Christopher Blattman
Donald Green
Daniel Ortega
Santiago Tobón

Hotspot interventions at scale

The effects of policing and city services on crime in Bogotá, Colombia

December 2018

Impact Evaluation Report 88

Public Sector Management

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality-assure evidence of what works, for whom, how, why and at what cost.

3ie impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, and uses the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of the report, *Hotspot interventions at scale: the effects of policing and city services on crime in Bogotá, Colombia*, as partial fulfilment of requirements under grant DPW1.1044 awarded through Development Priorities Window 1. The content has been copy-edited and formatted for publication by 3ie.

The 3ie technical quality assurance team comprises Monica Jain, Kanika Jha, Avantika Bagai, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Marie Gaarder. The 3ie editorial production team for this report comprises Sahib Singh and Akarsh Gupta, with Beryl Leach providing overall editorial supervision.

All of the content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. All affiliations of the authors listed in the title page are those that were in effect at the time the report was accepted. Please direct any comments or queries to the corresponding author, Christopher Blattman at blattman@uchicago.edu.

3ie received funding for the Development Priorities Window from 3ie's donor, the Bill & Melinda Gates Foundation.

Suggested citation: Blattman, C, Green, D, Ortega, D and Tobón, S, 2018. *Hotspot interventions at scale: the effects of policing and city services on crime in Bogotá, Colombia,* 3ie Impact Evaluation Report 88. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/DPW1IE88

Cover photo: Hemis / Alamy Stock Photo

Hotspot interventions at scale: the effects of policing and city services on crime in Bogotá, Colombia

Christopher Blattman
University of Chicago

Donald Green
Columbia University

Daniel Ortega
Development Bank of Latin America (CAF)

Santiago Tobón
University of Chicago and Innovations for Poverty Action

Impact Evaluation Report 88

December 2018

International Initiative for Impact Evaluation

Acknowledgements

This research would not have been possible without the collaboration of the National Police of Colombia and the Mayor's Office of Bogotá – in particular, Bogotá's Secretary of Security, Daniel Mejía, who co-conceived the interventions and experiment.

Innovations for Poverty Action in Colombia and the Center for the Study of Security and Drugs (Centro de Estudios sobre Seguridad y Drogas; CESED) at Universidad de los Andes coordinated all research activities; survey data were collected by Sistemas Especializados de Información, and for research assistance we thank Juan Carlos Angulo, Peter Deffebach, Marta Carnelli, Daniela Collazos, Eduardo Garcia, Sofia Jaramillo, Richard M Peck, Patryk Perkowski, Oscar Pocasangre, María Aránzazu Rodríguez Uribe and Pablo Villar.

For comments, we thank Thomas Abt, Roseanna Ander, Adriana Camacho, Aaron Chalfin, Marcela Eslava, Claudio Ferraz, 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 numerous conference and seminar participants.

Data collection and analysis were funded by Abdul Latif Jameel Poverty Action Lab's (J-PAL's) Governance Initiative, the International Initiative for Impact Evaluation (3ie), the Development Bank of Latin America (CAF), ProBogotá Fundación, Organización Ardila Lülle through CESED at the Universidad de los Andes, the Administrative Department for Science, Technology and Innovation of the National Government of Colombia (COLCIENCIAS) and the J. William Fulbright Programme.

Summary

This study investigates whether intensive policing and municipal services improved security for targeted hotspots and, importantly, whether these place-based strategies displaced crimes to nearby streets. This study was the largest evaluation of intensive policing ever conducted, and also one of the first randomised evaluations of its kind ever conducted in Latin America.

After working with police to identify 1,919 of the highest-crime street segments in Bogotá, researchers randomly assigned each street segment to one of four groups:

- **Intensive policing:** For eight months, police increased daily patrolling time from 92 to 168 minutes in 756 targeted street segments. Patrols occurred mainly during the day, though in hotspots located near bars and nightclubs, patrols were evenly distributed between day and night. Apart from spending more time in these streets, police did not alter any of their usual behaviour or activities.
- **Municipal services:** In a randomly assigned subset of 201 street segments that showed signs of physical disorder before the study began, the mayor's office instructed municipal maintenance crews to visit, diagnose which services were needed and deliver the appropriate services to these streets. The maintenance teams repaired streetlights and collected garbage.
- **Intensive policing and municipal services:** A subset of street segments received both intensive policing and intensified municipal services.
- **Control:** Police and municipal teams did not receive special instructions about how to work in these areas, nor did they know where the control street segments were.

To measure the impact of these interventions, researchers used police administrative data on reported crimes, a survey of 24,000 citizens and the location of each police patrol in the city every 30 seconds. The survey measured self-reported crimes, perceptions of security, and attitudes towards the police and the local government. To measure the effects on neighbouring blocks, researchers used administrative crime data from the entire city (over 138,000 street blocks) and collected survey data from the experimental crime hotspots as well as a representative sample of 480 street segments in the non-experimental sample.

When assessed in isolation, intensive policing and municipal services interventions do not lead to a statistically significant increase in security in hotspots. These results – assuming the presence of spillovers and no interaction between both treatments – suggest a decrease in the number of reported crimes of approximately 12.6 per cent in streets targeted with intensive policing, and 10.2 per cent in streets targeted with municipal services. These differences, however, are sufficiently small and imprecise to prevent us from saying with confidence that there was an improvement.

However, when both interventions were implemented concurrently, their effect was intensified, resulting in a large and statistically significant impact on security. The results suggest a decrease of about 45.6 per cent in the number of reported crimes in streets targeted with both intensive policing and municipal services simultaneously. Moreover, the effects of both interventions exceeded the sum of the individual effects. These results generally meet conventional levels of statistical significance. Survey measures also

suggest an improvement in perceptions of security in the streets that received both interventions. These results point to increasing returns on state presence on these streets. Also, the combined effects of both interventions were largest in hotspots with the highest crime rates.

In aggregate, the total crime deterred in targeted hotspots was modest. The results suggest that crimes might have been displaced to neighbouring streets, with a potential slight increase in each of the nearly 77,000 streets located within 250 metres of treated hotspot segments. When these displacement effects are added together, the study cannot rule out the possibility that all directly deterred crimes were displaced to other, neighbouring streets. However, there are some indications of a fall in the total number of violent crimes, particularly the most serious: homicides and sexual assaults saw an 8 per cent decrease. The fall in violent crime is sensitive to some assumptions on the distance and behaviour of crime spillovers, however, and must be taken with caution.

When proceeding with place-based security interventions, policymakers should experiment with two kinds of changes to improve security. The first are changes that make these interventions more effective on directly treated street segments. Another alternative is to have less predictable policing. The policy implications of the overall place-based approach hinge on the presence of adverse spillovers. Hence, the second type of change are those that reduce the chances of such spillovers. One way to achieve this is by increasing general police presence alongside the intensification of patrolling time in targeted streets.

Policymakers may also consider complementing place-based approaches with other evidence-based programmes tackling risky people and behaviours, as there are promising interventions in this realm. Finally, another aspect that deserves attention is the possibility that police presence and service delivery function as deterrents in high-crime areas.

Contents

Acknowledgements ... **i**
Summary ... **ii**
Contents ... **iv**
List of figures and tables ... **v**
Abbreviations and acronyms ... **vi**
1. Introduction ... **1**
2. Intervention ... **5**
2.1 Primary and secondary outcomes ... 6
2.2 Hypotheses ... 6
2.3 Hotspot policing ... 7
2.4 Municipal services ... 7
2.5 Interaction ... 7
2.6 Theory of change ... 7
3. Context ... **9**
3.1 Bogotá ... 9
3.2 Security policy ... 11
3.3 Patrolling ... 12
3.4 External validity ... 13
4. Timeline ... **13**
5. Evaluation: Design, methods and implementation ... **14**
5.1 Ethical research ... 14
5.2 Methodology ... 14
5.3 Sample size ... 16
5.4 Data ... 20
6. Interventions ... **23**
6.1 Intensive policing ... 23
6.2 Municipal services ... 24
6.3 Compliance ... 24
7. Impact analysis and results of the key evaluation questions ... **29**
7.1 Estimation ... 29
7.2 Inverse probability weighting (IPW) ... 30
7.3 Balance tests and summary statistics ... 30
7.4 Programme impacts on officially reported crime ... 33
7.5 Direct treatment effects ... 35
7.6 Spillover effects ... 35
7.7 Aggregate effects ... 35
7.8 Heterogeneity by type of crime and by initial level of crime ... 36
7.9 Programme impacts on insecurity ... 37
7.10 Programme impacts on state trust and legitimacy ... 41
7.11 Cost–benefit considerations ... 42
8. Discussion ... **43**
8.1 How do our results line up with the US evidence? ... 43
8.2 Methodological lessons ... 44
9. Specific findings for policy and practice ... **44**
9.1 Lessons for place-based security interventions ... 44
9.2 Lessons for crime prevention and state building ... 45
Appendix A: Sample size and power calculations ... **46**
Online Appendixes ... **48**
References ... **49**

List of figures and tables

Figure 1: Timeline of the project. 13
Figure 2: An example of assignment to the four treatment conditions 15
Figure 3: Proportion of crime reported, by crime (survey based) 22
Figure 4: Heterogeneity of security impacts by pre-treatment administrative crime levels 37

Table 1: Descriptive statistics: Bogotá vs rest of the country – baseline 10
Table 2: Distribution of crime in Bogotá 2012–2015 11
Table 3: Estimates of the municipal services treatment assignment and the inaccessible streets 18
Table 4: Breakdown of municipal services need 19
Table 5: Distribution of hotspots assigned to municipal services treatment 19
Table 6: Distribution of treatment and spillover 20
Table 7: Summary statistics for the primary security outcomes, all experimental conditions 23
Table 8: 'First-stage' effects of treatment on measures of compliance and effectiveness 25
Table 9: Municipal services eligibility and compliance 26
Table 10: Correlation between administrative crime and lighting quality 28
Table 11: Average treatment effect on the index of lighting quality 28
Table 12: Descriptive statistics for the experimental sample and test balance (treatment vs all control streets, including potential spillover streets) 31
Table 13: Estimated aggregate impacts of the interventions, accounting for spillovers within < 250m 34
Table 14: Aggregate impacts on crime by type (mean and confidence intervals) 36
Table 15: Programme impacts on security in the experimental sample, accounting for spillovers within 250m, with p-values from randomisation inference (N = 1,916) 39
Table 16: Programme impacts on security in the experimental sample using exponential decay function 40
Table 17: Impacts on state legitimacy allowing spillover within 250m, with randomised inference p-values 41

Appendix figure and tables

Figure A1: Statistical power in the intensive policing literature 46

Abbreviations and acronyms

CAI Comandos de Atención Inmediata (Commands of Immediate Attention police units)

CI Confidence interval

HSP Hotspot policing

IE Interaction effect

IPA Innovations for Poverty Action

IPW Inverse probability weighting

ITT Intention-to-treat

MS Municipal services

PPP Purchasing power parity

RI Randomisation inference

SD Standard deviation

SEI Sistemas Especializados de Información (Specialised Information Systems)

WLS Weighted least squares

1. Introduction

Police and city workers are the everyday face of the state. These street-level bureaucrats provide the most basic public goods we expect from government, especially security. Responding to crime, picking up garbage and lighting streets – it is impossible to ignore when they are done poorly. When crime and violence start to get out of control, these are also the first levers that governments pull. Cities step up enforcement, they put more police on the streets, or they light up or clean up high-crime places.

In the US, more than 90 per cent of police agencies use some form of 'intensive policing', or intensifying police attention to high-crime areas (Weisburd and Telep 2016).[1] These tactics typically target units as small as a street segment or a specific corner. Some cities also change the quality of policing in hotspots, enforcing minor infractions with a 'zero tolerance' approach.

Another tactic is to reduce disorder in hotspots through municipal services. Such services can make it more difficult to commit crimes by lighting dark areas or increasing the amount of people on the street (Becker 1968).[2] Services may also signal order and state presence, telling criminals to stay away and alerting citizens that the state is present. Altogether, policing and services interventions grow out of the famous 'municipal services' hypothesis (Wilson and Kelling 1982; Apel 2013).[3]

This is state building on a different margin than in weaker states, but it uses the same tools and rationale. From Afghanistan to Iraq or the Philippines, militaries use security forces and public services to establish order and legitimacy (Police Executive Research Forum 2008).[4] In more stable places, such as Bogotá, the state already has some control and legitimacy on most city streets. Here, governments are increasing state presence on the intensive margin – the last mile of state building.

This raises a number of questions. How much can an increased state presence reduce crime and violence? Which levers are most effective? Are there increasing or decreasing returns for state presence? Perhaps the most important but difficult question raised, however, is whether targeted state presence reduces overall crime, or merely displaces it elsewhere. These are the original questions we included in our pre-analysis plan.

We tackle these questions in Bogotá, the capital of Colombia. Two per cent of the city's 136,984 streets accounted for all murders and a quarter of all crimes from 2012–2015. These 'hotspots' received less than 10 per cent of police time and limited public services. In January 2016, a new city government decided to try increasing state presence in hotspots. They wanted to improve security and raise citizens' trust in police and local government.

---

[1] Interventions include greater police time, greater traffic enforcement, aggressive enforcement of infractions and problem-oriented policing.
[2] Police presence and street lighting are meant to raise the risk of detection and capture for offenders – a tenet of the economic approach to crime prevention, where crime is a gamble and increasing expectations of apprehension and punishment deter people from crime.
[3] 'Broken windows policing' can mean intensive, zero tolerance policing. However, a more visible state presence and physical order should send similar signals.
[4] Besides fighting insurgents, intensifying security and public services are designed to win the 'hearts and minds' of citizens. The idea is that they will be more likely to inform on offenders or collaborate against insurgents. See Berman and Matanock (2015) for a review.

We worked with the police to identify an experimental sample of 1,919 hotspot street segments. A segment is a length of street between two intersections, a common unit of police attention (Weisburd et al. 2012). The segments in the sample are mostly middle and low income (91% combined), while the surrounding areas are mainly dedicated to residential (50%) and commercial activities (38%). When we look at the data from 2012–2015, property crime comprises 70 per cent of all crime in hotspot street segments and the city. In order to target crime in hotspots, the city first doubled police patrol time on 756 segments (intensive policing). They then targeted 201 segments for clean-up and better lighting (municipal services). We randomised each area's assignment to intensive policing, more municipal services, both, or neither.

The city modelled its interventions on standard US practices and evidence. As with Bogotá, crime in large US cities is concentrated in a small number of hotspots. Based on several experimental trials, there is a consensus in the US that targeting hotspots with more state presence reduces crime within treated areas.[5] The enthusiasm for intensive policing is bolstered by two systematic reviews that argue that the evidence also points to reductions in crime in nearby streets (Braga et al. 2012; Weisburd and Telep 2016).[6]

We only have data for the years 2012–2015, as data did not exist prior to this. There is a positive correlation in the available data, and in all cases, it is significant at 1 per cent. The correlation between 2015 data and 2014 data is 0.838, and when correlated with 2012 data it is 0.942, which allows us to conclude that the locations of crimes do not tend to move around in the short term.

Spillovers and the aggregate effects on crime are difficult to pinpoint, however, because of the small size of most studies.[7] The median study in existing reviews has fewer than 30 treated hotspots per treatment arm, and the largest has 104. These sample sizes make it difficult to detect large effects, even those as large as 0.4 or 0.5 standard deviations in size (Appendix A). As a result, these studies cannot rule out huge spillovers in either direction. Given the scale of Bogotá's experiment, however, this study can identify direct effects of 0.15 standard deviations and spillovers as small as 0.02 standard deviations.

---

[5] Chalfin and McCrary (2017) review the evidence on increased policing and find that more police are usually associated with falling crime city-wide. Looking at targeted hotspot interventions, a systematic review of intensive policing identified 19 eligible studies (including 9 experiments). Among 25 tests of the core hypothesis, 20 report improvements in crime (Braga et al. 2012). These evaluations are largely in the US. Exceptions include quasi-experimental studies such as Di Tella and Schargrodsky (2004) in Buenos Aires, and ongoing experimental evaluations in Medellín (Collazos et al. 2017) and Trinidad and Tobago (Sherman et al. 2014). The evidence on interventions that tackle disorder is limited. Braga et al. (1999) and Braga and Bond (2008) report significant reductions in crime following a combined treatment of intensive arrests and environmental interventions in small US cities. There is some evidence that street lighting reduces crime (Farrington and Welsh 2008). Cassidy et al. (2014) review five studies, suggesting that there is weak evidence that urban renewal reduces youth violence.

[6] Banerjee et al. (2017) see displacement from drunk driving checkpoints in India. We consider this an important but distinct phenomenon from property and violent crime.

[7] Beyond methodological difficulties, prior studies have been designed mainly to address direct treatment effects and study spillovers as a secondary outcome. One exception is Weisburd et al. (2005), who study drug and prostitution hotspots. Their findings suggest that the benefits from the intervention diffuse to nearby areas.

Latin America is an important place to study the state's crime-fighting abilities. It is the most violent region in the world, home to 42 of the 50 most dangerous cities and one-third of the world's homicides (UNODC 2014). Major cities also have fewer police per person than the US or Europe. Policymakers are interested in the returns on a higher quality or quantity of policing.

In Bogotá, the mayor's office first reallocated existing police patrols to spend more time on high-crime streets. No new police were added in the city. Within their patrol area (a quadrant), officers were told to double their time on two hotspots from roughly one hour to two hours a day, in multiple visits. This intensive policing lasted from February to October 2016. With an average of 130 segments per quadrant, there was little effect on patrol time on other segments. Patrols simply went about their normal duties, interacting with citizens, and stopping and frisking suspicious people. Shortly afterwards, the city decided to tackle social disorder by repairing lights and cleaning up trash.

We designed the study to measure spillovers flexibly. Treating one hotspot can affect the outcomes in control hotspots. For example, criminals may shift activities to nearby hotspots, and areas close to treated segments must be crossed to deliver interventions. Thus, spillovers pose an identification problem for direct effects. We are also interested in spillovers to nearby streets outside the experimental sample, or non-hotspots. Taken together, these two spillovers tell us whether crimes are deterred or pushed around the corner.

Since we do not know the structure of spillovers, we pre-specified a more flexible design over many possible catchment areas. We divided control hotspots into categories: 0–250 metres from a treated hotspot; 250–500 metres; and more than 500 metres. By comparing outcomes across treatment and control categories, we can first test for spillovers in the 0–250m and 250–500m regions, and then use unaffected regions as a control group for estimating the effects of direct treatment. We estimate spillovers into the non-experimental sample the same way.

Spillovers present other estimation challenges, however. By simulating the experiment many times, we show that the close proximity of hotspots leads to hard-to-model patterns of clustering, also known as "fuzzy clustering" (Abadie et al. 2016). In most randomisations, hotspots close to other hotspots tend to be assigned spillover status. This biases treatment effects and understates standard errors. Without a fixed geographic unit of clustering, we cannot use standard correction procedures. This is a common but relatively underexplored problem with experiments in dense social or spatial networks. We show that randomisation inference provides exact p-values in such settings.

To evaluate impacts, we first looked at police administrative data on reported crimes. Police data are problematic, however, if errors in crime reporting are correlated with treatment. Therefore, we also conducted a survey of approximately 24,000 citizens. The survey measured unreported crimes, perceptions of security, and attitudes towards the state. Besides providing new outcomes, these data help us test whether official crime reporting is correlated with treatment.

Broadly speaking, the interventions deterred some crimes on directly treated streets, but we cannot reject the hypothesis that this crime was displaced to neighbouring segments. First, we find that intensive policing and municipal services each improved security by roughly 0.1 standard deviations using both survey and administrative data. Once we account for spillovers, however, neither decrease is statistically significant. For example, if we consider crimes reported to the police, our best guess is that eight crimes were averted on directly treated streets (a 1% fall). Other estimates are as high as 86 reported crimes deterred (a 12% fall). None of these decreases, however, are statistically significant.

The crime impacts were greatest in the 75 hotspots that received both interventions. In this case, security increased by more than 0.3 standard deviations, which is statistically significant at the 5 per cent level. This is equivalent, for example, to a 45 per cent decrease in officially reported crimes in the segment. The difference between getting both treatments or only one treatment is not always statistically significant, but it points in the direction of increasing returns for state presence on these streets. The combined effects of both interventions were also largest on the highest-crime hotspots.

Meanwhile, we see some evidence of adverse spillovers. We estimate spillovers into nearby control hotspots, as well as onto the nearly 77,000 non-hotspot segments within 250m of the experimental sample. At the segment level, these spillovers are small in magnitude: for example, in one specification, intensive policing led to an increase of only 0.016 crimes in neighbouring non-hotspot segments. Across more than 50,000 segments close to intensively policed hotspots, however, these small effects add up. We estimate a 90 per cent confidence interval for the effect of both interventions on aggregate crime on all city streets. It ranges from -735 (a 2% decrease in crime) to 2,033 (a 5% increase). The average, 768, represents a 2 per cent increase in crime. Thus, at a minimum, our study likely rules out a decrease of over 2 per cent in city-wide crime.

In our main specification, it is mainly property crime, as opposed to violent crime, that is displaced. There is some evidence that the interventions led to a decrease of nearly 100 homicides and sexual assaults, which is an 8 per cent decline in these most-serious crimes. This difference between violent crime and property crime is statistically significant. This result seems to be sensitive to specification, however. For example, violent crimes significantly increase city-wide if we estimate spillovers using a continuous rate of decay. Thus, the property/violent crime distinction must be made with caution.

These results show the importance of small spillovers and the statistical power of the experiment. Prior studies have not been powered to detect economically important spillovers. Yet the cumulative effect of many tiny spillovers is obviously important in evaluating the interventions and understanding the relationship between state presence and violence. This is especially true when we need to assess the aggregate effects on crime or distinguish between types of crime. Even with a sample size that is an order of magnitude greater than the previous four experiments, the spillover and aggregate effects are difficult to identify. Thus, methodologically, this study illustrates the importance of scale in estimating the effects of place-based interventions, as well as the importance of accounting for interference between treatment and control units. It also shows the importance of using randomisation inference to avoid overstating precision.

If our results are also accurate more generally, they add nuance to a common argument in criminology: that crime and violence are concentrated among a small number of people, places and behaviours, and that targeted interventions stand the best chance of being effective (Braga et al. 2012; Abt and Winship 2016; Weisburd and Telep 2016; Weisburd et al. 2017). Alongside another large-sample study of policing – a study of drunk driving checkpoints by Banerjee and others (2017) – our evidence reinforces the idea that crime is concentrated, but targeting particular places may not be generally effective, as crime may simply be pushed around the corner.[8] If place-based interventions simply displace crime, then targeting high-risk people and behaviours could be more impactful in addressing this kind of criminal behaviour. Nonetheless, we conclude with a discussion of how place-based policing and services could be more effective.

There are parallels between our results and the historical literature on states, where the most common response to state coercion has been to elude the state or run away (Scott 2014). The perennial problem of state building is controlling people, not land. The evidence from Bogotá suggests this could prove true even in the last mile of state building.

We begin this report by explaining the intervention, the research hypotheses and the underlying theory of change in section 2. In section 3, we describe the general crime situation in Bogotá, its security policy and a general outline of how police patrolling is organised. In section 4, we provide a timeline of the intervention and the evaluation. In section 5, we talk about the design and methods used for the study, and in section 6 we give detailed explanations of how the mayor's office and the police conducted the intensive policing and municipal services interventions, and their level of compliance. In section 7, we show the results, which are discussed in section 8. Lastly in section 9, we offer policy recommendations based on the findings of the evaluation.

2. Intervention

This experiment had two interventions. The first was an intensive policing intervention that consisted of increasing patrolling time from approximately 55 minutes per day per hotspot street segment to 90 minutes per day, which was divided into 6 entries of 15 minutes each. Police patrols were given specific instructions on how to distribute entries across the day. For hotspots in the control group, police did not receive any special instructions and were free to patrol as they saw fit. Activities while patrolling were standard (i.e. criminal record checks, door-to-door visits to the community, arrests, drug seizures, etc.).

The second intervention was a municipal services intervention, which consisted of sending a municipal team to selected hotspots to clean up streets in order to signal state presence and order. The municipal team was charged with repairing street lights and cleaning graffiti, and collecting garbage every few weeks.

[8] Similarly, Blanes i Vidal and Mastrobuoni (2017) use natural, high-frequency variation in police presence in the UK to argue that the deterrence effect of police lasts for a maximum of 30 minutes.

These policies were already being implemented to some extent by the police and the city. The police already targeted some of their patrolling to particularly difficult neighbourhoods, and the city offered municipal services such as street lighting and garbage collection throughout the city. What was new about the intervention was the targeting and the intensity of services to the highest-crime streets to observe their impact on the margins, as well as monitoring activities around police compliance.

2.1 Primary and secondary outcomes

Our primary outcomes are two insecurity measures: perceived risk and crime incidence.

2.1.1 Perceived risk of crime and violence on the segment

Our citizen survey asked respondents to rate perceived risk on a four-point scale from ‘very unsafe’ to ‘very safe’ in five situations: a young woman walking alone after dark on this street; someone talking on their smartphone on this street; a young man walking alone after dark on this street; and simply the perceived risk of crime ‘during the day’ and ‘at dusk’. We constructed an index of perceived risk that determines the average across all respondents in the segment. All indices in the paper are standardised to have mean zero and unit standard deviation.

2.1.2 Crime incidence on the segment

We constructed a standardised index of crime that gives equal weight to the survey and administrative data. The two components include: (1) survey of respondents’ opinions of the incidence of crime in that segment, as well as personal victimisation on that segment since the beginning of the year; and (2) the total number of crime incidents in that segment reported in the administrative crime data since the beginning of the intervention. We can subdivide all measures into property and violent crimes, although our main measure pools all crimes into one index.

The survey measured perceived incidence and personal victimisation by walking respondents through a list of 11 criminal activities. After finding out whether any of these activities had happened on the street since the beginning of the year, we asked respondents about each crime to establish perceived frequency (ranging from ‘everyday’ to ‘never’ on a 0–6 scale), and whether it had happened to the respondent him/herself in 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.

We pre-specified three secondary outcomes capturing impacts on trust in and legitimacy of the state. First, an opinion of the police index determined the average of four attitudes towards police: trust, quality of work, overall satisfaction, and likelihood they would give information to police. Second, an opinion of the mayor index asked the same four questions about the city government. Third, a crime reporting measure captured the likelihood that people would report a crime to the police. This helps us to understand whether administrative crime reporting changes with treatment, while also measuring the level of collaboration with police and perceptions of their legitimacy.

2.2 Hypotheses

Our pre-analysis plan (Online Appendix D) focused on intention-to-treat (ITT) estimates to identify the impact of our intervention regardless of whether the police and/or municipal team complied with the treatment status of each street segment. This analysis addressed the policy question of whether an intensive policing and/or municipal services strategy affects crime months after the intervention. We hypothesised that not only will treatment affect segments assigned to 90 additional minutes of policing, but there will also be spillover effects onto nearby untreated units.

2.3 Hotspot policing

We hypothesised that treatment would increase policing time in treated hotspots and decrease policing time in non-treated spillover segments. We also hypothesised that treatment would decrease crime in segments assigned to intensive policing. We tested these hypotheses using a two-tailed test. For spillover control units, the direction of the crime effect is uncertain: spillover segments might see increases in crime due to displacement or decreases in crime due to deterrence and a diffusion of benefits. For this reason, we used a two-tailed test.

Furthermore, the effects on spillover segments may vary across time. We hypothesised that if the effect of police presence on crime is gained mostly through deterrence, the reallocation of police presence should only have a transitory effect on crime unless this reallocation leads to a permanent shift in offenders' perceived probability of apprehension. Even if there are positive spillovers in the short run, there may still be some crime displacement over a longer period, as offenders seek out new low-cost locations.

2.4 Municipal services

We hypothesised that treatment would increase the number of times a segment is visited by the municipal team, and it might also reduce crime. We tested these hypotheses using a two-tailed test. We anticipated decreases in crime for spillover units and tested this using a two-sided test. However, we hypothesised that the spillover effect would be less pronounced than that in the intensive policing case.

2.5 Interaction

We hypothesised that segments receiving both the intensive policing and municipal services treatments would see larger decreases in crime than segments receiving just one. We tested this hypothesis using a one-tailed test.

2.6 Theory of change

The City of Bogotá has begun with two of the most commonplace, high-profile theories of crime reduction in the US (e.g. Abt and Winship 2016). Two main theories underlie the current interventions examined in this study. The first is the economic theory of crime introduced by Becker (1968), which argues that individuals take into account the probability of apprehension and punishment when making the decision to engage in a criminal act. Therefore, increasing the likelihood of apprehension and punishment should prevent criminals from taking part in illegal activities. The second theory is the municipal services hypothesis (Wilson and Kelling 1982), which reconciles with the previous theory by introducing criminals' subjective perceptions of apprehension and punishment. That is, the conditions of an environment may carry signals about social norms and enforcement capacity. If the environment presents itself as highly disordered, criminals may believe that police presence and other enforcement efforts are weak in that location.

The first intervention (intensive policing) increased police presence in hotspots by reallocating police from street segments with less crime to those with more. We theorised that police officers in charge of policing those hotspot segments assigned to treatment would increase their number of patrol minutes from 55 minutes per day to 90 minutes per day. The activities the police performed while patrolling would be standard (criminal record checks, door-to-door visits to the community, arrests, drug seizures, etc.). Potential criminals and the community would then become aware of the increased police presence in these streets.

The reaction of potential criminals to the increased patrol time can be examined through the economic theory of crime. The increased amount of time that police patrols spend in these segments increases the likeliness of apprehension and punishment. Because potential criminals are more likely to be caught, the expected cost of engaging in criminal activities in these areas rises. For some individuals, these higher costs will now outweigh the benefits of committing the crime, thereby leading to a decrease in crime. This increase in police presence may also have an effect through the municipal services hypothesis mechanism, in the sense that the mere presence of police patrols, even if they are not taking action, may change the perception of surveillance and therefore decrease incentives to commit crimes among potential offenders.

The second intervention (municipal services) aimed to reduce street disorder and create an environment of lawfulness. Municipal services teams were sent into hotspots to clean up trash and graffiti, and repair non-functional streetlights.

The reaction of potential criminals to the improved physical environment can also be viewed through the lens of Becker's theory of crime, although it has roots in Wilson and Kelling's (1982) theory. The physical environment of streets carries signals about social norms and enforcement and helps promote some forms of criminal behaviour. Potential criminals will become aware of the improved physical environment and believe that police presence and other enforcement efforts are stronger at this location. Therefore, the subjective perception of apprehension and punishment will rise. Similar to the policing intervention, this will increase the cost of engaging in criminal activity and thus decrease crime.

There are several challenges to these theories, however, which we feel the existing literature has yet to address:

First, there is the risk that both interventions displace, rather than reduce, crime. If the expected cost of crime increases for a criminal in his/her preferred location, he/she may move to another less costly (less patrolled) place instead of choosing not to commit the crime. Existing studies are generally too limited, or do not have the appropriate data, to test for such displacement, especially spillovers to non-adjacent streets. We addressed this concern by looking at spatial displacement within police quadrants.

A second challenge is that increased contact between the police and the community may actually lower intergroup trust and cooperation. For example, if the additional patrolling time is spent on making arrests that the community deems to be unfair or unnecessary, the police may face backlash from the community, causing trust to fall. While we did not have control over the actions of the police during the additional patrolling time, we collected various sets of data to understand the community's response to the additional policing. Our endline survey contained questions about perceptions of police power.

A third challenge is that interventions like these may increase violence in the short term but improve development outcomes in the medium term (e.g. Berman and Matanock 2015; Berman et al. 2016). For example, gangs that hold control in some areas of the city may respond to the increased state presence with additional violence. On the other hand, if such gangs are removed, they may leave a power vacuum that creates a struggle for local control.

A fourth challenge is that the deterrence effect of policing will depend on the likelihood and severity of punishment; if prosecution rates are low, the expected cost of crime for a criminal will decrease.

Because of the essential role that our theory of change plays in understanding the impact of our intervention and various risks associated with this theory, a portion of our evaluation will include a qualitative component to understand and document the channels through which the changes take (or do not take) place.

3. Context

3.1 Bogotá

Bogotá, a city of roughly eight million people, is the industrial and political centre of Colombia. In 2015, Bogotá's gross domestic product per capita was US$9,612 at market exchange rates, or about $22,000 adjusted for purchasing power parity (PPP). Ten per cent of the population were below the national poverty line for metropolitan areas of PPP$6 a day, and 2 per cent were below the extreme poverty line for metropolitan areas of PPP$2.50 a day. Many poor were displaced by a low-intensity civil war that ran for half a century until a peace agreement in 2016.

Crime is one of the most pressing social problems in Bogotá. In the 1990s, Bogotá was one of the most violent cities in the world, with 81 murders per 100,000 people.[9] In 2016, this figure was 15.6. This is much lower than the most violent cities in the world, such as Caracas (120), Cape Town (65), Detroit (64), and Cali, Colombia (64). It is comparable to a US city like Chicago, with 15 murders per 100,000 in 2015, but greater than the 7 recorded in Los Angeles or 4 in New York.[10] As in cities like Chicago, despite improvements, crime remains one of the foremost social and political concerns.

Table 1 presents a comparison of outcomes between Bogotá and the national population of Colombia with regard to the mean and standard deviation for violent and property crimes. It compares crime rates (per 100,000 people) between Bogotá and the rest of the country from 2010 to 2015. This information comes from a police database for reported

[9] It had 81 murders per 100,000 people in 1993. A number of factors are said to have contributed to the improvement, including the decline in civil war, as well as advances in police capacity, gun control policies, restrictions on alcohol consumption, and a major local security push. Figures from the World Atlas.
[10] US figures come from the FBI Uniform Crime Report and others from the World Atlas.

crimes. As seen in the table, Bogotá generally has higher crime rates than the national average.[11]

Table 1: Descriptive statistics: Bogotá vs rest of the country – baseline

	Bogotá		Rest of the country	
	Mean	SD	Mean	SD
Reported crime rate (100,000 pop). From 2010 to 2015.				
Violent crimes	171.94	20.92	140.83	125.51
Homicide	17.54	2.32	24.91	30.68
Injuries	154.40	21.97	115.92	119.61
Property crimes	356.21	59.23	66.81	99.22
Personal robbery	294.61	59.62	47.97	78.39
Car theft	32.78	6.10	4.06	9.59
Motorcycle theft	28.82	6.33	14.78	30.04
Municipalities	1		1121	
Years	6		6	
Observations	6		6726	

Note: SD = standard deviations. It is important to note that the data in the table include all areas of Bogotá (urban and rural), while the study only considered Bogotá's urban area.
Source: National Police of Colombia.

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 are crimes of passion. The mayor's office estimates that 81% of all homicides in the city in 2015 were the result of fights, 12% were contracted killings, and 5% resulted from violent robberies. Most offenders are individual young people. There are some semi-organised youth gangs and some organised crime, but they do not seem to be responsible for the vast majority of street crime or violence.

Bogotá is divided into geographic units called *localidades*. Table 2 displays the distribution of crime by *localidades* and the most prominent type of crime in each.

[11] We do not have information about risk perception and police performance rating for the rest of the country or for the covariates, as these data are not collected for the national population; therefore, we could not compare these outcomes.

Table 2: Distribution of crime in Bogotá 2012–2015

Name	Total crimes		Name	Total crimes	
	%	Most prominent		%	Most prominent
	(1)	(2)		(3)	(4)
Antonio Nariño	2.68	Property	Puente Aranda	3.95	Property
Barrios Unidos	3.33	Property	Rafael Uribe	4.87	Property
Bosa	5.78	Violent	San Cristobal	4.08	Property
Candelaria	1.47	Property	Santa Fe	4.83	Property
Chapinero	6.41	Property	Suba	11.38	Property
Ciudad Bolivar	7.60	Violent	Teusaquillo	4.38	Property
Engativa	8.00	Property	Tunjuelito	2.52	Property
Fontibon	4.40	Property	Usaquen	6.70	Violent
Kennedy	10.61	Property	Usme	3.38	Violent
Los Martires	3.63	Property			

Notes: The table displays the proportion of crime by *localidad*. The first and third columns represent the percentage of the total crime in the city for each *localidad*, while the second and fourth columns display the most prominent type of crime, namely property or violent crime. Source: National Police of Colombia.

Like many cities, crime in Bogotá is also highly concentrated. According to official crime statistics, from 2012 to 2015 just 2 per cent of the city's 136,984 street segments accounted for all murders, as well as one quarter of all other reported crimes. These hotspots are distributed around the city. 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. Hotspots also include popular nightlife areas.

Bogotá has moderate to low levels of police compared with large US and Latin American cities. Bogotá has about 18,000 police officers in operational activities, including approximately 6,200 patrol agents. We estimate that there are about 239 police per 10,000 people. The Colombian average is 350, and most cities are above Bogotá's ratio. The national ratio in the US was 230 in 2013 but it is greater in large cities, including 413 in New York, 444 in Chicago, 611 in Washington or 257 in Los Angeles.[12]

3.2 Security policy

Patrols are instructed to spend more time in high-crime places but do not necessarily comply. One indication is that 2 per cent of streets account for one quarter of all crime, but we estimate that they received roughly 10 per cent of police patrol time during 2012–2015.

[12] Data for Colombia was reported by the Secretariat of Security of Bogotá; data for the US are from the Department of Justice Statistics; and other data are from the United Nations Office on Drugs and Crime.

We discreetly observed police patrols and performed qualitative interviews with residents in 100 of the treated hotspots. The police freely patrol almost all city streets. Our assessment is that patrols are reasonably well regarded. The broader police force is not without problems, but our citizen survey (detailed below) suggests that street patrol officers are generally regarded as competent and non-corrupt. If anything, residents complained that officers were not present often enough.

In January 2016, Enrique Peñalosa came to power as the new mayor. Crime reduction and increasing trust in government were central to his platform. In his first 100 days, the mayor pledged to dedicate more municipal services and law enforcement in 750 hotspots.

Municipal services included trash collection, tree pruning, graffiti clean-up and streetlight maintenance. The performing agencies report directly to the mayor's office, but the mayor's power is limited by contracts and difficulties in monitoring and enforcing instructions.

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 Defence; however, the city has the power of the purse, as it controls the budget and pays for police equipment. The Colombian constitution also calls on police to comply with the mayor's 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 and quality, especially street patrols.

3.3 Patrolling

The quadrant (*cuadrante*) is the basic patrolling unit. Bogotá has 19 urban police stations. Stations are divided into CAIs (Comandos de Atención Inmediata), which are small local police bases that coordinate patrol agents and take civilian calls. Each CAI has about 10 quadrants. There are 1,051 quadrants, each with an average of 130 street segments.

Each quadrant has six permanent patrol officers. They patrol in pairs, on motorbikes and on foot, in three shifts of eight 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, passers-by and suspicious people.

Patrols carry a handheld 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 any suspicious people, and will seize illegal weapons (usually knives) and other contraband. Patrols tend to focus interrogations on young men. An arrest means both patrollers must take the suspect to the station for hours of paperwork and processing. This keeps them from meeting performance goals, and so patrols may avoid minor arrests.

The handheld computer also contains a global positioning system (GPS) chip that records the patrol's location roughly every 30 seconds (when operational). The city first 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 entailed increased monitoring and measurement of patrol activity.

3.4 External validity

In most cities in the world, crime is highly concentrated in a small number of places, and intensive policing has been used in places as diverse as Minneapolis, Trinidad and Tobago, and now Cali, Medellín and Bogotá. Similarly, municipal services interventions have been tried in places like Jersey City, New Jersey and Lowell, Massachusetts. Therefore, the transferability of the intervention itself to different contexts is possible.

The fact that we are working in a single city limits external validity. This experiment in Colombia offers the first large-scale randomised trials of intensive policing outside a developed country. Therefore, the external validity is fundamentally uncertain. At the same time, this entire trial is a test of external validity outside the US.

4. Timeline

Figure 1 shows the timeline of the project's development.

Figure 1: Timeline of the project

Year	**2016**				**2017**				**2018**
Quarter	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**
Activities									
Start of Peñalosa's administration									
Citizen baseline survey									
Ensuring involvement of IA and other stakeholders in the design of the study									
Intervention hotspots									
Intervention municipal services									
Citizen endline survey									
Qualitative research									
Lighting survey									
Policy brief									
Working paper									
Meetings with stakeholders to present the results									
Data and coding publicly available									

5. Evaluation: Design, methods and implementation

5.1 Ethical research

Policing involves coercion, and Bogotá police have mixed levels of legitimacy in the eyes of the community. Police tactics are sometimes controversial, including the routine use of 'stop and frisk' tactics. Randomising police intensity has real personal consequences for citizens on treatment and control streets. These consequences are, however, unknown, and it is the aim of this study to assess them.

There are a number of factors that argue for the importance of this research, most of all a democratic mandate for the interventions themselves and a careful identification of what works:

- Insecurity and crime are now the top concern of Bogotá residents according to public opinion polls;
- The democratically elected Mayor Peñalosa put criminal reform and policing as his central campaign pledge;
- The mayor identified intensive policing and targeted municipal services as one of his ten major objectives for the first 100 days of his administration;
- The first senior official the mayor selected for his administration was Daniel Mejía for Secretary of Security, largely based on Mejía's record of rigorous research and evidence-based policy, which was also an aim of Peñalosa's;
- Virtually no evidence has informed policing and security to date; and
- The central objective of the study is to closely track citizen well-being and security.

Moreover, the project has already been approved by Institutional Review Boards in Colombia, Chicago, Los Andes and Innovations for Poverty Action (IPA), covering all Principal Investigators involved.

5.2 Methodology

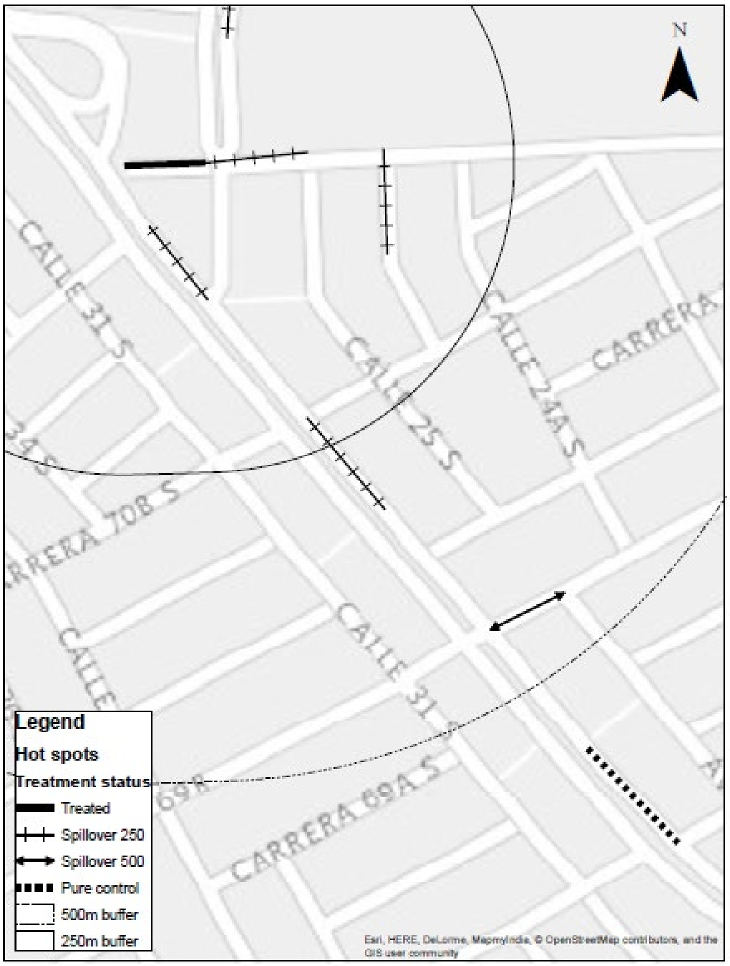
The size and direction of spillovers drive the policy implications of place-based anti-crime programmes. Failing to account for spillovers could also bias our estimates of direct treatment effects. If control hotspots are close enough to treated hotspots to experience displacement or diffusion, then spillovers violate the standard assumption of 'no interference between units'. Previous studies have generally ignored the possibility of interference between treatment and control hotspots, and focused instead on spillovers into nearby non-hotspots. This is reasonable in small samples where hotspots are widely dispersed and the spillover regions do not overlap. But interference between units grows large as we scale up to hundreds of treated hotspots in a city. The same would be true of any intervention in a spatial or social network. This is a growing source of experimental work. We illustrate how to approach these challenges through the experiment design and randomisation inference.

We did not know the range of spatial spillovers, and so we pre-specified a flexible design that tested for spillovers in radii of 250m and 500m around treated streets. There are many other ways to model spillovers, and we test robustness to a continuous rate of decay, as well as different radii. Previous literature on hotspots policing has focused mainly on catchment areas of about two blocks or 150m (Braga et al. 1999; Braga and Bond 2008; Mazerolle et al. 2000; Taylor et al. 2011; Weisburd and Green 1995). We felt 150m to be too conservative, however, and opted for 250m instead. We also specified a 500m option in case spillovers were unexpectedly large. Wider radii seemed implausible and would have eliminated the pure control category in a single city.

Our preferred approach partitions control segments into one of three experimental conditions according to their distance from the treated segment: less than 250m, 250–500m, and more than 500m. Figure 2 illustrates this partition. The hotspot segment at the centre of the two radii was assigned to the intensive policing treatment. For simplicity, Figure 2 ignores municipal services. Nearby hotspots are classified by their distance to the treated segment. One virtue of this approach is that all treatment effects estimates are simply differences in the means of the experimental conditions. We can also use this design to assess spillover effects on non-hotspots outside the experimental sample. We opt for regression-based estimates to control for possible confounders, as described below, but these preserve the spirit of the mean differences approach.

Our approach ignores the possibility of spillovers beyond 500m, as well as non-spatial spillovers. Some crime is undoubtedly displaced in non-Euclidean ways (e.g. to possibly distant hotspots where the benefits of crime are high and the risk of detection is low).

Figure 2: An example of assignment to the four treatment conditions



Source: Base map – ESRI, Open Street Maps and GIS user community.

5.3 Sample size

Our experimental sample consists of the 1,919 hotspot street segments with the highest aggregate crime index from January 2012 to September 2015 that were verified by the police to be areas of crime.

To estimate the minimum detectable effects our experiment is powered to detect, we regressed crime from 2015 (mean of 12.99, standard deviation of 51.25) on treatment assignments from our randomisation procedure, baseline controls including crime from 2012 to 2014 and previous police patrolling time, and block fixed effects using the weighted pairwise regression described below in section 7. All power calculations were based on a two-tailed hypothesis with a significance level of 0.05 and 80 per cent statistical power. To account for the interrelationship between segments, we posited four sets of potential outcomes: being treated; being untreated but within 250m of a treated segment; being untreated but between 250m and 500m of a treated hotspot; and being untreated and more than 500m away from any treated hotspot. Our power calculations are all relative to this final group.

Our power calculations reveal that for our policing intervention, we are powered to detect an effect of 0.10 standard deviations for treated units, 0.13 standard deviations for spillover units within 250m of a treated hotspot, and 0.12 standard deviations for spillover units between 250m and 500m of a treated hotspot. For the municipal services treatment, these numbers are 0.08, 0.11 and 0.10, respectively. We are also powered to detect an interaction effect of 0.13 standard deviations for units receiving both treatments. For the non-experimental sample, we are powered to detect effects of about 0.03 standard deviations for both types of spillover units.

Because we are using administrative data to evaluate our intervention, attrition was not an issue.

The mayor of Bogotá promised to deliver at least 750 treated hotspots, so our goal was to randomly assign between 750 and 770 hotspots that would receive at least 90 minutes of policing per day, with the remaining hotspots assigned to a control group for which the police station would receive no special instructions but would be free to patrol them as they saw fit. Restrictions on the operational capacity of the police implied that any given quadrant could not have more than two treated hotspots. We began by randomly assigning each quadrant with at least one hotspot to either treatment or control with a treatment probability of 0.60, blocking by police station.

Because of our blocking strategy, the number of treated hotspots falls below 750 in some randomisations. We restricted our randomisations to only those where at least 750 hotspots were assigned to treatment, but no more than 770 (the maximum the police could handle). For quadrants assigned to treatment, we then assigned hotspots to treatment or control using the following rule:

- Quadrants with one or two hotspots: assign both to treatment.
- Quadrants with more than two hotspots: randomly assign two to treatment and the rest to control.

A randomisation was deemed successful only if the number of hotspot segments assigned to treatment was between 750 and 770. Our randomisation procedure assigned 756 hotspots to treatment and 1,163 to control. Treated hotspots account for 24 per cent of the aggregate crime index, while untreated hotspots account for 31 per cent of the aggregate crime index.

Determining eligibility for the municipal services treatment required a segment-level measure of disorder. For this reason, we sent out enumerators to all 1,919 hotspots to take five photographs and rate hotspots on the presence of graffiti, municipal services, garbage, boarded-up buildings and run-down buildings. Enumerators were able to reach 1,534 out of 1,919 hotspots.

We ran the following weighted least squares (WLS) regression to test whether there were differences between the accessible and inaccessible street segments:

$$Y_{sqp} = \beta_1^{IA} M_{sqp} + \beta_2^{IA} I_{sqp} + \beta_3^{IA} (M \times I)_{sqp} + \gamma_p + \theta^{IA} X_{sqp} + \epsilon_{sqp}^{IA}$$

Table 3 shows the estimates of the municipal services treatment assignment and inaccessible streets. The inaccessible streets seem to have less crime, though there is no difference between accessible and inaccessible streets except the crime rate in their quadrants. The missing streets are less likely to be high or middle income. There are differences between accessible and inaccessible streets in terms of the number of segments in the quadrant.

Table 3: Estimates of the municipal services treatment assignment and the inaccessible streets

	Control	Municipal services		Inaccessible areas		Interaction	
	mean	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
# of reported crimes on street, 2012–2015 (original)	4.842	-0.335	0.406	-1.373	**0.000**	1.083	0.114
# of violent crimes	1.860	-0.147	0.389	-0.092	0.657	0.429	0.259
# of property crimes	2.983	-0.188	0.548	-1.281	**0.000**	0.655	0.206
# of reported crimes on street, 2012–2015 (updated)	6.123	-0.522	0.762	-4.242	**0.001**	1.550	0.497
# of violent crimes	1.443	0.039	0.949	-1.211	**0.011**	0.830	0.475
# of property crimes	4.680	-0.561	0.624	-3.031	**0.001**	0.720	0.575
# of crimes on quadrant	3.706	-0.207	0.404	-0.931	**0.004**	2.480	**0.080**
Average daily patrolling time (11/2015), minutes	37.498	-0.959	0.857	-1.752	0.720	17.208	0.338
Metres from police infrastructure	560.156	-4.179	0.867	3.582	0.879	-23.260	0.730
Zoned for industry/commercial	0.373	0.062	0.156	-0.032	0.369	-0.008	0.924
Zoned for services	0.130	0.027	0.350	-0.011	0.660	0.007	0.897
High-income street segment	0.094	-0.020	0.182	-0.045	**0.006**	0.051	0.108
Medium-income street segment	0.583	0.035	0.340	-0.224	**0.000**	-0.066	0.359
# of segments in quadrant	121.815	4.472	0.450	23.812	**0.000**	-33.676	**0.006**
# of experimental units in quadrant	4.034	-0.060	0.771	-0.552	**0.007**	-0.199	0.537
# of HSP treated units in quadrant	1.212	-0.084	0.329	-0.069	0.311	-0.030	0.858
# of MS treated units in quadrant	0.305	0.915	**0.000**	-0.084	**0.079**	0.031	0.685
Intensive policing assignment: treated	0.506	-0.085	**0.058**	-0.016	0.684	0.072	0.426
Intensive policing assignment: proximal spillover	0.283	0.078	**0.052**	-0.010	0.749	-0.128	**0.056**
Intensive policing assignment: distant spillover	0.140	0.020	0.548	-0.028	0.188	0.006	0.925
Intensive policing assignment: pure control	0.071	-0.012	0.537	0.053	**0.014**	0.051	0.379
Municipal services assignment: proximal spillover	0.338	-0.336	**0.000**	-0.058	**0.069**	0.091	**0.009**
Municipal services assignment: distant spillover	0.278	-0.284	**0.000**	0.025	0.430	-0.006	0.846
Municipal services assignment: pure control	0.384	-0.380	**0.000**	0.034	0.288	-0.084	**0.023**

Note: HSP = hotspot policing. MS = municipal services.

Source: National Police of Colombia and Mayor’s Office of Bogotá.

We created a 0–5 index of need for the municipal services treatment using the data mentioned above. Table 4 displays the breakdown of this index for the 1,534 segments our enumerators were able to rate. Thirty per cent of these hotspots show no need for the municipal services treatment.

Table 4: Breakdown of municipal services need

Index of need (0–5)	#
0	460
1	655
2	348
3	57
4	13
5	1
Total rated	1,534

Notes: The index of need is defined as the sum of graffiti presence, broken lights, garbage on the street, boarded-up buildings and run-down buildings. Photographers could only reach 1,534/1,919 hotspots due to safety concerns.

We restricted eligibility for the municipal services treatment to hotspots with an index score of 1 or greater, and all hotspots our enumerators could not access. To randomise segments, we assigned eligible hotspots to treatment with a probability of p = 0.25, blocking by station and intensive policing assignment (treated; less than 250m spillover; more than 250m and less than 500m; and greater than 500m from a treated unit). We then randomised these selected hotspots into different batches to roll out the intervention over time. Table 5 displays the distribution of hotspots assigned to municipal services by station. The average *localidad* has 22 hotspots assigned to receive the treatment.

Table 5: Distribution of hotspots assigned to municipal services treatment

	Assigned to treatment		Name	Assigned to treatment	
Name	#	% of all hotspots		#	% of all hotspots
	(1)	(2)		(3)	(4)
Antonio Nariño	15	0.29	Puente Aranda	13	0.29
Barrios Unidos	11	0.17	Rafael Uribe	26	0.26
Bosa	15	0.26	San Cristobal	15	0.31
Candelaria	16	0.27	Santa Fe	23	0.27
Chapinero	35	0.13	Suba	46	0.19
Ciudad Bolivar	33	0.22	Teusaquillo	16	0.25
Engativa	21	0.21	Tunjuelito	15	0.31
Fontibon	21	0.20	Usaquen	21	0.19
Kennedy	43	0.22	Usme	8	0.26
Los Martires	22	0.25			

Notes: The table displays the distribution of hotspots assigned to receive municipal services by *localidad*. The first and third columns represent the total number of hotspots receiving the treatment in each *localidad*, while the second and fourth columns display the percentage of allhotspots that receive the treatment.

We batched the units receiving municipal services into two groups. The first group started receiving treatment on 11 April. We sent photographers to analyse compliance with the intervention from 1 July. After analysing the data, we decided not to move onto the second batch but instead to increase the intensity for the first batch.

In order to measure spatial spillovers and retrieve the direct causal effect of intensive policing and/or municipal services on crime, we differentiate between control units depending on their distance to treated hotspots. Table 6 breaks down how hotspots are distributed in 16 potential outcomes.

Table 6: Distribution of treatment and spillover

Distribution of treatment assignments						
		Broken windows assignment				
		Treated	< 250m	250–500m	> 500m	All
Hotspots policing assignment	Treated	75	196	192	293	756
	< 250m	74	281	185	165	705
	250–500m	32	47	102	113	294
	> 500m	20	22	16	106	164
	All	201	546	495	677	1,919

Notes: The table breaks down our sample of 1,919 hotspots into 16 groups based on treatment assignment and distance to other treated units.

Out of the sample of 1,919 hotspot segments, 756 are assigned to hotspot treatment, 705 are spillover segments within 250m of a treated hotspot, 294 are spillover segments between 250m and 500m of a treated hotspot, and 164 are controls more than 500m from any treated hotspot. Similarly, 201 are treated by the municipal services treatment, 546 are within 250m of a treated hotspot, 495 are between 250m and 500m of a treated hotspot, and 677 are controls more than 500m away from any hotspot receiving the municipal services treatment. 106 units are considered ‘pure control’ in that they are more than 500m away from any hotspot receiving either treatment.

5.4 Data

We have administrative data on crimes, police patrolling time, socio-economic characteristics of all land plots in Bogotá, geo-coded urban infrastructure and location of public surveillance cameras. We complement the administrative data with primary data collection. In the end, we draw on six main sources of data prior to, during, and at the conclusion of the interventions.

5.4.1 Administrative data on police and municipal services compliance

The police shared the full database of GPS patrol locations for all 136,984 streets, 2015–2017. Twenty-three city agencies also shared reports on their diagnosis of each street and compliance with treatment for all streets assigned to the municipal services treatment.

5.4.2 Crime and policing

Police shared data on reported crimes and operations from 2012 to 2017, geolocated to 136,984 streets.

5.4.3 Survey of Bogotá residents[13]

In October 2016, we surveyed 24,000 citizens in 2,399 segments: the 1,919 in the experimental sample, plus a representative sample of 480 segments outside the experimental sample. We interviewed a convenience sample of 10 people per segment, and averaged responses across each segment. As the respondents had to be very familiar with the segment, we limited our sample to individuals who know, live or work in the specific segment. The enumerators found respondents in the segment in the following ways: at their work station, in their homes in the segment, or as they were passing by. Passers-by had to walk through the segment frequently to be able to answer the survey. The enumerators were told to vary the profile of the respondent. The survey collected outcomes such as: perceptions of security risks, perceived incidence of crimes, crimes personally experienced, crime reporting, and trust in and perceived legitimacy of the police and the mayor's office. Figure 3 illustrates the difference between actual and officially reported crimes. We asked whether people had experienced a crime since the beginning of the year, whether they had attempted to report it, and if they were successful. Homicides were reported by the police if individuals did not report them, so administrative data probably captured most murders. However, for all other crimes, about 27 per cent of people said they reported the crime, and an additional 9 per cent said they attempted to report the crime but were unsuccessful. Reporting rates are highest for vehicle theft, because insurance claims require a report.

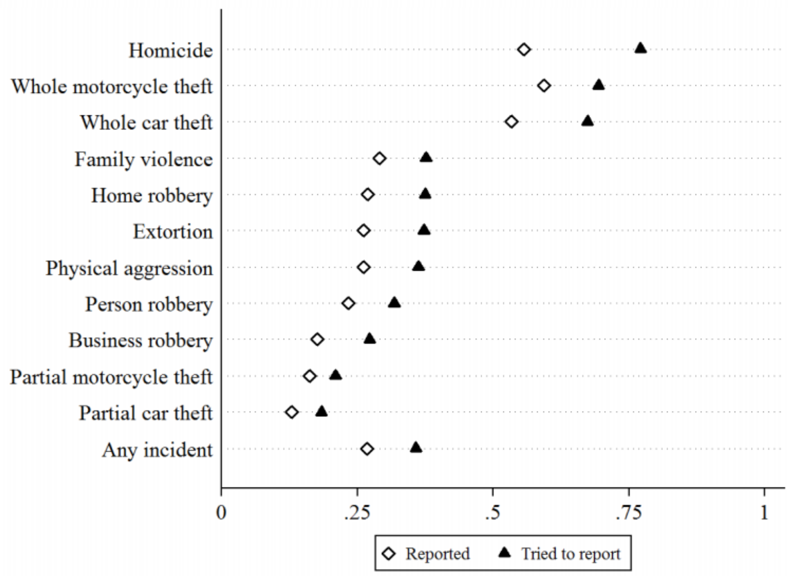
The survey was conducted by Sistemas Especializados de Información (Specialised Information Systems; SEI), who was selected through a merit competition. We invited four survey firms to participate in the competition and SEI won as its proposal had the highest quality (supervisors per enumerators, data entry procedures, etc.). SEI and IPA trained the enumerators and supervisors in a three-day workshop, plus an additional day of training for the supervisors.

Citizens were incentivised to participate in the survey with a weekly raffle of US$15 of phone credit for all who took part.

IPA monitored the survey implementation and put in place quality control measures such as field supervision, high frequency checks, backcheck calls and data entry checks (Online Appendix A).

[13] We created a convenience sample because we had insufficient funds for a representative survey. We are aware that survey data are vulnerable to potential bias, as are administrative data, and so they are useful in cross-checking each other.

Figure 3: Proportion of crime reported, by crime (survey based)



Notes: The figure includes data on all street segments surveyed. Each observation is a survey. The white diamonds denote the proportion of people that effectively reported a crime out of all victims. The black triangles denote the proportion of people that tried to report a crime out of all victims.

5.4.4 Survey of street disorder

To measure levels of street disorder before and after treatment, we sent enumerators hired by IPA to take photographs and rate the presence of graffiti, garbage and boarded-up buildings on a 0–5 scale.

5.4.5 Administrative data on pre-treatment street characteristics

The city also shared data on pre-treatment street characteristics: urban density; income level (high, medium, low); economic use (housing, services, industry); presence of public surveillance cameras; and distance to the closest police station, commercial area, school, religious centre, health centre, transport station, or other public services such as justice.

5.4.6 Qualitative interviews

A PhD student did informal qualitative interviews with dozens of police officers and citizens about their experiences with the intervention and police tactics in general. IPA also hired observers to discreetly visit 100 streets in the experimental sample for a day and passively observe police behaviour. They also interviewed citizens in each segment about police behaviour and attitudes.

5.4.7 Street light survey

Enumerators hired by IPA went to the street segments and rated the quality of the street lights and perceptions about security in the segments.

Our primary outcomes are two insecurity measures (as mentioned in section 2.1): perceived risk and crime incidence. Table 7 reports summary statistics on a standardised index of each outcome for each of the 4 × 5 experimental conditions, using inverse probability weights for assignment into each of the treatment conditions. We created our own survey measures, as existing literature tends to rely on administrative data only.

Table 7: Summary statistics for the primary security outcomes, all experimental conditions

			Municipal services assignment				
			Treated	< 250m	250–500m	> 500m	Ineligible
			(1)	(2)	(3)	(4)	(5)
A: Perceived risk (z-score)							
Intensive policing assignment	Treated	Mean	-0.073	0.430	0.138	-0.013	-0.373
		SD	0.876	1.017	0.864	0.943	0.934
		N	75	154	150	201	174
	< 250m	Mean	0.168	0.335	0.223	0.160	-0.124
		SD	1.061	1.005	0.859	1.369	1.013
		N	74	213	130	125	162
	250–500m	Mean	-0.105	0.291	0.057	0.256	-0.337
		SD	1.042	0.883	0.938	0.942	0.974
		N	32	32	75	80	75
	> 500m	Mean	-0.174	0.320	0.124	-0.218	-0.651
		SD	0.914	1.078	1.042	0.912	0.994
		N	20	14	13	68	49
B: Crime incidence (z-score)							
Intensive policing assignment	Treated	Mean	-0.079	0.379	-0.056	-0.047	-0.179
		SD	0.808	1.010	0.790	0.868	0.877
		N	75	154	150	201	174
	< 250m	Mean	0.157	0.425	0.139	0.169	0.248
		SD	1.032	1.056	0.849	1.769	1.230
		N	74	213	130	125	162
	250–500m	Mean	-0.143	0.207	-0.053	0.096	-0.105
		SD	0.825	1.024	0.889	0.921	0.874
		N	32	32	75	80	75
	> 500m	Mean	-0.215	0.361	-0.147	-0.325	-0.419
		SD	1.092	1.297	1.024	0.745	0.862
		N	20	14	13	68	49

Notes: SD = standard deviation. We report weighted means for each experimental condition, where weights are the inverse of the probability of falling in the corresponding treatment condition. We estimate that probability with repeated simulations of the randomisation procedure. The ineligible condition in column 5 reflects those streets that did not exhibit any disorder at baseline. Technically there are 3 ◊ 4 ineligible conditions for each dependent variable, one for each relative distance from municipal services treated streets, but we pool those columns here for simplicity.

6. Interventions

6.1 Intensive policing

Intensive policing began on 9 February 2016 and ended on 14 October 2016.[14] Intensive policing generally meant a two-thirds increase in police patrol time. As we will see below, during the intervention control streets received roughly 92 minutes of patrol time on

[14] The government, however, did not publicise 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 four to six months. They extended the intervention in part to permit the research team enough time to fund and conduct a survey of citizens.

average, with treated streets receiving an additional 77 minutes – an 84 per cent increase.[15] When we allow for one spillover, control streets received 62 minutes of patrol time on average, while treated streets received an additional 82 minutes and spillover streets less than 250m away received an additional 12 minutes.[16] In order not to overextend patrols, the police required us to assign no more than two hotspots to treatment per quadrant so as not to distort regular duties too much. A 77-minute increase in two hotspots implied that patrol time fell in other segments in the quadrant by roughly one minute each.

Commanders told patrols to visit treatment hotspots at least six times per day for roughly 15 minutes each, mostly during the day, unless near a bar. The police generally did not know what hotspots were in the control group, but in principle they could make reliable guesses. Commanders instructed patrols to continue their normal duties in treated hotspots: 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.

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.

6.2 Municipal services

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 appropriate services. The municipal services intervention began on 11 April 2016 and continued until the end of the intensive policing intervention.

The municipal service and the policing intervention overlapped for six months.

6.3 Compliance

The police patrols and municipal services complied with instructions and treatment assignment. Police did so for the full eight months, while municipal services agencies likely complied for a shorter period. Table 8 reports the effects of assignment to intensive policing or municipal services on various first-stage outcomes.

---

[15] Before the intervention, 1–2 weeks of GPS data suggested that hotspots received at least 38 minutes of patrol time per day. It is doubtful that actual time rose from 38 to 86 minutes. Rather, the 38 minutes was probably an understatement of average patrolling time per hotspot. The police did not have data on pre-intervention patrol times, since the handheld computers with GPS chips were piloted November 2015 through January 2016.

[16] When we allow for the presence of two spillovers, control streets received 64 minutes of patrol time on average. Treated streets received an additional 92 minutes, spillover streets less than 250m away received an additional 21 minutes, and the 250–500m spillover streets region received an additional 6 minutes.

Table 8: 'First-stage' effects of treatment on measures of compliance and effectiveness

	Control mean	ITT assignment to:			
		Intensive policing		Municipal services	
	(1)	(2)	(3)	(4)	(5)
Panel A: Hotspot policing compliance					
Survey: Believes police presence increased in last 6 mon	0.129	0.076	[.011]***	0.017	[.013]
Daily average patrolling time, top-coded and excluding quadrant-days without data	92.001	76.571	[4.424]***	-3.333	[4.371]
# of arrests	0.333	-0.053	[.082]	0.026	[.102]
# of drug seizure cases	0.041	-0.002	[.020]	0.029	[.024]
# of gun seizure cases	0.009	0.006	[.008]	0.007	[.013]
# of recovered car cases	0.003	0.000	[.001]	-0.003	[.001]***
# of recovered motorbike cases	0.006	-0.028	[.019]	0.032	[.027]
Panel B: Municipal services compliance					
Survey: Believes mayor presence increased in last 6 mon	0.144	0.005	[.010]	0.016	[.012]
Compliance measures from city					
Eligible for lights intervention	0.349	-0.007	[.048]	-0.139	[.048]***
Received lights intervention	0.000	-0.010	[.020]	0.199	[.026]***
Eligible for garbage intervention	0.000	0.011	[.025]	0.627	[.032]***
Received garbage intervention	0.000	0.015	[.026]	0.382	[.033]***
June 2016 assessment					
Graffiti on block	0.749	-0.018	[.050]	0.077	[.043]***
Garbage on block	0.251	0.071	[.061]	0.015	[.049]
Broken street light on block	0.000	0.012	[.012]	0.008	[.008]
December 2016 assessment					
Graffiti on block	0.624	0.019	[.053]	0.059	[.047]
Garbage on block	0.245	0.021	[.051]	0.001	[.043]
Broken street light on block	0.029	0.022	[.016]	-0.015	[.017]

Notes: This table reports ITT estimates of the effects of the two interventions, via a WLS 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. 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%, ** significant at the 5%, *** significant at the 1%.

6.3.1 Intensive policing

Calculating the time spent on street segments is difficult because of periodically malfunctioning units or outages. We estimate that control streets received 92 minutes of patrolling time per day, on average. Treated streets received an extra 77 minutes, an 84 per cent increase. Streets outside the experimental sample received an average of 33 minutes of patrolling time per day. Without pre-treatment data on patrol times it is impossible to say whether the increase in patrol time in treatment hotspots came at the expense of control hotspots. What we can say is that the 77-minute rise in two segments means roughly a minute less time in each of the 130 other segments in the quadrant. Some citizens noticed an increase in patrols in the previous six months. In control segments, 13 per cent reported an increase, compared with 21 per cent on treatment segments.

Our compliance analysis was also informed by qualitative interviews with police officers and passive observations to the police while they patrolled the hotspots. In general, we found that police agents were carrying out the expected activities while policing.

We do not see any effect of increased policing on arrests or police actions such as drug seizures, suggesting any effect of the policing may be through deterrence rather than incapacitation (Chalfin and McCrary 2017).

We also do not see any effect of decreased policing when we look at data gathered after the end of the intervention, from November 2016 to the end of June 2017. Although the coefficient points in the same direction as the intervention coefficients, the estimates are not significant.

6.3.2 Municipal services

Table 9 summarises compliance. After assigning 201 segments to municipal services, city agencies diagnosed each one in March 2016. They identified 123 segments needing clean-up services, and 47 needing lighting improvements. They performed these services from June through to August 2016. Tree pruning and graffiti cleaning were one-time treatments; garbage collection 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.

Table 9: Municipal services eligibility and compliance

		City's lighting assessment			% of eligible streets receiving lighting services
		Lights eligible	Lights ineligible	All	
City's cleanliness assessment	Eligible for garbage	21	102	123	41 (87.2%)
	Ineligible for garbage	26	52	78	
	All	47	154	201	
Eligible streets receiving clean-up		74 (60.2%)			

Notes: The table summarises compliance on the municipal services intervention for 201 streets assigned to treatment as reported by the corresponding agencies within the mayor's office.

The impacts were not obvious: 14.4 per cent of survey respondents in control segments noticed an improvement in service delivery in the past six months, and this figure was only 1.9 percentage points greater in treatment streets (not statistically significant, see Table 4). We also visited segments in the 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. We see no effect of treatment in Table 6. 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.

We conducted a survey in March 2017 to assess the quality of street lighting and perceptions of safety. We ran the survey first and foremost in order to understand the weak connection between the municipal services treatment and crime and safety. The survey allowed us to look for two relationships that underlie any reduced-form treatment effect: (1) whether street lighting is associated with reduced crime and perceived safety increases; and (2) whether the treatment improved lighting. If either one of these relationships is weak, then this can explain the absence of a treatment effect.

We find fairly suggestive evidence that the first relationship holds (street lighting is associated with less crime and more safety), though this relationship is not causally identified. But we do not see evidence that the municipal services treatment improved lighting materially. Therefore, the weak treatment effect of municipal services on crime and safety is insignificant in part due to poor implementation.

However, this raises some hope that an ongoing experimental evaluation of lights at scale may show impacts on crime if implementation problems are solved.

This survey's respondents were street light experts that went to the streets at night and answered questions regarding street light quality and safety perceptions. Table 10 summarises the correlation between the number of reported crimes post-intervention and an index of safety perception, with an index of street lighting quality. The index of lighting quality weights answers for different failures such as: whether there are light failures, presence of dark spots, lights that are off, bulbs exhausted, intermittent lights, stolen lights, hanging lights, turned lights, trees needing pruning, sectors without public lights, lights with low luminous flux, and new lights needed. All components range from 0 to 1, as well as the index of street lighting quality, where higher values mean more light or less failures.

The index of safety perception is based on the survey mentioned above, conducted in March 2017. The index averages the responses of safety perception for the following: using a cell phone, walking at 8pm, walking at midnight, a woman walking alone at night, a man walking alone at night, and a family walking at night. The responses were calculated on a scale from 1 to 4, where 4 is 'feels very safe' and 1 is 'feels very unsafe'. For both cases – the administrative crime and the index of safety – we controlled for being in the hotspots experimental sample from February to October 2016, for being assigned to receive municipal services, and for being assigned to receive hotspot intensive policing in the same period.

Table 10: Correlation between administrative crime and lighting quality

	# of crimes reported (admin), Nov 2016 – Jun 2017 (1)	Index of safety perception, Mar 2017 (2)
Index of lighting quality, 0–1 (+ more light), Mar 2017	-0.437	1.040***
	(0.305)	(0.119)
Being in the hotspots experimental sample, Feb–Oct 2016	0.153	-0.036
	(0.262)	(0.093)
Assigned to receive municipal services, Feb–Oct 2016	-0.009	0.021
	(0.571)	(0.188)
Assigned to receive hotspot policing, Feb–Oct 2016	0.007	-0.026
	(0.305)	(0.117)

Note: Police station (block) fixed effects and baseline covariates included. The index of lighting ranges from 0 to 1 where higher values mean more light. The index of safety perception is a z-score where higher values mean more safety. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

As expected, there is an inverse relationship between lighting and crime, although it is not significant. In the case of safety perception, there is a positive and significant relationship with lighting. Both results point in the same direction: the better the lighting, the better the security; however, since the coefficients for being in the hotspots experimental sample were not significant, this suggests that the municipal services intervention probably did not improve lighting in a sustained way. This can also be seen in Table 11, which shows the average treatment effect on the index of lighting quality. The effect of the municipal services intervention has a positive sign; however, it is not significant.

Table *11*: Average treatment effect on the index of lighting quality

	# of crimes reported (admin), Nov 2016 – Jun 2017 (1)	Index of safety perception, Mar 2017 (2)
Index of lighting quality, 0–1 (+ more light), Mar 2017	-0.437	1.040***
	(0.305)	(0.119)
Being in the hotspots experimental sample, Feb–Oct 2016	0.153	-0.036
	(0.262)	(0.093)
Assigned to receive municipal services, Feb–Oct 2016	-0.009	0.021
	(0.571)	(0.188)
Assigned to receive hotspot policing, Feb–Oct 2016	0.007	-0.026
	(0.305)	(0.117)

Note: Police station (block) fixed effects and baseline covariates included. The index of lighting ranges from 0 to 1 where higher values mean more light. The index of safety perception is a z-score where higher values mean more safety. Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

7. Impact analysis and results of the key evaluation questions

7.1 Estimation

We estimate treatment and spillover effects within the experimental sample using the following WLS regression:

$$Y_{sqp} = \beta_1 P_{sqp} + \beta_2 M_{sqp} + \beta_3 (P \times M)_{sqp} + \lambda_1 S^p_{sqp} \quad + \lambda_2 S^M_{sqp} \quad + \lambda_3 (S^P \times S^M)_{sqp} + \gamma_p + \theta X_{sqp} + \epsilon_{sqp}$$

using inverse probability weighting (IPW) for assignment to the conditions S^P and S^M. Thus, β_1 and β_2 estimate the marginal ITT effects of each treatment alone and β_3 estimates the marginal effect of receiving both. A negative sign on β_3 implies increasing returns. The effect of receiving both interventions is the sum β_1 + β_2 + β_3. Likewise, λ and λ^N estimate spillover effects of each treatment in each sample. To see the marginal effects of each treatment, we can perform the estimation under the constraints that β_3 = 0 and λ_3 = 0. These constraints are useful when we expect no interaction, such as the analysis of treatment compliance.

To calculate spillovers in non-hotspots we estimate:

$$Y_{sqp} = \lambda_1{}^N S_{sqp}{}^P + \lambda_2{}^N S_{sqp}{}^M + \lambda_3 (S^P \times S^M)_{sqp} + \gamma_p{}^N + \theta^N X_{sqp} + \epsilon_{sqp}{}^N$$

using IPW for assignment to the conditions S^P and S^M. Thus, β_1 and β_2 estimate the marginal ITT effects of each treatment alone and β_3 estimates the marginal effect of receiving both. A negative sign on β_3 implies increasing returns. The effect of receiving both interventions is the sum $\beta_1 + \beta_2 + \beta_3$. Likewise, λ and λ^N estimate spillover effects of each treatment in each sample. To see the marginal effects of each treatment, we can perform the estimation under the constraints that $\beta_3 = 0$ and $\lambda_3 = 0$. These constraints are useful when we expect no interaction, such as the analysis of treatment compliance.

7.1.1 Estimating spillovers using an exponential rate of decay

As an alternative to the above, we can estimate a continuous, monotonic spatial decay function with the following ordinary least squares regression:

$$Y_{sqp} = \beta_1^D P_{sqp} + \beta_2^D M_{sqp} + \lambda_1^D \sum_{t \in T_P} f(d_{sqp,t}) + \lambda_2^D \sum_{t \in T_M} f(d_{sqp,t}) + \gamma_p^D + \theta^D X_{sqp} + \epsilon_{sqp}^D$$

where $f(d_{sqp,t})$ is a spatial decay function with a standardised distribution. It is a weighted sum of distances to all treated hotspots, where *t* enumerates treated hotspots and *T* is the set of all treated hotspots. Treated segments receive no spillover from themselves but can receive spillovers from other treated segments. When applied to the non-experimental sample, the regression omits direct treatment effects. Our default functional form is exponential, $f(d_{sqp,t}) = 1/e^{d_{sqp,t}}$, but we examine alternatives. We can no longer employ IPW to weigh street segments because the exposure measures are continuous variables. Instead, we include in the control vector the expected spillover intensities (averaged across 1,000 simulations) and the probabilities of being treated by each intervention. We calculate standard errors using randomisation inference.

7.2 Inverse probability weighting (IPW)

Spillovers introduce spuriousness that can be corrected with IPW. Our randomisation procedure gives segments variable probabilities of being in each of the assigned conditions (treatment, spillover, pure control). This is especially true for segments in our non-experimental sample. For example, non-experimental segments in relatively safer areas of Bogotá have a 0 per cent chance of being a spillover for either treatment, since there are no experimental units in those neighbourhoods. In areas with lots of crime, non-experimental units have a higher probability of being a less-than-250m spillover because they are located in areas with more hotspots (experimental units). In areas like the south of Bogotá, however, many segments have a zero probability of being a less-than-250m spillover because there are no hotspots present. Thus, hotspots close to other hotspots such as those in the city centre or other dense areas will be assigned to the spillover condition in most randomisations. These streets may have unobservable characteristics that are associated with high levels of crime. This could mechanically lead us to conclude that spillovers increase crime. Thus, a simple spillover versus a control comparison will lead to biased estimates on the effect of crime because the outcome (crime) is correlated with treatment assignment.

Controlling for baseline characteristics and crime histories reduces, but does not eliminate, potential bias. With IPW, outcomes for the segments assigned to any given condition are weighted by the inverse of the probability of assignment to that condition. These weights ensure that all segments have the same probability (after weighting) of being exposed to spillovers. In other words, IPW adds greater weight to the segments that have a lower probability of being assigned to spillover (as they are far away from hotspots), and adds a lower weight to the segments that have a higher probability of being assigned to spillovers due to their proximity to hotspots.

7.3 Balance tests and summary statistics

Table 12 reports summary statistics and balance tests. In October 2016, the police updated all 2012–2016 crime data with more accurate GPS coordinates and additional crime categories, and we report original and updated data. Between 0 and 82 crimes were reported in hotspots in the previous four years (461 with the updated data, as we had information on more crime types), with an average of five crimes per hotspot. More than half were property crimes, but violent crimes such as murders and assaults were also significant. Overall, 95 per cent of hotspots had relatively low levels of physical disorder such as garbage.

Table 12: Descriptive statistics for the experimental sample and test balance (treatment vs all control streets, including potential spillover streets)

					WLS test of balance			
	Sample summary statistics				Intensive policing		Municipal services	
	Mean	SD	Min	Max	Coeff.	p-val	Coeff.	p-val
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Baseline crime (not top-coded)								
# of reported crimes on street, 2012–2015 (original)	4.53	5.72	0	82	-0.17	0.62	-0.13	0.70
# of violent crimes	1.88	2.94	0	56	-0.18	0.21	-0.05	0.75
# of property crimes	2.66	3.97	0	50	0.02	0.95	-0.08	0.76
# of reported crimes on street, 2012–2015 (updated)	5.18	18.24	0	461	-0.21	0.86	-0.36	0.79
# of violent crimes	1.40	5.38	0	78	0.39	0.38	0.22	0.68
# of property crimes	3.78	14.09	0	407	-0.60	0.45	-0.58	0.52
Average # of reported crimes per segment in quadrant, 2012–2015	3.56	5.13	0	61	-0.30	0.50	0.38	0.49
Average daily patrolling time (11/2015), minutes	38.03	70.27	1	1,029	-1.77	0.73	3.42	0.57
Rating of baseline disorder (0–5, + more disorder)	1.18	0.74	0	5	-0.05	0.31	0.35	0.00
Eligible for municipal services	0.86	0.35	0	1	-0.02	0.27	0.22	0.00
Metres from police infrastructure	551.37	351.46	6	2,805	-26.18	0.26	-11.95	0.64
Zoned for industry/commerce	0.38	0.49	0	1	-0.09	0.01	0.05	0.16
Zoned for service sector	0.13	0.34	0	1	0.02	0.33	0.03	0.25
High-income street segment	0.07	0.25	0	1	0.00	0.79	-0.01	0.54
Medium-income street segment	0.55	0.50	0	1	-0.06	0.06	0.00	0.98
# of segments in quadrant	127.21	86.99	2	672	2.05	0.71	-3.04	0.57
# of experimental units in quadrant	3.67	2.68	1	14	-0.30	0.08	-0.16	0.31
# of HSP treated units in quadrant	1.15	0.95	0	3	1.35	0.00	-0.01	0.91
# of MS treated units in quadrant	0.66	0.69	0	3	-0.08	0.06	0.91	0.00
Intensive policing assignment: treated	0.48	0.50	0	1	1.00	-	0.00	-
Intensive policing assignment: proximal spillover	0.29	0.46	0	1	-0.56	0.00	0.01	0.83
Intensive policing assignment: distant spillover	0.14	0.35	0	1	-0.28	0.00	0.00	0.96
Intensive policing assignment: pure control	0.09	0.28	0	1	-0.17	0.00	-0.01	0.72
Municipal services assignment: treated	0.41	0.49	0	1	0.00	-	1.00	-

	Sample summary statistics				WLS test of balance: Intensive policing		WLS test of balance: Municipal services	
	Mean	SD	Min	Max	Coeff.	p-val	Coeff.	p-val
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Municipal services assignment: proximal spillover	0.19	0.39	0	1	0.05	0.01	-0.31	0.00
Municipal services assignment: distant spillover	0.17	0.37	0	1	-0.01	0.71	-0.28	0.00
Municipal services assignment: pure control	0.23	0.42	0	1	-0.04	0.03	-0.40	0.00

Notes: HSP = hotspot policing. MS = municipal services. Columns 1–4 display the summary statistics for our sample of 1,919 hotspots, weighted by the probability of being in the observed experimental condition. In columns 5–8, we perform a balance test for treated versus all control units using WLS.

7.4 Programme impacts on officially reported crime

Table 13 reports estimations of direct treatment, experimental spillover, and non-experimental spillover coefficients, with and without the interaction terms between intensive policing and municipal services. Following our pre-specified rule, the table reports spillovers within 250m only. We do not see statistically significant spillovers in the 250–500m region.

Table 8 also calculates the total number of deterred crimes as the product of the estimated coefficients and the number of treatment and spillover segments in the city. We omit the 57,695 streets with zero probability of assignment to the spillover condition. There are 51,390 non-hotspots and 705 control hotspots for the policing intervention and 20,740 non-hotspots and 546 control hotspots for municipal services. Thus, even small estimated spillovers can have a large effect on total crime estimates. Since our coefficients are fairly uncertain, we must interpret aggregate impacts with caution.

Our best approximation regarding the overall impact on crime is that the interventions directly deter a relatively modest amount of crime, and that some or all of this crime is displaced to neighbouring streets. However, in our main specification, crime displacement is concentrated in property crime. Violent crimes may not be displaced as easily.

Table 13: Estimated aggregate impacts of the interventions, accounting for spillovers within < 250m

	Dependent variable: # of crimes reported to police on segment (administrative data)							
	No interaction between treatments				Interaction between treatments			
	Coeff.	RI p-value	# segments	Estimated total impact = (1) X (3)	Coeff.	RI p-value	# segments	Estimated total impact = (1) X (3)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Direct treatment effect								
Intensive policing	-0.094	0.512	756	-70.7	0.009	0.817	756	7.1
Municipal services	-0.076	0.783	201	-15.2	0.089	0.367	201	17.9
Both					-0.437	0.043	75	-32.8
Subtotal				-86.0				-7.8
B. Spillover, experimental sample								
Intensive policing	0.061	0.595	705	42.7	0.143	0.315	705	100.7
Municipal services	0.176	0.056	546	96.3	0.255	0.025	546	139.1
Both					-0.272	0.196	281	-76.5
Subtotal				138.9				163.3
C. Spillover, non-experimental sample								
Intensive policing	0.016	0.101	51,390	844.7	0.013	0.205	51,390	677.4
Municipal services	-0.003	0.394	20,740	-65.8	-0.006	0.484	20,740	-124.9
Both					0.005	0.973	15,491	85.0
Subtotal				778.9				637.5
				831.9				793.0
			95% CI	(-780, 2054)			95% CI	(-1001, 2199)
			90% CI	(-434, 1848)			90% CI	(-689, 1989)

Notes: RI = randomisation inference; CI = confidence interval. Columns 1–4 refer to the non-interacted results (equation 1 under the constraint that —3 = 0 and ⁄3 = 0) while columns 5–8 refer to the interacted results (equation 1 with no constraints). Columns 1 and 5 display the bias-adjusted treatment effect while columns 2 and 6 display RI p-values. Columns 3 and 7 display the number of units in each group. Columns 4 and 8 display 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 randomisation inference. First, we create a fake schedule of potential outcomes for each observation by adding or subtracting RI-adjusted treatment or spillover effects. This process gives us a potential outcome for each unit depending on its treatment assignment. Second, we simulate a randomisation and take the potential outcome associated with the treatment assignment of the new randomisation. Third, we estimate treatment and spillover effects using this new outcome and apply the RI bias adjustment from our main set of results. Fourth, we multiply these bias-adjusted treatment effects by the number of segments in each group, and sum across both the experimental and non-experimental samples to get the aggregate effect. We repeat steps two to four 1,000 times to get the distribution of the test statistic, which is roughly centred on the actual number of deterred crimes. The 2.5 and 97.5 percentiles of this distribution give us the 95 per cent confidence interval.

7.5 Direct treatment effects

Starting with columns 1–4 of Table 13 (no interaction), both intensive policing and municipal services reduce officially reported crimes on average, although these coefficients are not statistically significant. Control segments report an average of 0.743 crimes over the intervention period (column 1 in Table 15). Thus, the coefficient on intensive policing of -0.094 represents a 12.6 per cent improvement. The municipal services coefficient is about two thirds as large. In total, these estimates suggest that the reallocation of police and municipal services deterred 86 crimes in targeted streets over the intervention period (not statistically significant).

Turning to columns 5–8, we see larger and most statistically significant impacts of state presence in the segments that were assigned to both interventions. The coefficients on policing and municipal services are positive but imprecise. We see no evidence that either intervention on its own reduced crime. The coefficient on the interaction is -0.437; however, it has a randomised inference p-value of 0.043. The sum of the three coefficients is -0.339 with a p-value of 0.110 (column 5 in Table 15). This sum corresponds to a 45.6 per cent decrease in reported crimes on the 75 streets that received both interventions. The fact that the coefficient on the interaction is large, negative and statistically significant implies that there may be increasing returns to security investments, at least over this range of variation. Of course, given that the sum of effects is weakly statistically significant, we cannot say with confidence that both interventions reduced crime on these 75 streets. Moreover, the aggregate direct effect of the programme looks even smaller when we account for the interaction. According to these estimates, our best estimate is that only eight crimes were deterred directly by both interventions, which equates to approximately one per month during the policing and municipal services interventions.

7.6 Spillover effects

Meanwhile, the spillover coefficients suggest that any crime deterred is more than made up for by a rise in crime in streets within 250m. With regard to intensive policing, all four spillover coefficients are positive. The spillover effects in the experimental sample are imprecise, but given the large number of nearby non-hotspots, spillovers in the non-experimental sample suggest a positive effect, albeit one that falls short of conventional levels of significance (even jointly). The sufficiently large number of non-hotspot segments results in these small coefficients adding up to high levels of crime: 841 crimes in aggregate when we do not allow for the interaction, and 654 when we do. In contrast, we see no evidence that municipal services pushed crime around the corner. The coefficients on spillovers in the non-experimental sample are actually negative, although they are imprecise. In aggregate, however, this estimate adds up to between 56 and 121 crimes deterred in nearby streets, depending on the specification.

7.7 Aggregate effects

We use these estimates to approximate the aggregate effect on crime. It is unlikely that reallocating police and municipal services reduced total crime in the city. On the contrary, the estimates suggest that crimes increased by about 800 in both specifications (2% relative to the total number of reported crimes). This must be interpreted with caution, however, for two reasons. First, neither aggregate effect is statistically significant even at the 10 per cent level. Second, this estimate would not capture general equilibrium effects if they exist (e.g. if the intervention is disrupting city-wide criminal networks). These estimates suggest that we can rule out the possibility that crime decreased in the city by even a modest amount.

7.8 Heterogeneity by type of crime and by initial level of crime

Police prioritise violent crimes over property crimes. Table 14 disaggregates the impacts on total crime into violent and property crimes. Our best approximation is that aggregate violent crime fell by 174 to 411 crimes in total (1% to 3% relative to the total number of violent crimes), depending on whether we use the interaction or not, although neither estimate is statistically significant. Property crimes rose by 1,006 to 1,204 in aggregate (4% to 5% relative to the total number of property crimes), however, and these estimates are statistically significant at the 10 per cent level when we include the interaction. The two most socially costly crimes, homicides and sexual assaults, fell by 61 to 97 crimes (5% to 8% relative to the total number of homicides and sexual assaults). This difference in property and violent crimes is statistically significant. We interpret the different results for violent and property crimes with caution, however, since the aggregate effects change once we introduce minor changes in specification. We estimate an alternative version of equation 1 with different dummies for streets located within 250m and between 250m and 500m from treatment hotspots. In this case, we now observe crime displacement also for violent crime.

Table 14: Aggregate impacts on crime by type (mean and confidence intervals)

	No interaction between treatments			Interaction between treatments		
	Effect (1)	95% CI (2)	90% CI (3)	Effect (4)	95% CI (5)	90% CI (6)
All crime	831.9	(-780, 2054)	(-434, 1848)	793.0	(-1001, 2199)	(-689, 1989)
Property crime	1006.5	(-206, 2029)	(-41, 1873)	1204.1	(-338, 2370)	(27, 2203)
Violent crime	-174.7	(-858, 378)	(-744, 296)	-411.1	(-1112, 204)	(-997, 82)
Homicides and sexual assaults only	-60.8	(-179, 55)	(-165, 40)	-97.6	(-233, 33)	(-210, 15)
Property – violent crime	1181.2			1615.2		
p-value	0.063			0.000		

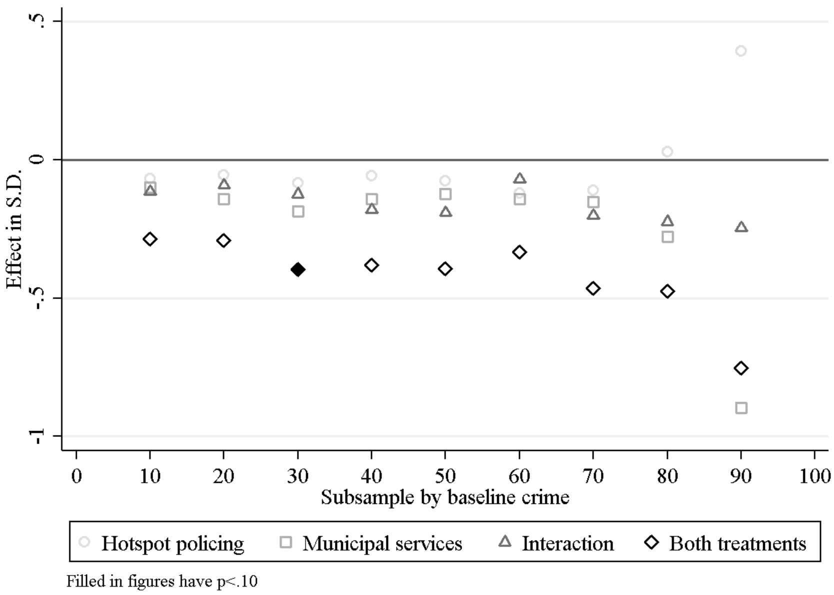
Notes: This table presents the aggregate effect calculation for various crime subgroups assuming spillovers within 250m. Calculations are based on the aggregate effect and confidence interval described in Table 6.

We pre-specified one major form of heterogeneity analysis by baseline levels of crime. In part, this helps us to imagine an experiment where we targeted a much smaller number of the hottest hotspots.

Broadly, we observed what we predicted: that improvements in security are greater in the higher-crime streets. Figure 4 reports the results of estimating treatment effects on the n per cent highest-crime hotspots. The treatment effect is fairly constant up until the point we reach the street segments in the 70th percentile and above, when the impact of receiving both interventions climbs first to 0.5 standard deviations and then to about 0.75 standard deviations. The effect is imprecise, as the sample size drops dramatically. These results are consistent with increasing returns to treating the least secure hotspots.

Note, however, that in the two highest-crime deciles, intensive policing alone is not associated with decreases in crime. Any decrease is driven by municipal services (itself not statistically significant) or the combination of both (imprecise though the estimates may be). Thus, our results are not diluted by the inclusion of less hot hotspots. We explore this further in Online Appendix B, where we estimate the effects of one additional hour of patrolling time over different levels of baseline crime. The results suggest that the marginal effects of intensive policing on security decrease as the baseline crime levels increase. This is consistent with results from the pre-specified heterogeneity analysis, and suggests that the larger effects in higher-crime hotspots are due to municipal services and the interaction of both treatments rather than by a larger effect of intensive policing.

Figure 4: Heterogeneity of security impacts by pre-treatment administrative crime levels



7.9 Programme impacts on insecurity

Table 15 reports impacts on our main security measures: the perceived risk index, based on surveys; and the index of crime, which averages survey-reported and officially reported crime. Treatment effects can be interpreted as average standard deviation changes in the outcome. The table also reports treatment effects on components of the crime index. Our focus is centred on the two pre-specified indices, but we also report results for an equally weighted average of both. Table 15 reports direct treatment effects and spillover effects on hotspots within 250m. Below the estimated coefficient, the table reports the randomisation inference p-value instead of either the standard error-based p-value or the standard error as previously explained.

The survey data tell a similar story to police data. We see the largest and most statistically significant impacts of state presence in the segments that received both interventions. Those 75 segments reported a 0.327 standard deviation decrease in overall insecurity, significant at the 10 per cent level (column 5). The coefficients on perceived risk and crime indices are similar, although only the perceived risk index is statistically significant alone.

Alone, the interventions are associated with improvements in security, but none of the estimates are individually significant. Nonetheless, the coefficients all point in the direction of better security: intensive policing alone reduces perceived risk by 0.12 standard deviations, crime by 0.06 and overall insecurity by 0.11 (column 2); municipal services alone reduces perceived risk by 0.09 standard deviations, crime by 0.08 and overall insecurity by 0.10 (column 3). The coefficient on the interaction term (column 4) is statistically significant for officially reported crimes only. We take this result as suggestive of increasing returns to state presence.

7.9.1 Spillovers

There is also evidence of crime displacement to control hotspots in columns 6–9 of Table 15. Intensive policing alone and municipal services alone are associated with increases in crimes in nearby hotspots of 0 to 0.26 standard deviations. Only the municipal services impacts are statistically significant, with a 0.15 standard deviation increase in insecurity. The interaction terms are generally negative (column 8) and generally statistically significant, such that there is generally no evidence of spillovers into hotspots near other hotspots that received both intensive policing and municipal services.

Table 15: Programme impacts on security in the experimental sample, accounting for spillovers within 250m, with p-values from randomisation inference (N = 1,916)

		ITT of assignment to:				Impact of proximal spillover:			
	Control mean (1)	Hotspot policing (HSP) (2)	Municipal services (MS) (3)	Interaction effect (IE) (4)	HSP + MS + IE (5)	Hotspot policing (HSP) (6)	Municipal services (MS) (7)	Interaction effect (IE) (8)	HSP + MS + IE (9)
Insecurity index, z-score (+ more insecure)	-0.003	-0.106	-0.100	-0.121	-0.327	0.045	0.150	-0.217	-0.022
		0.391	*0.536*	*0.447*	*0.095*	*0.322*	*0.020*	*0.039*	*0.577*
Perceived risk index, z-score (+ riskier)	0.049	-0.122	-0.086	-0.084	-0.292	0.002	0.083	-0.160	-0.075
		0.259	*0.494*	*0.644*	*0.094*	*0.511*	*0.129*	*0.085*	*0.808*
Crime index, z-score (+ more crime)	-0.054	-0.054	-0.080	-0.118	-0.252	0.073	0.166	-0.200	0.039
		0.701	*0.659*	*0.412*	*0.196*	*0.231*	*0.010*	*0.059*	*0.361*
Perceived & actual incidence of crime, z-score (survey)	0.059	-0.081	-0.158	0.066	-0.173	0.027	0.099	-0.137	-0.011
		0.514	*0.153*	*0.423*	*0.507*	*0.417*	*0.092*	*0.171*	*0.578*
# crimes reported to police on street segment (admin)	0.743	0.009	0.089	-0.437	-0.339	0.143	0.255	-0.272	0.125
		0.817	*0.367*	*0.043*	*0.110*	*0.315*	*0.025*	*0.196*	*0.289*

Notes: p-values generated via randomisation inference are in italics. This table reports ITT estimates of equation 1, estimating the direct effects of the two interventions (Columns 2 to 4) and the spillover effects (Columns 6 to 8) via a WLS regression of each outcome on treatment indicators, spillover indicators, police station (block) fixed effects, and baseline covariates. Columns 5 and 9 report the sum of the three preceding coefficients. The measures of perceived risk, perceived incidence of crime and proportion reporting crime come from our citizen survey, and the number of crimes reported to the police come from police administrative data.

7.9.2 Estimating spillovers with an exponential decay function

We also consider an exponential rate of decay rather than our fixed radii, which calculates spillovers in non-hotspots using equation 2.[17] The coefficients represent the expected increase in crimes as a segment moves a standard deviation closer to a treated hotspot (Table 16). The signs on the policing coefficients are all positive but not statistically significant, as is consistent with the analysis above.

One difference is that the evidence of displacement is no longer confined to property crimes. Here, the majority of displacement seems to be associated with violent crimes. The signs on municipal services, meanwhile, are negative, implying a diffusion of benefits to nearby streets. The decrease is roughly significant at the 10 per cent level for all crimes, and roughly significant at the 5 per cent level for violent crimes alone. These signs are consistent across most functional forms, although the statistical significance is not.

Table 16: Programme impacts on security in the experimental sample using exponential decay function

		Impact of a 1 standard deviation change in the average exponential distance to a hotspot treated with:	
	Control mean	Intensive policing	Municipal services
	(1)	(2)	(3)
# crimes reported to police on street segment	0.274	0.049	-0.050
		0.280	*0.086*
# property crimes only	0.100	0.004	0.001
		0.798	*0.963*
# violent crimes only	0.174	0.045	-0.051
		0.288	*0.044*

Notes: Randomisation inference p-values are in italics. This table estimates the coefficients on spillovers using equation 2 above. We estimate the regression on the non-experimental sample of non-hotspots alone. The weighted distance measures have been standardised to have zero mean and unit standard deviation.

[17] This functional form places some of the heaviest weight on immediately proximate streets. Linear, logarithmic and inverse square decay functions produce qualitatively similar conclusions, even though they give more weight to more distant segments. We ignore interactions between treatments for simplicity, as they yield similar results.

7.10 Programme impacts on state trust and legitimacy

Table 17: Impacts on state legitimacy allowing spillover within 250m, with randomised inference p-values

		ITT of assignment to:				Impact of proximal spillover			
	Control mean (1)	Hotspot policing (HSP) (2)	Municipal services (MS) (3)	Interaction effect (IE) (4)	HSP + MS + IE (5)	Hotspot policing (HSP) (6)	Municipal services (MS) (7)	Interaction effect (IE) (8)	HSP + MS + IE (9)
Opinion of police, z-score (+ better opinion)	0.024	0.143	0.210	-0.308	0.045	-0.024	0.043	0.123	0.141
		0.150	*0.107*	*0.017*	*0.867*	*0.591*	*0.817*	*0.338*	*0.797*
Opinion of mayor, z-score (+ better opinion)	-0.014	0.001	0.179	-0.414	-0.234	-0.024	0.068	-0.025	0.020
		0.912	*0.078*	*0.003*	*0.008*	*0.523*	*0.982*	*0.919*	*0.668*
Likeliness to report crime (0–3, + higher likelihood)	2.046	0.004	0.021	0.035	0.060	-0.007	0.007	0.026	0.026
		0.921	*0.802*	*0.522*	*0.385*	*0.688*	*0.991*	*0.637*	*0.837*

Notes: p-values generated via randomisation inference are in italics. This table reports ITT estimates of equation 1, estimating the direct effects of the two interventions (columns 2 to 4) and the spillover effects (columns 6 to 8) via a WLS regression of each outcome on treatment indicators, spillover indicators, police station (block) fixed effects and baseline covariates. Columns 5 and 9 report the sum of the three preceding coefficients. The three measures come from our citizen survey.

We pre-specified three secondary outcomes capturing impacts on trust in and legitimacy of the state. The first was an opinion of the police index, averaging four attitudes towards police: trust, quality of work, overall satisfaction and likelihood that they would give information to police. The second was an opinion of the mayor index, which asked the same four questions regarding the city government. The third was a crime reporting measure, which captured the likelihood that people would report a crime to the police. This helps us understand whether administrative crime reporting changes with treatment and is also a measure of collaboration and resultant legitimacy.[18]

Overall, we see little evidence that the interventions increased trust in or legitimacy of the state. Table 17 reports ITT effects using equation 1 and the randomisation inference p-values. We see an unexpected pattern: intensive policing and municipal services alone are associated with an increase in positive opinions of the police and the mayor; however, this is effectively cancelled out when both treatments are received. This pattern is statistically significant when we ignore spillovers, but less robust when accounting for spillovers. This heterogeneity across arms is hard to interpret and could reflect noise. In analysis ignoring any interactions (not shown), intensive policing and municipal services are associated with little change in opinions of police, and a slightly negative effect on opinions about the mayor: a 0.13 standard deviation fall, significant at the 10 per cent level.

7.11 Cost–benefit considerations

Cost-effectiveness in this case is in the eye of the beholder. The city sees the interventions as having little or no marginal cost, since they simply reallocated existing resources from some streets to others without raising their budgets or personnel. Therefore, the main question is whether a high likelihood of reducing roughly 100 murders and rapes (8% relative to the total number of cases) is worth a rise in property crime. This is a trade-off that many police chiefs and mayors might reasonably make.[19]

On the other hand, reallocating street-level bureaucrats had real costs. There was a logistical cost of coordinating patrols, especially management time. It also made police patrols spend more time in unpleasant places. Officers told us they disliked the loss of autonomy and flexibility. There are also opportunity costs to consider. Intensive policing was a major reform, and like any bureaucracy, the police can only undertake so many reforms in a year. The mayor's office used scarce social and political capital to implement it. We believe one should measure this reform against the others it supplanted.

[18] We also have survey data on 399 non-experimental street segments, and Online Appendix H estimates these non-experimental spillovers within 250 metres. This sample is generally too small to estimate non-experimental spillovers precisely, but the patterns are generally consistent with what we see in the large-sample dataset on reported crimes. In particular, the coefficients on intensive policing are positive.

[19] Indeed, we ran the aggregate effects reported in Table 9 using a weighted crime index as outcome instead of the simple sum (not reported). The weights are the average prison sentence in the Colombian penal code for each crime. For instance, the weight for one homicide is about 13 times that for a shoplifting case. In such a case, the aggregate effects are negative for the interacted version, although imprecise.

8. Discussion

Intensive policing is probably the most common security tactic in the world. We evaluate it on an unprecedented scale in Bogotá. Unexpectedly, we do not see evidence in this Latin American capital of large or statistically significant impacts of doubling of policing time on the top 2 per cent highest-crime streets. The evidence also suggests that any crimes deterred may simply be displaced to nearby hotspots and non-hotspots, although with the caveat that these subtle spillovers are hard to estimate precisely even in a sample this large. Our most robust finding is that a combination of both policing and municipal services was most effective at deterring crime. We cannot reject the possibility that these crimes were displaced to nearby streets but, as we discuss below, this and other results point to lessons for future place-based interventions.

Our study is also a good example of a policy evaluation where the implications hinge on how to interpret estimates and significance levels under uncertainty. Despite our large sample, confidence intervals are wide, especially on the aggregate effects. In our main specification, for instance, state presence seems to have reduced the most serious violent crimes city-wide. This result, however, is sensitive to specification. This points to the importance of statistical power in future security evaluations.

8.1 How do our results line up with the US evidence?

This experiment provides some of the first experimental evidence on place-based crime interventions outside the US.[20] At first glance, it might seem that the displacement of total crime to nearby streets runs against US literature. We must compare with caution, as Bogotá and the US are different contexts. Policing interventions also take different forms, and vary in terms of intensity, concentration, crimes targeted, duration and quality of approach. That said, on close inspection, our results are not so different.[21] The previous literature has not ruled out positive or negative spillovers in a definitive way. These studies are split on whether they observe displacement of crime or diffusion of benefits on average. Moreover, most prior studies' sample sizes are so small that the confidence intervals on spillovers include sizable displacement effects.[22] Perhaps the biggest lesson for place-based crime studies is that small sample sizes will simply not help to answer the crucial question of spillovers.

[20] Two ongoing projects in Latin America and the Caribbean are Collazos et al. (2017) in Medellín, Colombia, and Sherman et al. (2014) in Trinidad and Tobago. Compared with the Medellín study, we find generally different results. We observe direct treatment effects on both property and violent crimes, while they only find evidence of a decrease on car thefts. We observe displacement mainly on property crimes, and they find a decrease in car thefts in nearby targeted hotspots. The context is radically different regarding both criminal behaviour and implementation capability, and we believe this could be driving the differences. For instance, Medellín has about 60 per cent more police than Bogotá in relative terms.

[21] Most previous studies use only post-intervention data to conduct the evaluation. We follow a similar approach (not reported) and find no evidence of an enduring deterrent effect. If anything, the (equivocal) evidence points in the opposite direction.

[22] This can be difficult to judge, however, since several studies do not report standard errors or confidence intervals. Given that sample sizes are often under 100 or even under 30, it seems reasonable to assume that the confidence intervals include displacement effects.

8.2 Methodological lessons

We believe one of the major contributions of this study, apart from adding empirical evidence to citizen security interventions, is the methodological lesson for future policy experiments in dense networks of streets or people. When small spillovers matter, anything that could bias spillover effects or make them less precise matters a great deal. This points to the importance of eliminating these biases and having accurate, efficient estimates. Failure to account for the biases arising from spillover estimation will have profound effects on our conclusions, whether it is bias correction through IPW and re-centring, or randomisation inference for calculating exact p-values.

Randomisation inference has yet to gain currency in randomised trials, in part because it tends to provide generally the same conclusion as the usual clustered standard errors. A textbook case for randomisation inference, however, is design-based estimation of spillovers where units have widely different probabilities of assignment to different experimental conditions. This problem extends to any other situation in which the structure of the clustering of experimental units in a given treatment condition is difficult to model, which is a condition that is prevalent in dense networks with a high chance of outcomes or even treatments spilling over to close units.

Flexibility in measuring spillovers is also crucial, and we illustrate how this can be a design-based choice, regardless of the inference method used. In Bogotá, we find evidence of spillovers in a catchment area considerably wider than the norm, which if true could mean that the aggregate effect of displacement is considerably greater. Continuous rates of decay impose a fair degree of structure on the nature of the spillover, which is fine if that structure is well understood.

9. Specific findings for policy and practice

9.1 Lessons for place-based security interventions

We can group policy changes into two categories. The first comprises changes that make place-based security, especially policing, more effective on directly treated streets. The second comprises changes that reduce the chances of adverse spillovers.

Our results suggest that there may be increasing returns to state presence, especially the combination of policing and municipal services. This combination deserves to be tested at scale. The evidence also suggests that focusing on the higher-crime hotspots could have larger proportional impacts. Qualitatively, our interactions with the government and police patrols suggest other ways to increase direct impacts. One is less predictable policing, such as changing hotspots month to month. This has the advantage of increasing statistical power in an evaluation. Another is organising hotspots in a more sophisticated manner (e.g. according to their risk at particular times of day or days of the week, such as schools at the start and end of the school day, or nightclubs in the evening).

Increasing direct treatment effects will decrease crime only if these crimes are not pushed around the corner. Here our study has fewer insights to offer. But the broader policing literature has consistently found that more police are associated with lower crime (Chalfin and McCrary 2017). Increasing general police presence alongside intensive policing could reduce crime displacement. Holding the number of police constant and shrinking high-crime patrol areas (quadrants) could have the same effect. Treating hotspots at the level of a neighbourhood rather than a street segment could also plausibly reduce adverse spillovers. These all deserve testing at scale.

9.2 Lessons for crime prevention and state building

From the perspective of crime and violence reduction, our results are consistent with a tenet of criminology: that crime and violence are highly concentrated in specific places. But if crime is easily displaced, then targeting, coordinating and concentrating resources in high-crime places may not be the right approach. Rather, it might be wiser to target the specific people who commit crimes or exhibit particular behaviours. Displacement may be inherently less likely than that in place-based approaches. This is the spirit of focused deterrence, which identifies the small group of people who commit serious crimes and uses threats and incentives to keep them from offending (Kennedy 2011). This is also the spirit of cognitive behavioural therapy, which fosters skills and norms of non-violent behaviour in high-risk young adults (Heller et al. 2017; Blattman et al. 2017). These all deserve testing at scale to make an accurate cost-effectiveness comparison. There are not enough data about the cost of alternatives; to our knowledge there has never been a formal evaluation of alternatives outside the US; and even within the US, there is a scarcity of rigorous and statistically powered evaluations of alternatives.

From the broader perspective of state building, the effort to build the last mile of the state in Bogotá parallels a broader set of cases. The tendency for people to elude the state, or simply run away, is as old as state coercion. Targeted state interventions simply create the illusion of local control. It may be that state coercion and state presence must be much more general, and much more widely spread, in order to be effective. Urban crime and violence literature have pushed theory and interventions to a more and more micro level; however, to be effective, interventions might have to be more broad based and stronger in order to keep crime from being pushed to nearby places. The monopoly of violence is necessarily broad, and order is inconsistent with an ungoverned periphery. Small-scale trials may have led us to the opposite conclusion. Larger-scale investigations, which are sorely needed in the US and more globally, provide more precise tests.

Appendix A: Sample size and power calculations

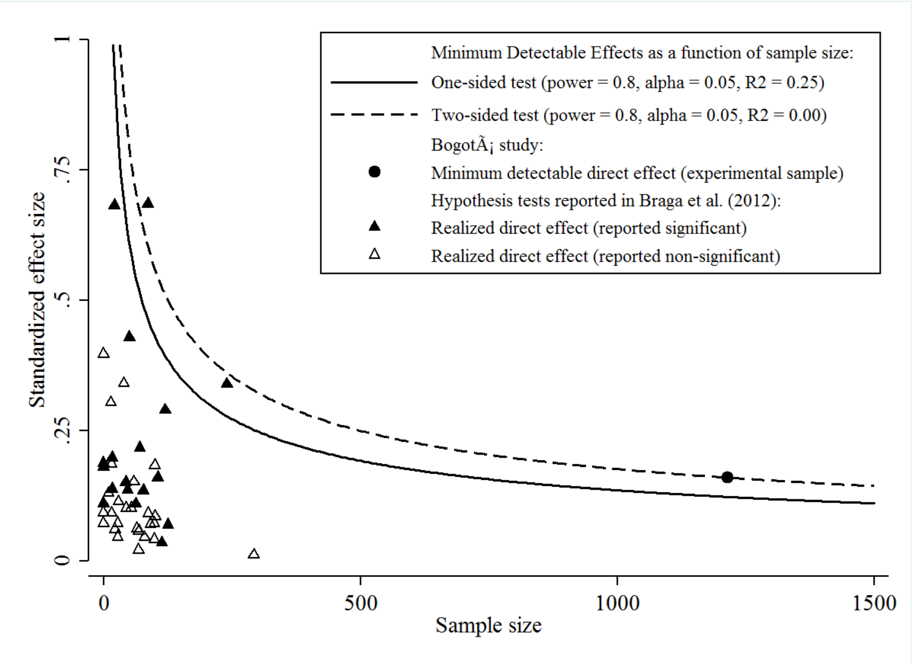
Power analysis

The aggregate effects on crime are difficult to pinpoint because of the small size of most studies. Figure A1 plots the systematically reviewed studies by sample size and effect sizes, for both direct and spillover effects. We calculate statistical power curves, representing the minimum effect size that we would expect to be able to detect with 80 per cent confidence. Note that even the largest studies do not exceed 50 or 100 treated hotspots, with a similarly modest number of spillover segments. The average effect size for direct hotspots treatment across the studies is 0.17 standard deviations, and 0.24 if statistically significant. While covariate adjustment and blocking strategies could improve statistical power slightly, these would produce marginal gains in precision at best.

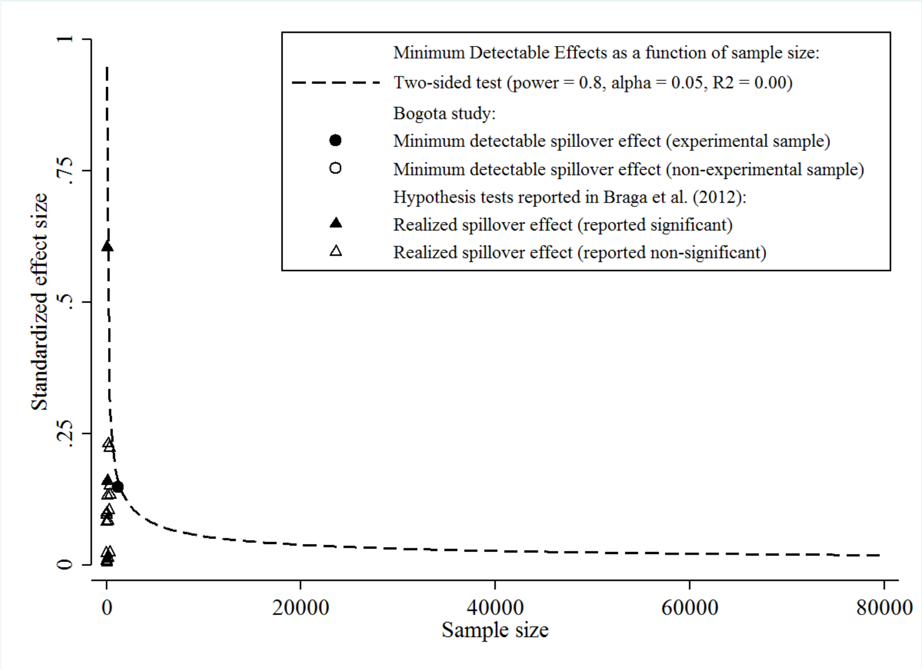
In Bogotá, the city tested two place-based security interventions on a scale large enough to identify direct treatment effects of 0.15 standard deviations, and spillovers as small as 0.02 standard deviations. We plot these in Figure A1. For fairness in the comparison, we plot the power of our study measured also on the basis of sample size and the number of treated units.

Figure A1: Statistical power in the intensive policing literature

(a) Direct and spillover effects within the experimental sample of hotspots



(b) Spillover effects into 'non-hotspots' proximate to the experimental sample



Notes: The figure depicts minimum detectable effects and realized effect sizes as a function of sample size. The vertical axis is in standard deviation units and measures minimum detectable effects for power curves and realize effect sizes for previous studies, and the horizontal axis measures sample size. The equations for power curves are $y = m \times 2\sqrt{\frac{1-R^2}{x}}$, where y is the standardized effect size, x is the sample size, and m is a multiple relating the standard deviation to the effect size. This multiple is 2.49 for one sided tests and 2.80 for two sided. Triangles represent a hypothesis test from previous studies and circles represent the minimum detectable effects in our study.

Online Appendixes

Online Appendix A: Results

http://www.3ieimpact.org/sites/default/files/2019-01/IE88-Online-Appendix-A-Results.pdf

Online Appendix B: Marginal effects of extra patrolling time

http://www.3ieimpact.org/sites/default/files/2019-01/ie88-online-appendix-b-marginal_effects_of_patrolling_time.pdf

Online Appendix C: Survey instruments

http://www.3ieimpact.org/sites/default/files/2019-01/ie88-online-appendix-c-survey_instruments.pdf

Online Appendix D: Pre-analysis plan

http://www.3ieimpact.org/sites/default/files/2019-01/ie88-online-appendix-d-pre-analysis_plan.pdf

Online Appendix E: Monitoring plan

http://www.3ieimpact.org/sites/default/files/2019-01/ie88-online-appendix-e-monitoring_plan.pdf

Online Appendix F: Descriptive statistics

http://www.3ieimpact.org/sites/default/files/2019-01/ie88-online-appendix-f-descriptive_statistics.pdf

Online Appendix G: .do files

The .do files have been made available online on 3ie’s Dataverse.

Online Appendix H: Cost data for the programme implementation to provide the ‘ingredients’ into CEA, CBA or CUA

http://www.3ieimpact.org/sites/default/files/2019-01/ie88-online-appendix-h-basic_costing_template.xls

References

Abadie, A, Athey, S, Imbens, G and Wooldridge, J, 2016. Clustering as a design problem. Unpublished working paper.

Abt, T and Winship, C, 2016. *What works in reducing community violence: a meta review and field study for the Northern Triangle*. Washington, DC: United States Agency for International Development.

Apel, R, 2013. Sanctions, perceptions, and crime: implications for criminal deterrence. *Journal of Quantitative Criminology*, 29(1), pp.67–101.

Aronow, M and Samii, C, 2017. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals of Applied Statistics*, 11(4), pp.1912–47.

Banerjee, AV, Duflo, E, Keniston, D and Singh, N, 2017. *The efficient deployment of police resources: theory and new evidence from a randomized drunk driving crackdown in India*. Working Paper. Boston, MA: MIT.

Becker, G, 1968. Crime and punishment: an economic approach. *Journal of Political Economy*, 76, pp.169–217.

Berman, E and Matanock, A, 2015. The empiricists' insurgency. *Annual Review of Political Science*, 18, pp. 443–64.

Berman, E, Shapiro, JN and Felter, JH, 2011. Can hearts and minds be bought? The economics of counterinsurgency in Iraq. *Journal of Political Economy*, 119(4), pp.766–819.

Blanes i Vidal, J and Mastrobuoni, G, 2017. *Police patrols and crime*. London: Centre for Economic Policy Research.

Blattman, C, Jamison, J and Sheridan, M, 2017. Reducing crime and violence: experimental evidence on cognitive behavioural therapy in Liberia. *American Economic Review*, 107(4), pp.1165–206.

Braga, A and Bond, BJ, 2008. Policing crime and disorder hotspots: a randomised controlled trial. *Criminology*, 46(3), pp.577–608.

Braga, AA, Weisburd, DL, Waring, EJ, Mazerolle, LG, Spelman, W and Gajewski, F, 1999. Problem-oriented policing in violent crime places: a randomised controlled experiment. *Criminology*, 37, pp.541–80.

Braga, AA, Papachristos, AV and Hurreau, DM, 2012. An ex post factor evaluation framework for place-based police interventions. *Campbell Systematic Reviews*, 8, pp.1–31.

Cassidy, T, Inglis, G, Wiysonge, C and Matzopoulos, R, 2014. A systematic review of the effects of poverty deconcentration and urban upgrading on youth violence. *Health and Place*, 26, pp.78–87.

Chalfin, A and McCrary, J, 2017. Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1), pp.5–48.

Chalfin, A and McCrary, J, 2018. Are US cities under policed? Theory and evidence. *Review of Economics and Statistics*, 100(1), pp.167–86.

Collazos, D, Garcia, E, Mejia, D, Ortega, D and Tobon, S, 2017. *Hotspots policing in a high crime environment: An experimental evaluation in Medellin*. In progress.

Di Tella, R and Schargrodsky, E, 2004. Do police reduce crime? Estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, 94(1), pp.115–133.

Draca, M, Machin, S and Witt, R, 2011. Panic on the streets of London: Police, crime, and the July 2005 terror attacks. *The American Economic Review*, 101(5), pp.2157–81.

Ehrlich, I, 1973. Participation in illegitimate activities: a theoretical and empirical investigation. *The Journal of Political Economy*, 81(3), pp.521–65.

Farrington, DP and Welsh, BC, 2008. Effects of improved street lighting on crime: a systematic review. *Campbell Systematic Reviews*, 59(13), pp.1–51.

Ferraz, C, Monteiro, J and Ottoni, B, 2016. *State presence and urban violence: Evidence from Rio de Janeiro's favelas*.

Gerber, AS and Green, DP, 2012. *Field experiments: Design, analysis, and interpretation*. New York: WW Norton.

Heller, SB, Shah, AK, Guryan, J, Ludwig, J, Mullainathan, S and Pollack, HA, 2017. Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago. *Quarterly Journal of Economics*, 132(1), pp.1–54.

Horvitz, DG and Thompson, DJ, 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260), pp.663–85.

Kennedy, DM, 2011. *Don't shoot: one man, a street fellowship, and the end of violence in inner-city America*. New York: Bloomsbury USA.

Kling, JR, Liebman, JB and Katz, LF, 2007. Experimental analysis of neighbourhood effects. *Econometrica*, 75(1), pp.83–119.

Levitt, SD and Miles, TJ, 2006. Economic contributions to the understanding of crime. *Annual Review of Law and Social Science*, 2(1), pp.147–64.

Mazerolle, LG, Prince, JF and Roehl, J, 2000. Civil remedies and drug control: a randomised field trial in Oakland, CA. *Evaluation Review*, 24, pp.212–41.

Muggah, R, de Carvalho, IS, Alvarado, N, Marmolejo, L and Wang, R, 2016. Making cities safer: citizen security innovations from Latin America, Strategic Note, Igarapé Institute (Rio de Janeiro: Igarapé Institute, World Economic Forum y BID).

Police Executive Research Forum, 2008. *Violent crime in America: What we know about hotspots enforcement*. Washington, DC.

Ratcliffe, JH, Taniguchi, T, Groff, ER and Wood, JD, 2011. Philadelphia foot patrol experiment: a randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology*, 49, pp.795–831.

Sherman, LW, Williams, S, Ariel, B, Strang, LR, Wain, N, Slothower, M and Norton, A, 2014. An integrated theory of hotspots patrol strategy: Implementing prevention by scaling up and feeding back. *Journal of Contemporary Criminal Justice*, 30(2), pp.95–122.

Sherman, L, Buerger, M and Gartin, P, 1989. *Beyond dial-a-cop: A randomised test of repeat call policing (recap)*. Washington, DC: Crime Control Institute.

Sherman, L and Rogan, DP, 1995. Deterrent effects of police raids on crack houses: A randomised, controlled experiment. *Justice Quarterly*, 12(4), pp.755–81.

Sherman, L and Weisburd, D, 1995. Does patrol prevent crime? The Minneapolis hotspots experiment. *Crime Prevention in the Urban Community*. Boston: Kluwer Law and Taxation Publishers.

Taylor, B, Koper, CS and Woods, DJ, 2011. A randomised controlled trial of different policing strategies at hotspots of violent crime. *Journal of Experimental Criminology*, 7(2), pp.149–81.

Telep, C, Mitchell, R and Weisburd, D, 2014. How much time should police spend at crime hotspots? Answers from a police agency directed randomized field trial in Sacramento, California. *Justice Quarterly*, 31(5), pp.905–33.

UNODC, 2014. *Global Study on Homicide 2013.* Vienna: United Nations Office on Drugs and Crime. Available at: <https://www.unodc.org/documents/data-and-analysis/statistics/GSH2013/2014_GLOBAL_HOMICIDE_BOOK_web.pdf> [Accessed 23 November 2018].

Vazquez-Bare, G, 2017. *Identification and estimation of spillover effects in randomized experiments*. arXiv preprint arXiv:1711.02745.

Weisburd, D and Gill, C, 2014. Block randomized trials at places: Rethinking the limitations of small n experiments. *Journal of Quantitative Criminology*, 30(1), pp.97–112.

Weisburd, D and Green, L, 1995. Measuring immediate spatial displacement: Methodological issues and problems. In *Crime and Place: Crime Prevention Studies*. Monsey, NY: Willow Tree Press, pp.349–59.

Weisburd, D and Telep, C, 2016. Hotspots policing: What we know and what we need to know. *Journal of Experimental Criminology*, 30(2), pp.200–20.

Weisburd, DL, Wyckoff, L, Ready, J, Eck, J, Hinkle, J and Gajewski, F, 2005. Does crime just move around the corner? A study of displacement and diffusion in Jersey City, NJ. *Criminology*, 44(August), pp.549–92.

Weisburd, D, Groff, D and Yang, S, 2012. *The criminology of place: street segments and our understanding of the crime problem*. New York: Oxford University Press.

Weisburd, D, Farrington, DP and Gill, C, 2017. What works in crime prevention and rehabilitation: An assessment of systematic reviews. Criminology & Public Policy, 16(2), pp.415-449.

Wilson, J and Kelling, G, 1982. Broken windows: the police and neighbourhood safety. *Atlantic Monthly*, 1982(March), pp.29–38.

Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/en/publications/3ie-impact-evaluation-reports/3ie-impact-evaluations/

Impact evaluation of the Philippine Special Program for Employment of Students, 3ie Impact Evaluation Report 87. Beam, E, Linden, L, Quimbo, S and Richmond, H, 2018.

Community-based distribution of oral HIV self-testing kits: experimental evidence from Zambia, 3ie Impact Evaluation Report 86. Hensen, B, Ayles, H, Mulubwa, C, Floyd, S, Schaap, A, Chiti, B, Phiri, M, Mwenge, L, Simwinga, M, Fidler S, Hayes, R, Bond, V and Mwinga, A, 2018.

Evaluating the economic impacts of rural banking: experimental evidence from southern India, 3ie Impact Evaluation Report 85. Field, E and Pande, R, 2018.

Direct provision versus facility collection of HIV tests: impacts of self-testing among female sex workers in Uganda. 3ie Impact Evaluation Report 84. Ortblad, K, Musoke, DK, Ngabirano, T, Oldenburg, C and Bärnighausen, T, 2018.

Increasing female sex worker HIV testing: effects of peer educators and HIV self-tests in Zambia, 3ie Impact Evaluation Report 83. Chanda, MM, Ortblad, KF, Mwale, M, Chongo, S, Kanchele, C, Kamungoma, N, Fullem, A, Bärnighausen, T and Oldenburg, CE, 2018.

Community delivery of antiretroviral drugs: a non-inferiority matched-pair pragmatic cluster-randomized trial in Dar es Salaam, Tanzania, 3ie Impact Evaluation Report 82. Francis, JM, Geldsetzer, P, Asmus, G, Ulenga, N, Ambikapathi, R, Sando, D, Fawzi, W and Bärnighausen, T, 2018.

Nourishing the future: targeting infants and their caregivers to reduce undernutrition in rural China, 3ie Impact Evaluation Report 81. Cai, J, Luo, R, Li, H, Lien, J, Medina, A, Zhou, H and Zhang, L, 2018.

Impacts of the World Food Programme's interventions to treat malnutrition in Niger. 3ie Impact Evaluation Report 80. Brück, T, Ferguson, NTN, Ouédraogo, J and Ziegelhöfer, Z, 2018.

Impact evaluation of the World Food Programme's moderate acute malnutrition treatment and prevention programmes in Sudan. 3ie Impact Evaluation Report 79. Guevarra, E, Mandalazi, E, Balegamire, S, Albrektsen, K, Sadler, K, Abdelsalam, K, Urrea, G and Alawad, S, 2018.

Impact evaluation of WFP's programs targeting moderate acute malnutrition in humanitarian situations in Chad. 3ie Impact Evaluation Report 78. Saboya, M, Rudiger, J, Frize, J, Ruegenberg, D, Rodríguez Seco, A and McMillon, C, 2018.

Improving midday meal delivery and encouraging micronutrient fortification among children in India, 3ie Impact Evaluation Report 77. Shastry, GK, Berry, J, Mukherjee, P, Mehta, S and Ruebeck, H, 2018.

Evaluation of infant development centres: an early years intervention in Colombia, 3ie Impact Evaluation Report 76. Andrew, A, Attanasio, O, Bernal, R, Cordona, L, Krutikova, S, Heredia, DM, Medina, C, Peña, X, Rubio-Codina, M and Vera-Hernandez, M, 2018.

Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone. 3ie Impact Evaluation Report 75*.* Cilliers, J, Dube, O and Siddiqi, B, 2018.

Impact evaluation of the Menabe and Melaky development programme in Madagascar, 3ie Impact Evaluation Report 74. Ring, H, Morey, M, Kavanagh, E, Kamto, K, McCarthy, N, Brubaker, J and Rakotondrafara, C, 2018.

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya, 3ie Impact Evaluation Report 73. Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018.

Impact and adoption of risk-reducing drought-tolerant rice in India, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India, 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll, 3ie Impact Evaluation Report 70*.* Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia, 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

Age at marriage, women's education, and mother and child outcomes in Bangladesh, 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

Evaluating agricultural information dissemination in western Kenya, 3ie Impact Evaluation Report 67. Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

Impact of the Uddeepan programme on child health and nutrition in India, 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya, 3ie Impact Evaluation Report 64*.* Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe, 3ie Impact Evaluation Report 63. Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

Increasing male partner HIV testing using self-test kits in Kenya, 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania, 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

Using HIV self-testing to promote male partner and couples testing in Kenya, 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

Increasing male partner HIV self-testing at antenatal care clinics in Kenya, 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Impact of free availability of public childcare on labour supply and child development in Brazil, 3ie Impact Evaluation Report 58. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico, 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme, 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

*The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional statu*s, 3ie Impact Evaluation Report 55. Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco, 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

Using advertisements to create demand for voluntary medical male circumcision in South Africa, 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania, 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

Voluntary medical male circumcision uptake through soccer in Zimbabwe, 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda, 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya, 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone, 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

Property tax experiment in Pakistan: Incentivising tax collection and improving performance, 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan, 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

Making networks work for policy: Evidence from agricultural technology adoption in Malawi, 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana, 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

*Evaluating the effectiveness of computers as tutors in China, 3*ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

Micro entrepreneurship support programme in Chile, 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh, 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India, 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya, 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation, 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda, 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests, 3ie Impact Evaluation Report 34. Persha, L and Meshack, C, 2016.

The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador, 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

The effect of transfers and preschool on children's cognitive development in Uganda, 3ie Impact Evaluation Report 32. Gillian, DO and Roy, S, 2016.

Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

Improving maternal and child health in India: evaluating demand and supply strategies, 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, 3ie Impact Evaluation Report 24. Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

Shelter from the storm: upgrading housing infrastructure in Latin American slums, 3ie Impact Evaluation Report 21. Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, 3ie Impact Evaluation Report 20. Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E, 2014.

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K, 2015.

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, 3ie Impact Evaluation Report 17. Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M, 2014.

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, 3ie Impact Evaluation Report 9. Ashraf, N, Bandiera, O and Jack, K, 2013.

Paying for performance in China's battle against anaemia, 3ie Impact Evaluation Report 8. Zhang, L, Rozelle, S and Shi, Y, 2013.

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, 3ie Impact Evaluation Report 7. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

The impact of daycare on maternal labour supply and child development in Mexico, 3ie Impact Evaluation Report 6. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, 3ie Impact Evaluation Report 5. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, 3ie Impact Evaluation Report 3. Casey, K, Glennerster, R and Miguel, E, 2013.

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, 3ie Impact Evaluation Report 2. Burwen, J and Levine, DI, 2012.

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, 3ie Impact Evaluation Report 1. Martinez, S, Naudeau, S and Pereira, V, 2012.

Two per cent of Bogotá's 136,984 streets accounted for all murders and a quarter of all crimes from 2012-2015. Increased state presence in the form of intensive policing and municipal services are often used to reduce disorder and deter crimes. This study evaluated the impact of both of these forms of state presence in targeted hotspots and assessed whether these place-based strategies displaced crimes to nearby streets. When assessed in isolation, intensive policing and municipal services interventions did not lead to a significant increase in security in hotspots. When both interventions were implemented together, their effect was intensified, resulting in a large and significant impact on security. Results also suggest that crimes might have been displaced to neighbouring streets that were located within 250 metres of treated hotspot street segments.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org