services appear to have been important. For example, we see no evidence that municipal services treatment effects were concentrated in the segments diagnosed as needing improved lights. Furthermore, we do not see larger treatment effects at nighttime (tables not shown).

We total these direct and spillover effects at the base of the table, and calculate RI confidence intervals. These aggregate estimates suggest the treatments increased crimes by about 813, a 3% increase relative to the 26,445 total reported crimes during that period in those segments. We have to take this increase with caution, as the estimates are not statistically significant at the 10% level. Also, this is not a true general equilibrium estimate; rather it is simply a means to place some policy-relevant magnitudes on total crime impacts.

While we cannot exclude zero spillovers, it is incorrect to view these aggregate estimates as imprecise. First, the 95% confidence interval rules out a 2.5% aggregate decrease, and the 90% confidence interval rules out a 1.2% decrease. Second, recall that we were ex ante powered to detect spillovers of roughly 0.02 standard deviations—an order of magnitude more power than prior studies (Online Appendix B). Most of these spillover coefficients are just below that threshold (table not shown). This is one reason why the confidence intervals on spillovers and aggregate effects include zero.

Robustness These results are largely robust to alternative estimation approaches, as seen in Table 4.[41] Direct and spillover estimates generally maintain their direction and general magnitude when we: (i) omit covariate adjustments; (ii) use a count of the number of treated streets within 250 meters for the spillover measure, instead of an indicator; (iii) estimate spillovers using an exponential rate of decay rather than our 250 meter rings; and (iv) use a linear instead of an exponential decay function. The only exception is the spillover count measure, where the direction of spillovers changes, but is not statistically significant. Note that the coefficients on the two decay functions represent the expected increase in crimes as a segment moves a standard deviation closer to a treated segment, so are not directly comparable in magnitude to the main specification.

5.3. Additional Analysis

5.3.1. Survey-based Outcomes. Direct program impacts using survey-outcomes are broadly consistent with the results using officially-reported crimes. Table 5 reports the direct effects of the intervention in hot spots, and can be interpreted as the average standard deviation changes in the outcome, unless specified otherwise. Each intervention is associated with a 0.13–0.14 standard deviation security improvement on directly treated streets, not statistically significant at conventional levels. The

41. To see the effect of other design and estimation choices, in Online Appendix C.4, we estimate “naïve” treatment effects ignoring IPWs and RI, but including covariates. Note that these are not robustness checks, however, because they are not causally identified. Direct treatment effects are slightly smaller than in Table 5, but the patterns remain similar. The estimated spillover effects in this “naïve” case, however, are much larger and highly statistically significant compared to Table 5. Hence, failing to account for interference between units and clustering of treatment conditions would have led us to severely exaggerate the degree to which the interventions push crime elsewhere.

TABLE 4. Estimated direct and spillover effects using alternative covariate adjustments and methods of spillover estimation, with RI *p*-values (N = 79,767).

		Dependent variable: Reported crime per segment			
		Direct effect		Spillover effect	
Specification	Control mean (1)	Policing (2)	Services (3)	Policing (5)	Services (6)
Main specification	0.283	– 0.098	– 0.133	0.017	0.002
		0.386	*0.185*	*0.112*	*0.645*
Drop covariates	0.283	– 0.112	– 0.117	0.015	0.021
		0.337	*0.326*	*0.200*	*0.905*
Spillover count measure	0.283	– 0.096	– 0.107	– 0.019	– 0.007
		0.393	*0.262*	*0.147*	*0.257*
Spillover exponential decay	0.283	– 0.113	– 0.138	0.003	0.020
		0.314	*0.178*	*0.500*	*0.322*
Spillover linear decay	0.283	– 0.107	– 0.128	0.011	0.021
		0.348	*0.197*	*0.208*	*0.350*

Notes: Each line reports an alternate specification model. Randomization inference *p*-values are in italics.

TABLE 5. Estimated direct program impacts on survey-based measures of security in the experimental sample, with RI *p*-values (N = 1,916).

		ITT	
Dependent variable	Control Mean (1)	Policing (2)	Services (3)
Insecurity index, z-score (+ more insecure)	– 0.003	– 0.130	– 0.139
		0.269	*0.223*
Perceived risk index, z-score (+ riskier)	0.049	– 0.139	– 0.111
		0.172	*0.275*
Crime index, z-score (+ more crime)	– 0.054	– 0.078	– 0.121
		0.508	*0.313*
Perceived & actual incidence of crime, z-score (survey)	0.059	– 0.060	– 0.129
		0.709	*0.220*
# crimes reported to police on surveyed segment (admin)	0.743	– 0.092	– 0.025
		0.416	*0.764*

Notes: This table reports intent to treat (ITT) estimates of the effects of the two interventions, via a ordinary least squares regression of each outcome on treatment indicators with, police station (block) fixed effects, and baseline covariates (see equation (1)). The equations are estimated and reported with and without the interaction between treatments. Columns 2–5 report the RI-adjusted estimates of the interventions on the listed measures. Note that the police and city invaded and cleared three streets as part of a special, one-time intervention, and hence we have no survey data from these areas. Two of the three streets were assigned to treatment and one to the control group. We exclude all three from the main analysis. Randomization inference *p*-values are in italics.

coefficients on the component indexes—perceived risk and actual crime—are generally similar.[42]

42. After completion of the experiment, we also received calls-for-service data from police. We are concerned that direct treatment would affect calls for service, especially the more frequent presence of

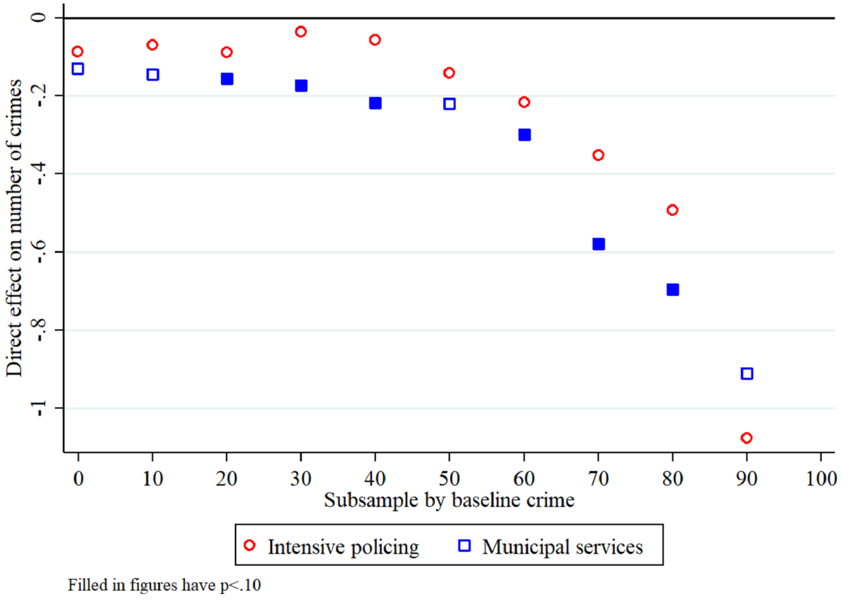

FIGURE 2. Direct program impacts in *n*th percentile highest-crime street segments. We estimate equation (1) ten times, each time interacting each treatment indicator with an indicator for whether a segment is above the *n*th percentile of baseline crime levels among our experimental sample of segments, for $n = 0, 10, 20, \ldots, 90$. The coefficients on the treatment indicators indicate the effect on the higher crime segments above that percentile (hence, the right side of the figure represents the highest crime "hot spots").

Heterogeneity by Level of Initial Crime. We pre-specified one major form of heterogeneity analysis, by baseline crime. This helps us compare our results to the US hot spot policing literature that tend to focus on higher-crime hot spots (because they have typically been smaller in scale). Figure 2 reports estimated direct treatment effects on the number of reported crimes for the $n\%$ highest-crime hot spots. Specifically, we estimate equation (1) nine additional times. Each time, we interact the treatment indicator with an indicator for whether a segment is above the *n*th percentile of baseline crime levels among our experimental sample of hot spots, for $n = 0, 10, 20, \ldots, 90$. The figure plots the coefficients on these higher crime streets, with the $n\%$ "hottest spots" on the right. Broadly speaking, the results are consistent with direct effects being roughly proportional to the levels of crime. For instance, at the 90th percentile of baseline crime, there are just over four crimes reported per segment. The treatment effect of roughly -1 crime is a decline equal to 25% of the the average crime totals at control experimental streets with similar baseline crime.

police. Hence, we omit these data from the final analysis. The average hot spot received 17.5 calls over the 8 months. Intensive policing alone reduced this by 3.9 calls ($p = 0.30$), municipal services increased calls by 1.7 ($p = 0.44$), and the cumulative effect of both interventions was to reduce calls by 2.3 ($p = 0.71$).

5.3.3. Heterogeneity over Time. A natural question is whether police or the criminal population adapt to the increased police patrols over 8 months. If potential offenders respond strategically, then we would expect program effects to be stronger near the start. Alternatively, officers may take time to build up place-specific knowledge and relationships, helping to sustain or increase program impacts over time. Figure 3 reports treatment effect estimates for cumulative reported crimes in the first w weeks of the intervention. We estimate equation (1) ten times, each time adding four more weeks until the end of the intervention. Figures 3a and 3b display the coefficients for direct effects from the intensive policing and the municipal services treatments, respectively. We see no statistically significant difference over time, though the point estimates trend downwards. Thus we do not see much evidence that the intervention is becoming less effective over time due to adaptation of potential offenders. Figures 3c and 3d display estimated spillovers. Consistent with the results on direct effects, we see no statistically significant difference over time, though the estimates are increasing over time, especially for the intensive policing treatment. This is broadly consistent with the larger direct effects. Speculatively, these results are consistent with patrols becoming more effective over time, perhaps because they take time to build up knowledge and relationships (though the crimes they counter appear to still spill over into nearby streets).

5.3.4. Heterogeneity by Type of Crime. Police tend to prioritize violent crimes such as assault, rape, and murder over property crimes such as burglary or theft. Table 6 takes the aggregate impacts on officially-reported crime from Table 3 and disaggregates these total effects into violent and property crimes (Online Appendix C.5 reports the full tables).

We need to take these non-prespecified subgroup analyses with caution. Nonetheless, it is interesting to note that the interventions have opposing effects on property and violent crime, however, imprecise. Property crimes rose by 990 in aggregate (6% relative to the total number of property crimes in potentially affected segments). This increase is statistically significant at the 10% level, but this p-value would be considerably higher if adjusted for multiple comparisons.

Meanwhile, our best estimate is that aggregate violent crimes fell by 177 crimes in total (1% relative to the total number of violent crimes in potentially affected streets). The two most socially costly crimes, homicides and sexual assaults, fall by 60. This represents a large proportion of very serious crimes—8% relative to the total number of homicides and sexual assaults reported in potentially affected segments. None of these estimates are statistically significant, which is one reason why we do not further adjust p-values for multiple comparisons.

6. Discussion

Bogotá's reorganization of police patrols and municipal services represents one of the largest security experiments ever undertaken worldwide. Its policing and urban

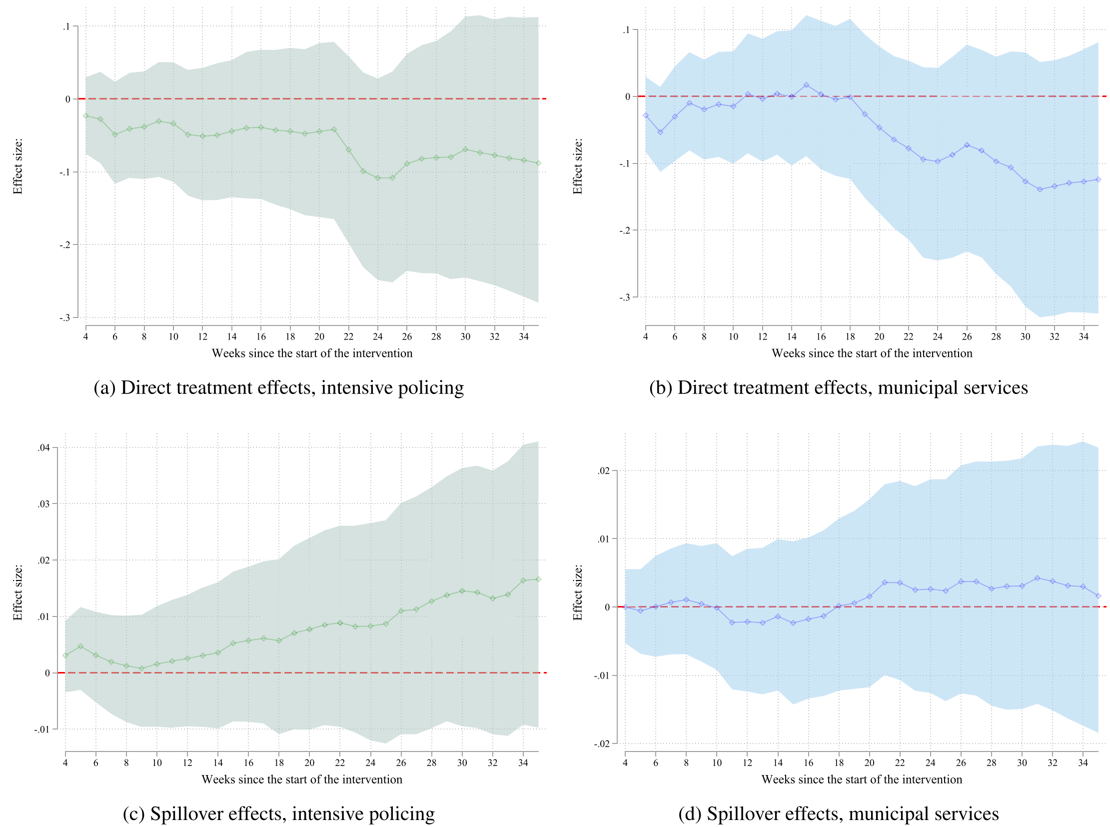

FIGURE 3. Heterogeneous direct and spillover effects over time, using police crime data. The figure depicts point estimates for heterogeneous treatment effects using police crime data. Each circle corresponds to the intent to treat effect in a regression with a restricted sample of hot spots, measured in standard deviations. In each case, the sample includes hot spots at the corresponding percentile or higher based on baseline crime levels. For instance, the circle at 50 for car thefts includes hot spots that are at the 50^{th} percentile or above based on car theft levels in 2014. We estimate the effects using equation (1) assuming short-range spillovers. Filled in figures have RI p-values < 0.1.

TABLE 6. Aggregate impacts on crimes by type (mean and confidence intervals, N = 79,767).

	Total crimes (1)	Est. total impact (2)	95% CI (3)	90% CI (4)
All crime	26,445	813	(−648, 2,192)	(−317, 1,986)
Property crime	17,844	990	(−141, 2,115)	(8, 1,943)
Violent crime	8,604	−177	(−803, 439)	(−695, 341)
Homicides and sexual assaults	794	−60	(−179, 53)	(−162, 40)
Property–violent crime difference		1,167		
p-value		0.071		

Notes: This table presents the aggregate effect calculation for various crime subgroups assuming spillovers within 250 m. Calculations are based on the aggregate effect and confidence interval described in Table 3.

renewal interventions are broadly comparable to US interventions in approach and duration. At first glance, however, the adverse spillovers we observe in Bogotá run against the policy consensus on place-based interventions. After all, meta-analyses suggest that instances of positive spillovers from policing outweigh negative ones (Braga et al. 2019; Weisburd and Telep 2016). On closer inspection, though, our results are broadly consistent with previous findings, and suggest that it is too soon for a policy consensus. A large fraction of past studies have observed adverse spillovers.[43] Moreover, the true confidence interval across these studies probably includes both very large positive and negative spillovers. To use one meta-analysis as an example, of the 13 individual spillover estimates documented in Braga, Papachristos, and Hureau (2014, Figure 2), eight estimates have a p-value smaller than 0.001. Yet these studies are fairly small, with a median number of treated units below 30.[44] Even in a study the size of Bogotá's, levels of precision were never as high as $p < 0.001$. Thus, we view the average direction of spillovers across studies as fundamentally uncertain. This is to be expected: when treated units are surrounded by a dozen or more nearby units, mechanically we should expect potential spillovers to be small, subtle, and difficult to detect without large samples. Bogotá illustrates how such small and imprecise average spillover estimates aggregate to something substantively important, to the point of outweighing the direct gains of an intervention. Going ahead, more large-scale studies are needed.

But, as more urban policy experiments go to scale in a confined space, we need practical tools and methods for dealing with the challenges that come from

43. In a recent meta-analysis, the average point estimate for spillovers for intensive policing was −0.086 standard deviations (Braga et al. 2019). Our 90% confidence interval for the spillover effects of intensive policing on our insecurity index ranges from −0.110 to 0.124, meaning that the US mean is within but at the extreme tail end of our range (table not shown).

44. Another challenge facing any meta-analysis is that the individual papers seldom report sufficient or comparable information. And only recently has it become common to move beyond simple t-tests of means. Few studies adjust standard errors for clustering of treatment assignment. Hence, the precision of past estimates is likely overstated, at least in some of the small-sample cases. This puts meta-analyses in the impossible position of re-litigating past papers where the data are no longer available.

spillovers in dense interconnected networks. This is not just important in cities, it is important for experiments in social networks and other settings where we worry about interference between units, and cannot experiment in separate and independent clusters.

Methodologically, we show how design-based approaches can help. First, researchers can structure the randomization to minimize estimation challenges, especially differential probabilities of assignment to spillover and control conditions. Second, designs can prespecify the estimation of spillovers in a flexible way, with a minimum of ex ante assumptions on the structure and nature of displacement. Such flexibility is especially important when we don't have a strong sense of this structure in advance. Finally, besides illustrating the uses of design, this paper is also a rare example of the practical uses of RI. Bogotá offers a textbook case: units of varying size, with widely different probabilities of assignment to experimental conditions, with spillovers that lead to fuzzy, difficult-to-model clustering. Large-scale urban interventions suffer from both problems, and we show how RI offers a practical solution requiring relatively few assumptions.

More broadly, the results of this experiment have implications for our understanding of criminal motives and behavior. An example comparing our study with a similar and close intervention can illustrate this point. Collazos et al. (2020) report results from a hot spots policing experiment in Medellín, Colombia. They find that violent crimes did not change, while some property crimes decreased (especially car thefts). This could imply that hot spots policing interventions could lead to net benefits when crime is highly concentrated and the motives for crime are not instrumental (such as violent crimes in Bogotá); or when crime is highly concentrated, the motives for crime are instrumental, but criminal rents are generally immobile or hard to displace (such as hot spot locations for car thefts in Medellín). On the contrary, hot spots policing interventions may backfire or lead to net losses when crime is highly concentrated, the motives for crime are instrumental, and criminal rents are generally mobile (such as contract killings in Medellín, where organized crime regulates violence as Blattman et al. (2020) show).

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey Wooldridge (2016). "Clustering as a Design Problem." Working paper, Harvard–MIT Econometrics Seminar.

Abt, Thomas and Christopher Winship (2016). "What Works in Reducing Community Violence: A Meta Review and Field Study for the Northern Triangle." Democracy International, Inc, USAID, Washington, DC.

Adda, Jerome, Brendon McConnell, and Imran Rasul (2014). "Crime and the Depenalization of Cannabis Possession: Evidence from a Policing Experiment." *Journal of Political Economy*, 122, 1130–1202.

Apel, Robert (2013). "Sanctions, Perceptions, and Crime: Implications for Criminal Deterrence." *Journal of Quantitative Criminology*, 29, 67–101.

Aronow, Peter M. and Cyrus Samii (2013). "Estimating Average Causal Effects Under General Interference, with Application to a Social Network Experiment." Working paper.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh (2017). "The Efficient Deployment of Police Resources: Theory and New Evidence from a Randomized Drunk Driving Crackdown in India." NBER Working Paper No. 26224.

Becker, Gary S. (1968). "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76, 169–217.

Blanes i Vidal, Jordi and Tom Kirchmaier (2018). "The Effect of Police Response Time on Crime Clearance Rates." *Review of Economic Studies* 85, 855–891.

Blanes i Vidal, Jordi and Giovanni Mastrobuoni (2017). "Police Patrols and Crime." CEPR Working Paper No. DP12266.

Blattman, C., G. Duncan, B. Lessing, and S. Tobon (2020). "Gangs of Medellin: How Organized Crime is Organized." Working Paper.

Braga, Anthony A., Andrew V. Papachristos, and David M. Hureau (2014). "The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-analysis." *Justice Quarterly*, 31, 633–663.

Braga, Anthony A., Brandon Turchan, Andrew V. Papachristos, and David M. Hureau (2019). "Hot Spots Policing of Small Geographic Areas Effects on Crime." *Campbell Systematic Reviews*, 15, e1046.

Braga, Anthony A., Brandon C. Welsh, and Cory Schnell (2015). "Can Policing Disorder Reduce Crime? A Systematic Review and Meta-Analysis." *Journal of Research in Crime and Delinquency*, 52, 567–588.

Cassidy, Tali, Gabrielle Inglis, Charles Wiysonge, and Richard Matzopoulos (2014). "A Systematic Review of the Effects of Poverty Deconcentration and Urban Upgrading on Youth Violence." *Health and Place*, 26, 78–87.

Chalfin, Aaron and Justin McCrary (2017). "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature*, 55, 5–48.

Chiba, Saori and Kaiwen Leong (2014). "Behavioral Economics of Crime Rates and Punishment Levels." Working paper, Dept. of Management, Universita Ca' Foscari Venezia.

Clarke, Ronald V. and David Weisburd (1994). "Diffusion of Crime Control Benefits: Observations on the Reverse of Displacement." *Crime Prevention Studies*, 2, 165–184.

Collazos, Daniela, Eduardo Garcia, Daniel Mejia, Daniel Ortega, and Santiago Tobon (2020). "Hotspots Policing in a High Crime Environment: An Experimental Evaluation in Medellin." *Journal of Experimental Criminology*, forthcoming.

Dell, Melissa (2015). "Trafficking Networks and the Mexican Drug War." *American Economic Review*, 105(6), 1738–79.

Di Tella, Rafael and Ernesto Schargrodsky (2004). "Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack." *American Economic Review*, 94(1), 115–133.

Draca, Mirko, Stephen Machin, and Robert Witt (2011). "Panic on the Streets of London: Police, Crime, and the July 2005 Terror Attacks." *American Economic Review*, 101(5), 2157–2181.

Ehrlich, Isaac (1973). "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *The Journal of Political Economy*, 81, 521–565.

Farrington, David P. and Brandon C. Welsh (2008). "Effects of Improved Street Lighting on Crime: A Systematic Review." *Campbell Systematic Reviews*, 4, 59.

Ferraz, Claudio, Joana Monteiro, and Bruno Ottoni (2016). "State Presence and Urban Violence: Evidence from Rio de Janeiro's Favelas." Working paper.

Gerber, Alan S. and Donald P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation*. WW Norton.

Guerette, Rob and Kate Bowers (2009). "Assessing the Extent of Crime Displacement and Diffusion of Benefits: A Review of Situational and Crime Prevention Evaluatoins." *Criminology*, 47, 1331–1368.

Horvitz, Daniel G. and Donovan J. Thompson (1952). "A Generalization of Sampling Without Replacement from a Finite Universe." *Journal of the American Statistical Association*, 47, 663–685.

Nussio, Enzo and Ervyn Norza Cespedes (2018). "Deterring Delinquents with Information. Evidence from a Randomized Poster Campaign in Bogota." *PLOS One*, 13, 1–20.

Police Executive Research Forum, (2008). “Violent Crime in America: What We Know about Hot Spots Enforcement.” Technical report, Police Executive Research Forum, Washingon, DC.

Sherman, Lawrence and David Weisburd (1995). “Does Patrol Prevent Crime? The Minneapolish Hot Spots Experiment.” In *Crime Prevention in the Urban Community*. Kluwer Law and Taxation Publishers.

Sherman, Lawrence, Stephen Williams, Ariel Barak, Lucinda R. Strang, Neil Wain, Molly Slothower, and Andre Norton (2014). “An Integrated Theory of Hot Spots Patrol Strategy: Implementing Prevention by Scaling Up and Feeding Back.” *Journal of Contemporary Criminal Justice*, 30, 95–122.

Taylor, B., C. Koper, and D. Woods (2011). “A Randomized Controlled Trial of Different Policing Strategies at Hot Spots of Violent Crime.” *Journal of Experimental Criminology*, 7, 149–181.

Telep, Cody, Renee Mitchell, and David Weisburd (2014). “How Much Time Should Police Spend at Crime Hot Spots? Answers from a Police Agency Directed Randomized Field Trial in Sacramento, California.” *Justice Quarterly*, 31, 905–933.

Vazquez-Bare, Gonzalo (2017). “Identification and Estimation of Spillover Effects in Randomized Experiments.” Working Paper No. 1711.02745.

Weisburd, David, D. Groff, and S. M. Yang (2012). *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*. Oxford University Press.

Weisburd, David and Cody Telep (2016). “Hot Spots Policing: What We Know and What We Need to Know.” *Journal of Experimental Criminology*, 30, 200–220.

Weisburd, David L., Laura A. Wyckoff, John E. Eck, and Joshua Hinkle (2006). “Does Crime Just Move Around the Corner? A Study of Displacement and Diffusion in Jersey City, NJ.” *Criminology*, 44, 549–592.

Wilson, J.Q. and G. Kelling (1982). “Broken Windows: The Police and Neighborhood Safety.” *Atlantic Monthly*, 249, 29–38.

Supplementary Data

Supplementary data are available at *JEEA* online.