

# No.1

ENERO DE 2019

## Documentos **CEDE**

ISSN 1657-7191 Edición electrónica.

Hot spots policing in a high crime  
environment: An experimental  
evaluation in Medellín

---

Daniela Collazos  
Eduardo García  
Daniel Mejía  
Daniel Ortega  
Santiago Tobón

**CEDE**  
CENTRO DE ESTUDIOS SOBRE DESARROLLO ECONÓMICO

 Universidad de  
**los Andes**  
Facultad de Economía



Serie Documentos Cede, 2019-01  
ISSN 1657-7191 Edición electrónica.  
Enero de 2019

© 2019, Universidad de los Andes, Facultad de Economía,  
CEDE. Calle 19A No. 1 – 37 Este, Bloque W.  
Bogotá, D. C., Colombia Teléfonos: 3394949- 3394999,  
extensiones 2400, 2049, 2467  
[infocede@uniandes.edu.co](mailto:infocede@uniandes.edu.co)  
<http://economia.uniandes.edu.co>

Impreso en Colombia – Printed in Colombia

La serie de Documentos de Trabajo CEDE se circula con propósitos de discusión y divulgación. Los artículos no han sido evaluados por pares ni sujetos a ningún tipo de evaluación formal por parte del equipo de trabajo del CEDE.

El contenido de la presente publicación se encuentra protegido por las normas internacionales y nacionales vigentes sobre propiedad intelectual, por tanto su utilización, reproducción, comunicación pública, transformación, distribución, alquiler, préstamo público e importación, total o parcial, en todo o en parte, en formato impreso, digital o en cualquier formato conocido o por conocer, se encuentran prohibidos, y sólo serán lícitos en la medida en que se cuente con la autorización previa y expresa por escrito del autor o titular. Las limitaciones y excepciones al Derecho de Autor, sólo serán aplicables en la medida en que se den dentro de los denominados Usos Honrados (Fair use), estén previa y expresamente establecidas, no causen un grave e injustificado perjuicio a los intereses legítimos del autor o titular, y no atenten contra la normal explotación de la obra.

Universidad de los Andes | Vigilada Mineducación  
Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento personería jurídica: Resolución 28 del 23 de febrero de 1949 Minjusticia.

# Hot spots policing in a high crime environment: An experimental evaluation in Medellín\*

Daniela Collazos      Eduardo García      Daniel Mejía      Daniel Ortega  
Santiago Tobón<sup>†</sup>

## Abstract

*Objectives:* Test direct, spillover and aggregate effects of hot spots policing on crime in a high crime environment. *Methods:* We identified 967 hot spot street segments and randomly assigned 384 to a six-months increase in police patrols. To account for the complications resulting from a large experimental sample in a dense network of streets, we use randomization inference for hypothesis testing. We also use non-experimental streets to test for spillovers onto non-hot spots, and examine aggregate effects city-wide. *Results:* Our results show an improvement in short term security perceptions and a reduction in car thefts, but no direct effects on other crimes or satisfaction with policing services. We see larger effects in the least secure places, especially for short term security perceptions, car thefts and assaults. We find no evidence of crime displacement but rather a decrease in car thefts in nearby hot spots and a decrease in assaults in nearby non-hot spots. We estimate that car thefts decreased citywide by about 11 percent. *Conclusions:* Our study highlights the importance of context when implementing hot spots policing. What seems to work in the U.S. or even in Bogotá is not as responsive in Medellín (and vice versa). Further research—especially outside the U.S.—is needed to understand the role of local crime patterns and police capacity on the effectiveness of hot spots policing. *JEL codes:* K42, O17, E26, J48, C93. *Keywords:* crime, spillover effects, police, hot spots, field experiment, Colombia.

---

\*This project was possible because of the collaboration of the Ministry of Defense of Colombia, the National Police of Colombia, and the City of Medellín. In particular, we are grateful to the former Minister of Defense Juan C. Pinzón, Director of the National Police General Jorge H. Nieto, and Mayor of Medellín Anibal Gaviria. CESED at Universidad de los Andes and the Latin American Development Bank (CAF) coordinated research activities. Cifras y Conceptos collected the data. For research assistance we are grateful to Juan Carlos Angulo and Eduardo Fagre. For comments we thank Chris Blattman, Marcela Eslava, Don Green, Patryk Perkowski, Juan F. Vargas, Hernando Zuleta and numerous conference participants. For financial support we thank the Ministry of Defense of Colombia, the Latin American Development Bank (CAF), the City of Medellín, Organización Ardila Lülle and Open Society Foundations.

<sup>†</sup>Collazos: Secretariat of Security of Bogotá, daniela.collazos@scj.gov.co; García: University of Rochester, Economics, egarc12@ur.rochester.edu; Mejía: Office of the Attorney General of Colombia, daniel.mejia@fiscalia.gov.co; Ortega: Latin American Development Bank (CAF), dortega@caf.com; Tobón (corresponding author): University of Chicago and Innovations for Poverty Action, tobon@uchicago.edu.

# Policía en puntos calientes en ambientes de alta criminalidad: Una evaluación experimental en Medellín

Daniela Collazos      Eduardo García      Daniel Mejía      Daniel Ortega  
Santiago Tobón \*

## Resumen

*Objetivos:* Evaluar los efectos directos, de desplazamiento y agregados de intervenciones de policía en puntos calientes en ambientes de alta criminalidad. *Métodos:* Identificamos 967 puntos calientes (segmentos de vía) y aleatoriamente asignamos 384 para un incremento de seis meses en patrullaje policial. Para controlar por las complicaciones que resultan de una muestra experimental grande en una red densa de vías, probamos hipótesis utilizando inferencia con permutaciones (randomization inference). También utilizamos la muestra no experimental de vías para estudiar el desplazamiento, y examinamos efectos agregados en toda la ciudad. *Resultados:* Nuestros resultados sugieren que hay una mejora en la percepción de seguridad en el corto plazo y una reducción en el hurto de carros, pero no se observan efectos directos sobre otros delitos o la satisfacción con el servicio de policía. Observamos efectos más grandes en los sitios más inseguros, especialmente para la percepción de seguridad en el corto plazo, el hurto de carros y los casos de lesiones personales. No encontramos evidencia de desplazamiento del crimen. Por el contrario, observamos caídas en el hurto de carros en puntos calientes cercanos, y caídas en casos de lesiones personales en puntos no calientes cercanos. Estimamos que el hurto de carros bajó un 11 % en la ciudad como resultado de la intervención. *Conclusiones:* Nuestro estudio resalta la importancia del contexto para la implementación de intervenciones de policía en puntos calientes. Lo que parece funcionar en EEUU o incluso en Bogotá no funciona en Medellín (y viceversa). Mayor investigación—especialmente por fuera de EEUU—es necesaria para entender el rol de los patrones locales de crimen y la capacidad policial en la efectividad del patrullaje en puntos calientes. *Códigos JEL:* K42, O17, E26, J48, C93. *Palabras clave:* crimen, efectos de desplazamiento, policía, puntos calientes, experimento de campo, Colombia.

---

\*Collazos: Secretaría de Seguridad de Bogotá, daniela.collazos@scj.gov.co; García: Universidad de Rochester, Economía, egarc12@ur.rochester.edu; Mejía: Fiscalía General de la Nación, daniel.mejia@fiscalia.gov.co; Ortega: Banco de Desarrollo de América Latina (CAF), dortega@caf.com; Tobón (autor correspondiente): Universidad de Chicago e Innovations for Poverty Action, tobon@uchicago.edu.

# 1 Introduction

In most cities, crime is highly concentrated in a small number of places usually known as crime hot spots. This situation is prevalent in both developed and developing countries. For instance, Weisburd (2015) analyzes data from eight cities in the U.S. and Israel and document that in large cities as New York or Tel Aviv, between 5 and 6 percent of the streets account for half of all reported crimes. In small cities such as Brooklyn Park, MN or Redlands, CA, only 2 percent of the city streets account for half the crimes. Similarly, Mejía et al. (2015) analyze data for the five largest Colombian cities and find that half the crimes are concentrated in only 3 to 5 percent of the streets in all cases.

A common policy response to this problem, fostered since the late 1980s when the study of micro-geographic units in criminology started to gain relevance (Sherman et al., 1989), is to direct disproportionate police efforts to these places. These tactics are commonly known as hot spots policing. The idea is that criminals would either be deterred by the increased risk of arrest due to police presence (Becker, 1968; Ehrlich, 1973), or incapacitated and taken out of the criminal market when they are effectively imprisoned. Hot spots policing tactics are backed by a large body of evidence.<sup>1</sup> Braga et al. (2014) conduct a systematic review of hot spots policing studies and report that 20 out of 25 tests of the core hypothesis point in the direction of large reductions in crime and disorder.<sup>2</sup> Moreover, systematic reviews conducted by Bowers et al. (2011), Braga et al. (2012) and Weisburd and Telep (2016) conclude that more than just moving crime around the corner, hot spots policing interventions also benefit places in the surroundings of targeted locations. As a result, a large number of police departments in the U.S. and other countries such as Argentina, Colombia, Trinidad and Tobago, Uruguay or Venezuela has adopted these tactics.<sup>3</sup>

---

<sup>1</sup>Some of the most prominent experimental studies in criminology are Sherman et al. (1989) in Minneapolis, MN; Sherman and Weisburd (1995) in Minneapolis, MN; Weisburd and Green (1995) in Jersey City, NJ; Sherman et al. (1995) in Kansas City, KS; Braga et al. (1999) in Jersey City, NJ; Mazerolle et al. (2000) in Oakland, CA; Braga and Bond (2008) in Lowell, MA; Taylor et al. (2011) in Jacksonville, FL; Ratcliffe et al. (2011) in Philadelphia, PA; Groff et al. (2015) in Philadelphia, PA; and Santos and Santos (2016) in Port St. Lucie, FL. There is also non-experimental evidence on hot spots policing in the criminology literature as Sviridoff et al. (1992) in New York, NY; Cohen et al. (2003) in Pittsburgh, PA; and Lawton et al. (2005) in Philadelphia, PA. Studies on police and crime in the economics literature include Di Tella and Schargrofsky (2004) in Buenos Aires, Argentina; and Draca et al. (2011) in London, U.K.

<sup>2</sup>See also Abt and Winship (2016) and Weisburd et al. (2017).

<sup>3</sup>See Police Executive Research Forum (2008) for U.S. data. See the Report on hot spots policing by one of the major national newspapers in Argentina (La Nación). In Colombia, the National Police Department requires that police patrols intensify their activities in crime hot spots. This is outlined in the the Quadrants Policing Guidelines. For the case of Uruguay, see the Report from the Ministry of the Interior. For Trinidad and Tobago see Sherman et al. (2014). In Venezuela, there was an unsuccessful initiative to implement hot spots policing strategies in Sucre. One of the coauthors of this study was involved in the early evaluation efforts.

Notwithstanding this enthusiasm, the body of evidence on hot spots policing in Latin America is virtually non-existent. Latin America is the most violent region in the world, with homicide rates per 100,000 people above 40 in different countries, some of the most murderous cities, and some of the larger and most pervasive criminal organizations. Latin America holds less than 10 percent of the world’s population but about a third of all homicides.<sup>4</sup> Moreover, it is not evident that implementation capability in Latin America matches that in the U.S., especially for rather complex programs as hot spots policing, where sound monitoring procedures, accountable police departments, and functioning bureaucracies are key to success.<sup>5</sup> Because of these differences in criminal behavior and implementation capability, it is unclear whether hot spots policing programs in Latin America would converge or diverge with the U.S. based evidence, both in terms of direct and spillover effects of the interventions. To the best of our knowledge, alongside the Blattman et al. (2018) study in Bogotá and the Sherman et al. (2014) study in Trinidad and Tobago, this is one of the first hot spots policing evaluations in Latin America, and one of the largest by an order of magnitude.

To set up this experiment, we first split the street network of Medellín into 37,055 segments—a length of street between two corners—and used geo-located police crime data, as well as qualitative inputs from police patrols throughout the city, to identify 967 crime hot spots. We focused on car and motorbike thefts, personal robberies, homicides and assaults to identify these crime hot spots. We randomly assigned 384 hot spots to a six-months increase in daily police time. Police patrols were expected to intensify their presence in these streets and conduct their usual activities: check background records on people and vehicles, conduct arrests, drug and merchandise seizures, and recover stolen property.

We used both police crime data and an original victimization and perception survey to evaluate the impacts of the intervention. Police crime data consists of crime reports registered mainly at police stations. These data include the type of crime, day and exact location coordinates of the event. Moreover, police patrols in Medellín and most large Colombian cities use a device to receive citizens’ calls and check criminal records. This device sends a signal with the exact location coordinates of the patrol regularly, and we used these data to measure and enforce compliance by sending two weekly reports to the Police on daily average patrolling times for every treatment street. We conducted the victimization and perception survey in all 967 streets beforehand, as well as during the last weeks of the intervention to study changes in security perceptions, satisfaction with policing services, and

---

<sup>4</sup>See Consejo Ciudadano para la Seguridad Pública y Justicia Penal, Global Study on Homicide 2013 and Transnational Organized Crime in Central America and the Caribbean.

<sup>5</sup>As Chong et al. (2014) show, there is wide variation in implementation capabilities even for simple policies as returning mail when the address is non-existent.

to retrieve direct data on victimization.

We estimate direct treatment effects by comparing average outcomes between targeted and non-targeted streets. However, as Blattman et al. (2018) show, in a setting where experimental streets are not isolated and indeed form large clusters in many parts of the city, there are several threats to the assumption of no interference between units. For instance, criminals could move their activities to neighboring control streets or patrols could increase patrolling time in control streets nearby targeted hot spots, as they need to traverse them to comply with the program. To account for this problem, we follow Blattman et al. (2018) and split the control group in three sub-groups: short-range spillover streets located within 125 meters of treatment hot spots, long-range spillover streets located between 125 and 250 meters from treatment hot spots, and pure control streets located farther than 250 meters from treatment hot spots. This allow us to estimate direct treatment effects by comparing targeted streets with pure, presumably uncontaminated controls. We can also estimate spillover effects by comparing short and long-range spillover units with pure controls. We defined these radii ad-hoc, hence we present our results using different aggregations of the control units.<sup>6</sup> Finally, we also follow Blattman et al. (2018) and use the non-experimental sample of streets to study spillover effects onto non-hot spots. We compare streets in the non-experimental sample located nearby targeted streets with those that are farther away.

The police largely complied with the required increase in daily patrolling time in targeted hot spots. Treatment streets received between 50 and roughly 80 percent more patrolling time, depending on how we estimate these differences. This adds up to about 50 to 70 more minutes of police presence per day. Also, surveyed citizens reported observing more police presence in targeted streets.

Our findings suggest that some of the conclusions of the U.S. based evidence are not borne out in Medellín, though some others are. First, we find a large decrease in reported car thefts in targeted streets. However, we see virtually no change in the number of motorbike thefts, personal robberies, homicides or assault cases.

Second, we see a major improvement in security perceptions. This change, however, is bounded by the six-month intervention period. Beyond that, we see no differences in citizens' perceptions between targeted and control streets. Also, we see no changes in citizens' satisfaction with policing services.

Third, we find no evidence of crime displacement. Instead, we see a large drop in car thefts in hot spots close to targeted streets. Moreover, when we analyze spillovers onto the

---

<sup>6</sup>Blattman et al. (2018) classify control streets using 250 and 500 meters radii. We use 125 and 250 meters for two reasons. First, in our case there are only 20 control streets beyond 500 meters and that small sample size would drive the results. Second, Blattman et al. (2018) don't find crime displacement beyond 250 meters.

non-experimental sample of streets, we see a statistically significant decrease in assault cases in streets nearby targeted hot spots. We did not expect this latter result, as we do not observe any change in assault cases in treated hot spots. However, to the extent that the non-experimental sample of streets provides us with a much larger statistical power resulting from the increase in sample size, this diffusion of the program’s benefits could point in the direction of a small drop in assault cases in targeted streets, however unobservable given the more limited statistical power.

Fourth, when we look at heterogeneous treatment effects based on baseline crime or perception levels, we see much larger effects in the least secure hot spots. Our more robust results are for car thefts and security perceptions over the intervention period, the same outcomes for which we observe more precise direct treatment effects. In the case of car thefts, for instance, the program effects in streets at or above the 90<sup>th</sup> percentile according to baseline car thefts levels are about five times larger relative to the average treatment effect. The direct effects are larger and statistically significant in the highest crime hot spots for assault cases. This result provides another possible explanation for the slight diffusion of benefits we observe onto non-experimental streets for this type of crime.

Fifth, since even small direct or spillover effects add up when a large number of streets is exposed to a given condition, we perform a back-of-the-envelope estimation of aggregate effects. In particular, we use our best guess of the average effects in each case (the estimated coefficients, even if they are not statistically significant) and the number of streets falling in each condition, to estimate the net effect of the intervention citywide. We find that the intervention led to a decrease of about 55 car thefts (11 percent relative to the total number of reported cases citywide). This estimate is not statistically significant, but a 10 percent confidence interval goes from -115 to 3, suggesting there is a high chance of an aggregate decrease in this specific type of crime. For other crimes, the confidence intervals on aggregate effects are generally large, and we cannot rule out the possibility of crime spillovers outweighing direct treatment effects.

Generally, our study adds some nuisance to the U.S. based hot spots policing research, as we do not observe direct treatment effects for major crimes as homicides or assaults. We believe the large increase in perceptions of security is promising, and should be studied further. Compared to the closer study conducted by Blattman et al. (2018) in Bogotá, we find some important differences. On one hand, they find that property crimes displace, while we observe a diffusion of benefits to nearby hot spots for car thefts. On the other hand, they find that violent crimes decrease in targeted hot spots and may even decrease in nearby streets. We generally do not observe direct treatment effects on violent crimes. Factors such as criminal behavior (violent crimes are more instrumental in Medellín than they are



in Bogotá), and police manpower (Medellín has 60 percent more police than Bogotá relative to the population) could probably explain a portion of such differences.

Importantly, we believe our results add caution to the immediate adoption of U.S. based programs in Latin America, a region with large contextual differences. Take for instance Scared Straight programs, which have been proven ineffective in the U.S. and yet they are prevalent in many Latin American countries.<sup>7</sup> The fact that we find hot spots policing programs to have some mild positive effects on one specific type of crime but not others, adds nuisance to the immediate adoption without accounting for both crime patterns and implementation capability in each context.

## 2 Institutional framework

### 2.1 The City of Medellín

Medellín is the second largest city in Colombia with a population of about 2.5 million. It is the capital and economic center of the department of Antioquia, which participates with about 14 percent of the national gross domestic product.<sup>8</sup> The city is so densely populated that a recent report ranks it third in the world with about 20,000 people per square kilometer.<sup>9</sup>

Medellín was known worldwide in the 1980s for housing one of the most violent and powerful drug cartels, whose war with both the local and national governments led to unprecedented levels of violence. The evolution of the homicide rate in the city is depicted in figure 1. Violence reached a maximum in 1991—at the outset of the war—with a homicide rate of 422 per 100,000 people. Indeed, more than 50,000 people were murdered in the decade between 1986 and 1995. Following Escobar’s death in 1993, the presence of urban guerrillas and the rise of paramilitaries led to continuous levels of violence until 2002, when a confluence of new national and local security policies, the demobilization of the paramilitaries and the hegemonic power over criminal groups in the city reached by paramilitary leader Diego Murillo, also known as *Don Berna*, drove down the levels of violence to a historical minimum (McDermott, 2014). In 2006 *Don Berna* was extradited to the U.S. The situation led to a new rise of violence as several criminal leaders aimed to re-claim *Don Berna’s* former power.<sup>10</sup> By the end of 2014, as organized crime in the city realigned under a new collective

---

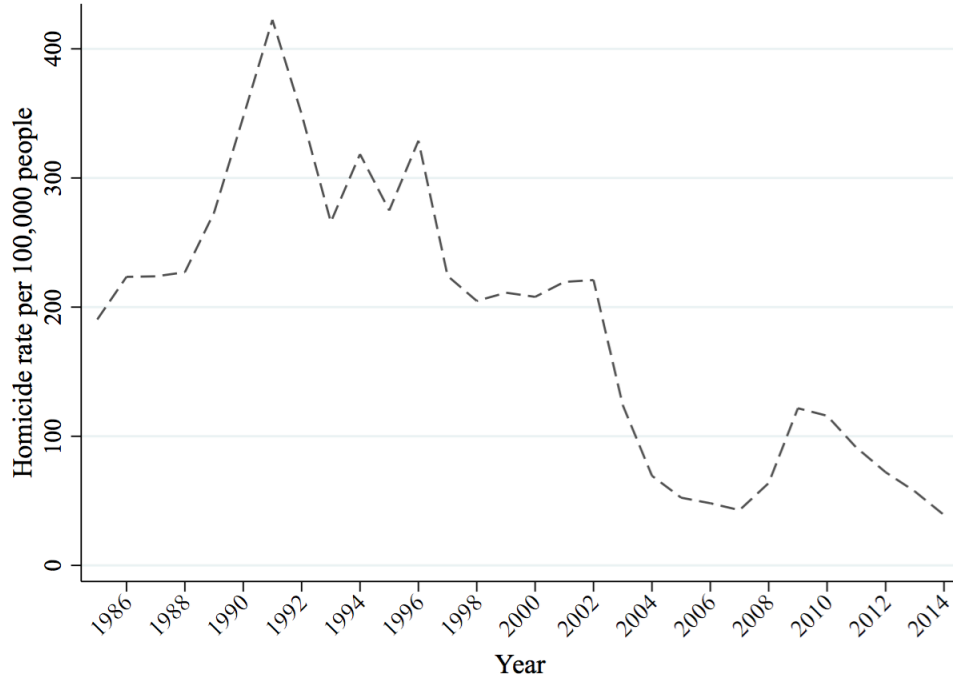
<sup>7</sup>See Petrosino et al. (2003). In Colombia, the program is called *Delinquir no paga*, and it is run by the prison authority countrywide.

<sup>8</sup>Data on population and gross domestic product is from the National Department of Statistics.

<sup>9</sup>See the report The Worlds’ Most Crowded Cities by the World Economic Forum.

<sup>10</sup>The war for power after *Don Berna’s* extradition was documented by local media. See for instance *La Guerra que Desangró a Medellín* in the major regional newspaper El Colombiano.

Figure 1: Evolution of the homicide rate in Medellín



*Notes:* The figure depicts the evolution of the homicide rate in Medellín. Data on the number of homicides is from the Secretariat of Security and data on population is from the National Department of Statistics.

leadership,<sup>11</sup> homicide rates were back to historical minimums. However, violence is heterogeneous throughout the city. Medellín is divided in 16 urban *comunas*, and by 2014 some had less than 10 homicides per 100,000 people as La América, El Poblado and El Popular, while La Candelaria—located downtown—reached 135.<sup>12</sup>

The Secretariat of Security and the Metropolitan Police estimate that about two-thirds of the city’s neighborhoods are under the control of organized crime. However, there seems to be variation in the degree of control that these organizations exert over their communities, ranging from a minor involvement in drug sales and prostitution in parts of the city with higher income levels, to the regulation of crime and violence, illegal drug markets, public space—including the imposition of curfews, and the provision of security and justice services in the most disadvantaged neighborhoods (Duncan et al., 2015; Giraldo et al., 2014). As a result of the high level of organization in criminal groups in Medellín, a relevant share of

<sup>11</sup>See for instance McDermott (2014) and the report *Así Funciona la Oficina*, published in the newspaper El Colombiano.

<sup>12</sup>The influence of organized crime in the regulation of violence becomes apparent with the case of El Popular, a low income *comuna* where the presence of some criminal organizations is prevalent. See for instance Duncan et al. (2015).

the violence is thought to be instrumental in nature. Indeed, the local authorities estimate that about 60 percent of all homicides are somewhat related to organized crime.<sup>13</sup> In sum, organized groups in Medellín regulate and use violence, and regulate and engage in other crimes—such as motor vehicle thefts.

Medellín is disproportionately affected by car and motorbike thefts. As of 2014, about 5 percent of the country’s population lived in Medellín, but 15 percent of reported car thefts and 19 percent of reported motorbike thefts in that year occurred within the city’s jurisdiction. Other crimes such as homicides, assaults and personal robberies are not particularly concentrated in Medellín. In 2014, 5 percent of all homicides, 3.2 percent of all assault cases and 5.2 percent of all personal robberies were reported to be committed within the city limits.<sup>14</sup>

## 2.2 Policing strategies in Colombia

The focus of this study is on police patrolling strategies. This leaves other policing activities—such as criminal intelligence—out of scope. The core of the patrolling scheme of Colombian Police is the Quadrants Model,<sup>15</sup> which establishes well defined patrolling areas known as quadrants—similar to police beats in standard U.S. policing—to be under the surveillance of a police patrol. Each quadrant is assigned to one police station. By the beginning of 2014, the urban area of Medellín had 13 police stations and a total of 411 quadrants.

A police patrol consists of two people with a motorbike. There are three patrols assigned to a quadrant, and each of them covers one of three eight-hours shift. The two members of a patrol are required to be always together, even if they need to leave the quadrant—as when formally registering an arrest at a detention center.<sup>16</sup> When the patrol leaves the quadrant, the closest patrol is required to cover for the high priority activities. The activities undertaken by police patrols are coordinated by the police station to which the quadrant is assigned. Each police station has a weekly meeting to define the activities by the hour for each quadrant, and patrols are expected to comply with every activity.<sup>17</sup> The usual activities for police patrols are to conduct background checks on people and motor vehicles—for which they have daily quotas, seize illegal drugs or other illegal merchandise, arrest people and traverse

---

<sup>13</sup>Estimates are from the Secretariat of Security of Medellín.

<sup>14</sup>Data is from the National Police of Colombia.

<sup>15</sup>Specifically, it is known as the *Modelo Nacional de Vigilancia Comunitaria por Cuadrantes*. For more details on the model see the Quadrants Policing Guidelines. Throughout the paper, we use quadrants as a translation for *cuadrante*.

<sup>16</sup>These are usually police stations or special locations managed by the National Prosecutor’s Office.

<sup>17</sup>The activities are recorded in the *Tabla de Acciones Mínimas Requeridas*, and specify the activity, the time of the day, general places to focus on, and other relevant details to provide surveillance to the quadrant.

the streets within the quadrant. The performance of police patrols is measured upon their compliance with patrolling activities, their operational results—arrests and seizures, and the number of crimes reported within the borders of the quadrant. Indeed, the most common benchmark for grading performance is the situation of the quadrant in terms of operational results and reported crimes by the same day of the previous year.

The Medellín Metropolitan Police has a moderate police to population ratio, with about 387 policemen per 100,000 people. As a benchmark, high crime U.S. cities such as Baltimore or Chicago have about 450 police per 100,000 people.<sup>18</sup> To account for the limitations on the availability of resources and personnel, police patrols are required to intensify patrolling activities—within the quadrant—in areas where crime is more prevalent or, put differently, crime hot spots.<sup>19</sup> However, police regulations are loose in the way they specify the size and characteristics of these areas, as well as the specific instructions on how to patrol them. As a result, such intensification of patrolling activities is not fulfilled—as reported by senior police officers in Medellín and other cities, and inefficient allocation of patrolling activities is fairly common.<sup>20</sup>

## 3 The Medellín hot spots policing experiment

### 3.1 Data

**Units of analysis** As suggested by Weisburd et al. (2012), we use street segments as our units of analysis for the identification of crime hot spots. Street segments—a length of street between two corners—provide a reasonably small geographic area to focus police patrolling activities. A smaller alternative would be a specific address, while a larger one would be the quadrant. We have at least two reasons to believe this is the adequate approach. First, if we define a larger area, it is possible that the hottest places within the area are going to be under patrolled. Second, we want a small number of agents to be accountable for the results within their jurisdiction, so that responsibilities are not diluted across a large number of policemen. The urban area of Medellín has a network of 37,055 street segments with a length of about 90 meters, on average. Our experimental sample consists of a subset of 967 high-crime streets. We explain how we selected this sample in Section 3.2.

---

<sup>18</sup>All figures are for active police and exclude those performing solely administrative tasks. Data from U.S. cities if from the FBI Uniform Crime Report .

<sup>19</sup>This is specified in the Quadrants Policing Guidelines.

<sup>20</sup>One major motivation for this hot spots policing intervention was the concern by senior officers at the National Police on how to identify and systematically target crime hot spots. Indeed, this concern allowed us to collaborate with the police in designing this and other similar interventions.

**Police crime data** We use data on reported crimes provided by the Metropolitan Police and the Secretariat of Security. In particular, we focus on five specific types of crime: homicides, assaults, car and motorbike theft, and personal robbery. For each reported crime, we have data on the specific day of occurrence and exact coordinates.<sup>21</sup> We match each reported crime to a street segment using 40 meters buffers around the segment. Any crime located within the buffer of a street segment is automatically matched to that segment. If there is a crime within the buffer of one or more street segments, we match the crime to the closest one using euclidean distances. For our main analysis, we use data for each type of crime. For the identification of crime hot spots we used a weighted sum of these five crimes. We explain these weights and the motivation to use them in section 3.2.

One limitation of these data is related to reporting rates, and whether the measurement error resulting from the willingness to report is correlated with treatment. On average, one-fourth of all crimes in Colombia were reported in 2014, and this varies by type of crime.<sup>22</sup> We expect homicides and motor vehicle thefts to have considerable reporting rates, while personal robberies or assault cases may suffer from under-reporting to a larger extent. We deal with the problem of under-reporting by using an original citizen survey, which we present in detail below.

**Patrolling data** Since the beginning of 2015, every police patrol in Medellín uses a device to receive citizens' calls and check the criminal history of an individual or a motor vehicle. This device also sends a signal with the exact location coordinates of the patrol, usually in windows of 30 seconds to one or two minutes. We match these signals to street segments using 40 meters buffers as we do for crime data. Using the time stamp of every signal, we approximate entries and exits to each street segment and aggregate them to estimate the daily patrolling time. Data on these signals was provided by the Metropolitan Police.<sup>23</sup> We estimate that high-crime streets—the top 3 percent in the distribution of pre-treatment

---

<sup>21</sup>In some crime reporting locations, the individual filing the report is allowed to point to the specific location in an interactive map. In these cases, the exact coordinates are automatically recorded. In other reporting locations, the individual filing the report specifies the address. In these cases, the Metropolitan Police records the exact coordinates after the crime is reported. The main reporting locations are police stations.

<sup>22</sup>Reporting rates are from the National Survey on Citizen Security conducted by the National Department of Statistics.

<sup>23</sup>These data was subject to some amount of uncertainty and was erratic during some periods. In the best cases, each device sent the signals frequently, and we were able to identify clearly the time of entry and exit to the hot spots. When this was not the case—for instance, when time stamps were separated by large periods of several hours with changing locations—we had to assign patrolling time making ad-hoc decisions. In general, when we observed only one signal from a street with the next one being in a different location, we assigned three minutes of police patrolling time. When we observed one signal from a street with the next one being assigned to the same, but separated by many hours, we top-coded the entry at the duration of the shift. These decisions resulted from discussions with police patrols and officials.

crime—had roughly one hour of police patrolling time every day before the intervention started.<sup>24</sup>

**Survey data** As we discussed above, police crime data has several limitations that result mainly from imperfect reporting rates, differences in these reporting rates across crimes and places, and the likelihood of correlation with treatment. For instance, it may be that more police presence incentivizes citizens to report crimes. On the other hand, since police patrols know which are the streets subject to the intervention, they could have incentives not to receive reports in those streets or to suggest they are located elsewhere whenever the citizen filing the report is dubious. To account for these issues we conducted baseline and endline surveys in all street segments in our experimental sample.

For the baseline we surveyed three people per street, while for the endline we surveyed two.<sup>25</sup> We retrieved information on perceptions of security (last 6 months, last 12 months, and general perception), victimization (to the respondent directly or a third person), and perceptions of police service (quality of work, satisfaction with service and presence). Since some of the questions had multiple ordinal answers, we build z-scores for all variables to ease comparability and interpretation.<sup>26</sup> For both the baseline and endline surveys, we averaged responses to get measures at the street level.

**Administrative data on socioeconomic characteristics** Finally, we also used administrative data on socioeconomic characteristics to correct for any imbalance in our analysis. In particular, we use the length of the street and the distance to: nearest police station, nearest community facility, nearest education facility, nearest justice facility, nearest transportation facility, nearest institutional facility,<sup>27</sup> nearest recreational facility, nearest health facility, and nearest religious facility.

We report summary statistics in baseline characteristics for high-crime streets in our experimental sample in Table 1. On average, these streets had 0.10 car thefts, 0.48 motorbike thefts, 0.73 personal robberies, 0.05 homicides and 0.30 assault cases reported during 2014.

---

<sup>24</sup>Author’s estimations based on data from the Metropolitan Police and the Secretariat of Security. Patrolling times were estimated during pilots between January and April 2015. We estimate the distribution of total reported crimes using a crime index that weights crimes according to the average prison sentence. See section 3.2 for more details.

<sup>25</sup>This was a result of budgetary constraints from the Secretariat of Security, which covered for the expenses for the end-line survey.

<sup>26</sup>The original survey measures were the following: (i) indicator for direct and indirect victimization of the respondent; (ii) score from 1 to 3 on perceptions of security for last 6 months; (iii) score from 1 to 4 on perceptions of security for last 12 months; (iv) score from 1 to 4 on general perceptions of security; (v) score from 1 to 4 on quality of police work; (vi) score from 1 to 4 on satisfaction with police service; (vii) indicator for an increase in police presence.

<sup>27</sup>Institutional facilities are mainly from the local government to provide different public services.

For each of these type of crimes there were streets with no reports. Indeed there are some streets with no report for any type of crime. This is a result of the validation process with the police that we explain below in section 3.2, when we included streets with no reported crimes in our experimental sample. Presumably, some of these streets had many crimes but no reports. Our baseline survey measures are generally close to zero, on average.

### 3.2 Selection of crime hot spots

We identified high-crime streets along with the Metropolitan Police in four steps. First, we created an aggregate crime index as a weighted sum of homicides, assaults, car and motorbike thefts, and personal robberies. The Secretariat of Security and the Metropolitan Police wanted to target crime hot spots that were socially more costly. Hence, we used the weights used by Mejía et al. (2015), which resemble the relative average sentence for each type of crime according to the Colombian penal code.<sup>28</sup> Second, we ranked the network of 37,055 streets according to the crime index for the 2012-2014 period, and pre-selected the top 3% streets in the distribution. Third, since the index is based solely on reported crimes, we validated with each police station and included or excluded streets in each case. We conducted the validation process with both senior officials and patrolling agents in all cases.<sup>29</sup> Finally, when a dense cluster of streets were identified as crime hot spots, we cleared intermediate streets to avoid contamination and joined contiguous streets into one hot spot.

Our final sample consisted in 817 clustered hot spots when we count contiguous streets as one hot spot (679 independent streets or one-street hot spots, 128 two-street hot spots, 8 three-streets hot spots, and 2 four-street hot spots), or 967 hot spots when we considered each street independently. After we handed over the treatment hot spots to the Metropolitan Police, senior officials required us to target independent streets rather than clustered hot spots. In principle, this would ease implementation and the understanding that patrols had on the instructions. As a result, we conducted the randomization using the 817 hot spots, but implemented the intervention using the 967 independent streets. Figure 2 depicts the map of hot spots.

---

<sup>28</sup>We used the following weights for each crime: 0.550 for homicides, 0.112 for assaults, 0.221 for car and motorbike theft, and 0.116 for personal robbery.

<sup>29</sup>There were different reasons to include or exclude streets. For instance, streets nearby metro stations had a disproportionate number of personal robberies reported, but the location was usually the transport system rather than the station. Other streets were not pre-selected because there were few reports.

Table 1: Summary statistics for individual hot spots in the experimental sample (N=967)

	Mean (1)	S.D. (2)	Min (3)	Max (4)
<i>a. Baseline crime in 2014</i>				
# of car thefts	0.10	0.33	0	2
# of motorbike thefts	0.48	0.93	0	9
# of personal robberies	0.73	2.21	0	52
# of homicides	0.05	0.23	0	3
# of assaults	0.30	1.01	0	11
<i>b. Baseline survey</i>				
Perception - 6 months, z-score	0.02	1.00	-2.27	1.61
Perception - 12 months, z-score	0.01	1.01	-2.82	1.94
Perception - general, z-score	0.02	1.00	-2.54	1.78
Direct or indirect victimization, z-score	-0.01	0.99	-0.75	5.03
Police - labor, z-score	0.00	1.02	-2.78	2.07
Police - satisfaction, z-score	-0.01	1.02	-2.62	2.29
Police - presence, z-score	0.01	1.00	-1.34	1.21
<i>c. Other characteristics</i>				
Average daily patrolling time, minutes	56.90	66.15	0	988
Meters from police infrastructure	911.03	482.52	36	3,553
Length, meters	87.24	51.57	4	623
Meters from community center	271.36	187.76	8	988
Meters from education facility	191.10	149.45	7	820
Meters from justice facility	704.07	414.93	18	2,461
Meters from public transportation	643.94	434.59	4	2,166
Meters from institutional facility	654.95	501.13	16	2,915
Meters from recreational facility	275.83	171.11	5	1,095
Meters from health center	399.74	251.34	11	1,727
Meters from religious center	250.78	165.57	7	1,116
<i>d. Quadrant characteristics</i>				
# of streets in quadrant	95.46	66.55	1.00	396.00
# of experimental streets in quadrant	5.42	3.20	1.00	15.00

Notes: In columns (1) to (4) we report summary statistics for 967 crime hot spots in our experimental sample (street segments). Each observation is weighted by the inverse of the probability of being observed in its experimental condition.



Figure 2: Experimental sample of crime hot spots (street segments) in Medellín



*Notes:* The figure depicts the experimental sample of 967 crime hot spots when considered as independent street segments, or 817 crime hot spots when we join contiguous streets into one hot spot.

### 3.3 Intervention

The idea of the intervention was to correct for a misallocation of police patrols: prior to the intervention, we estimate that those streets concentrating one-third of all crime were receiving about 8 percent of patrolling time.

We instructed police patrols to increase the dosage of police patrolling time from roughly one hour a day per hot spot to at least 105 minutes divided in 7 entries of about 15 minutes each. The activities while patrolling were expected to be the usual: check criminal records on people and cars, make door-to-door visits to the community, and conduct arrests and drug or merchandise seizures. The instructions were given to the six agents assigned to each quadrant, and we suggested that hot spots where most of the crimes were reported at night had more entries during the night shift.<sup>30</sup> To ease implementation and prevent a major loss in patrolling time in other streets—the non-hot spots, the Metropolitan Police required us to limit the number of treatment hot spots in each quadrant to three in La Candelaria station and four in any other station. Since the average number of streets per quadrant is 90, we estimate that non-hot spot streets in a quadrant with four treated hot spots would lose about five minutes of daily patrolling time, on average.

The intervention lasted for roughly six months from May 4 through November 19, 2015. To ensure compliance we sent two weekly reports to senior officers from the Metropolitan Police, who monitored the performance of all patrols. The reports included details on the compliance levels at the police station, quadrant, shift and street segment levels.

## 4 Empirical framework

### 4.1 Randomization and schedule of potential outcomes

Recall from Section 3.2 that we used the sample of 817 clustered crime hot spots to conduct the randomization. The majority of them were independent streets but some were between two and four clustered streets. We randomized the 817 hot spots to treatment and control considering the two restrictions required by the Metropolitan Police: assign a maximum of three hot spots to treatment in La Candelaria station, and assign a maximum of four hot spots to treatment in any other station. We imposed the restrictions at the street level so that, for instance, no quadrant in La Candelaria had more than three streets—rather than clustered hot spots—assigned to treatment.

When considering the sample of 817 clustered streets, we assigned a total of 334 hot

---

<sup>30</sup>In effect, these instructions were included in the weekly meeting to define the patrolling strategy and recorded in the *Tabla de Acciones Mínimas Requeridas*.

Table 2: Treatment assignment in the experimental and non-experimental samples

	Experimental sample (1)	Non- experimental sample (2)	Total streets in the city (3)
Treatment streets	384	0	384
Short-range spillovers (<125m from treatment streets)	271	8,738	9,009
Long-range spillovers (125-250m from treatment streets)	189	9,397	9,586
Pure control streets (>250m from treatment streets)	123	17,953	18,076
Total streets	967	36,088	37,055

*Notes:* In column (1) we present the distribution of streets in the experimental sample, in column (2) we present the distribution of streets in the non-experimental sample, and in column (3) we present the distribution of all streets in the city.

spots to treatment and 483 to control. When considering the sample of 967 independent streets, a total of 384 individual hot spots were assigned to treatment and 583 to control. Since the police required us to implement the intervention using the sample of individual streets, throughout the paper we consider these 967 streets—and the corresponding treatment assignment—as our experimental sample.<sup>31</sup> Importantly, because of a minor bug in the randomization code, 7 streets had a probability of being treated equal to 1. We drop these streets from the analysis (see section 4.2).

We did not pre-specify any potential outcome regarding spillovers.<sup>32</sup> However, to flexibly estimate spillovers and account for some of the challenges to identification detailed in section 4.3, we divided control hot spots in three categories: short-range spillovers (control streets located between 0 and 125 meters from treated hot spots), long-range spillovers (control streets located between 125 and 250 meters from treated hot spots), and pure control streets (control streets located at more than 250 meters from treated hot spots). Since treatment assignment is random, exposure to spillovers is also random, and thus we follow Blattman et al. (2018) and extend the spillover analysis to streets in the non-experimental sample. Table 2 presents the distribution of treatment status for both the experimental and non-experimental samples considering exposure to short and long-range spillovers. The city has 37,055 street segments in total, with 967 in the experimental sample and the remaining 36,088 in the non-experimental sample. As we explain in section 4.2, our regressions include subsamples of these streets.

We present balance tests on pre-intervention characteristics for streets in the experimental

<sup>31</sup>We account for the clustered assignment using clustered standard errors, as we explain in section 4.2.

<sup>32</sup>For a similar analysis with pre-specified spillover ranges, see Blattman et al. (2018).

sample in Table 3. All p-values in column (2) are above the conventional levels for statistical significance. However, as there are some imbalances resulting from chance, we control for baseline crime and other street characteristics in our main regression analysis.

## 4.2 Estimating equations

We estimate intent to treat, short-range and long-range spillover effects within the experimental sample using equation (1):

$$y_{sp} = \alpha T_{sp} + \beta S_{sp}^{SR} + \gamma S_{sp}^{LR} + \delta_p + \Gamma X_{sp} + \varepsilon_{sp} \quad (1)$$

where  $y$  is some crime or perception outcome in individual hot spot  $s$  and police station  $p$ ;  $T$  is an indicator for assignment to the hot spots policing treatment,  $S_{sp}^{SR}$  is an indicator of exposure to short-range spillovers (untreated streets located within 125 meters from treated hot spots),  $S_{sp}^{LR}$  is an indicator of exposure to long-range spillovers (untreated streets located between 125 and 250 meters from treated hot spots),  $\delta_p$  stands for police station fixed effects,  $X_{sp}$  is a vector of street characteristics and pre-intervention crime levels, and  $\varepsilon_{sp}$  are standard errors clustered at the unit of randomization. Specifically, since we randomized treatment using the 817 original clustered hot spots—which consider contiguous streets as one individual hot spot, we cluster standard errors using this structure. However, when we assume the presence of spillovers (i.e.  $\beta \neq 0$  or  $\gamma \neq 0$ ) we use randomization inference to estimate exact p-values. Moreover, we estimate equation (1) using weighted least squares (see section 4.3 below). The coefficient  $\alpha$  estimates intent to treat effects, while  $\beta$  and  $\gamma$  estimate short and long-range spillover effects to neighboring hot spots, respectively. Finally, in each regression we include the sample of experimental streets that have a positive probability of assignment to both treatment and control.<sup>33</sup>

As Blattman et al. (2018), we also use the non-experimental sample of streets to estimate spillovers. Specifically, we estimate equation (2):

$$y_{sp} = \beta^{NE} S_{sp}^{SR} + \gamma^{NE} S_{sp}^{LR} + \delta_p + \Gamma X_{sp} + \varepsilon_{sp} \quad (2)$$

where all variables and indicators follow from equation (1). The coefficient  $\beta^{NE}$  estimates short-range spillover effects to non-experimental streets, and  $\gamma^{NE}$  estimates long-range spillover effects to non-experimental streets. When estimating equation (2) we also use randomization inference to estimate exact p-values, and use weighted least squares rather than

---

<sup>33</sup>Recall from section 4.1 that a bug in the randomization code assigned 7 streets to treatment with probability 1—hence they had a probability of 0 to be in the control condition. We drop these streets from the analysis.

Table 3: Balance tests on pre-intervention characteristics for individual hot spots in the experimental sample (N=967)

	Treatment - Control (1)	p-value (2)
<i>a. Baseline crime in 2014</i>		
# of car thefts	-0.03	0.18
# of motorbike thefts	-0.06	0.35
# of personal robberies	-0.05	0.73
# of homicides	-0.01	0.70
# of assaults	0.03	0.75
<i>b. Baseline survey</i>		
Perception - 6 months, z-score	0.09	0.17
Perception - 12 months, z-score	0.03	0.69
Perception - general, z-score	0.06	0.37
Direct or indirect victimization, z-score	0.02	0.80
Police - labor, z-score	-0.02	0.81
Police - satisfaction, z-score	-0.04	0.61
Police - presence, z-score	0.03	0.69
<i>c. Other characteristics</i>		
Average daily patrolling time, minutes	-2.84	0.49
Meters from police infrastructure	32.54	0.28
Length, meters	1.62	0.64
Meters from community center	-9.05	0.48
Meters from education facility	2.74	0.79
Meters from justice facility	34.96	0.21
Meters from public transportation	-32.27	0.21
Meters from institutional facility	7.60	0.80
Meters from recreational facility	-6.62	0.55
Meters from health center	-9.34	0.55
Meters from religious center	-8.88	0.42

*Notes:* In column (1) we report the difference between treatment and control street segments, and in column (2) the corresponding p-value for the test of no difference. We run weighted least squares regressions, weighting each observation with the inverse of the probability of being observed in its experimental condition.

ordinary least squares to estimate the equation (see section 4.3 below).

The main issue with the non-experimental sample of streets is to reach the highest comparability between groups of streets in different spillover or control conditions. For instance, streets that are too far from treated experimental hot spots may not be comparable to those that are closer. Hence, in these regressions we include streets that have a positive probability of being assigned to all experimental conditions considered in each specific analysis. For instance, when we assume the presence of short-range spillovers, the spillover group consists of streets located at less than 125 meters from experimental streets that were assigned to treatment, that also meet these two conditions: (i) have a positive probability of being within 125 meters from treated hot spots (which is immediate in this case, as they were effectively exposed to spillovers); and (ii) have a positive probability of not being exposed to spillovers within 125 meters. Note the second condition leaves out of the analysis any non-experimental street that has a probability of 1 of being exposed to short-range spillovers. On the other hand, the control group consists of individual hot spots that are located at more than 125 meters from experimental streets that were assigned to treatment, that also meet the two conditions above. A similar rationale follows for the selection of the sample of streets when we assume the presence of both short and long-range spillovers.<sup>34</sup>

### 4.3 Challenges to identification

As Blattman et al. (2018) show, the scale of these kind of interventions in a dense network of streets leads to several identification problems—even with random treatment assignment. First, the stable unit treatment value assumption—SUTVA—can be violated if crime is displaced from treatment to control streets, or if control hot spots receive more or less patrolling time depending on their location relative to targeted streets. We account for this problem by dividing control streets into pure control and spillover categories, and estimating spillovers flexibly over different distance ranges. For instance, when we estimate equation (1) assuming that  $\beta \neq 0$  and  $\gamma \neq 0$ , we consider there could be potential violations of the assumption of no interference between experimental units up to 250 meters. Similarly, when we estimate equation (1) assuming that  $\beta \neq 0$  but  $\gamma = 0$ , we consider the violation of the assumption of no interference between units up to 125 meters. Finally, when we estimate equation (1) assuming that  $\beta = 0$  and  $\gamma = 0$  we are implicitly assuming there is no violation of the assumption of no interference between experimental units.<sup>35</sup>

---

<sup>34</sup>See also Blattman et al. (2018) for a similar analysis in the city of Bogotá, where all these considerations were pre-specified.

<sup>35</sup>We could assume even farther spillover effects (i.e. violations of the SUTVA) but the number of pure control units largely decreases. See Blattman et al. (2018) for a pre-specified design that allows to test for spillover effects within 500 meters.

Second, assuming the presence of spillovers within buffers surrounding treatment hot spots creates a clustering structure that is not easily identifiable. For example, when there is a dense area with a large number of hot spots that are relatively close (as the area to the center-right in figure 2), once one street is assigned to treatment all hot spots within 125 meters of that street are assigned to the short-range spillover status. Hence, these streets form a cluster and we cannot model the structure of the clustering using geographical areas as a quadrant or a police station. Indeed, some of those streets can be in a different quadrant or a different police station, and some that actually fall in the same quadrant or police station are not part of the cluster that resulted from the randomization. Blattman et al. (2018) show that, in such a situation, usual standard errors over-estimate the precision of treatment and spillover effects and the way to account for the problem is to use randomization inference to estimate exact p-values. As they do, we repeat the randomization procedure 10,000 times and estimate treatment and spillover effects under each randomization so that we obtain the sampling distribution under the sharp null hypothesis of no treatment effects—making no assumption on the distribution of the error term. Then, we estimate the p-value in each case as the probability of obtaining an estimate that is as large as the one generated by the experiment.<sup>36</sup>

Third, the restrictions imposed by the Metropolitan Police on the number of treated streets per quadrant and the fact that not all quadrants have the same number of hot spots, lead to different probabilities of treatment assignment across streets in the experimental sample. Moreover, the probabilities of being exposed to spillovers within the different radii also varies. Perhaps more important, since hot spot streets tend to be clustered in specific locations throughout the city (again, as the area to the center-right in figure 2 illustrates) the probabilities of assignment to treatment, spillover or pure control conditions are also correlated with crime occurrence. We follow Blattman et al. (2018) and weight each observation by the inverse of the probability of being exposed to its experimental condition. In practice, this procedure gives less weight to streets that had a high probability of being assigned to some condition and ended up effectively assigned to it.

Finally, Blattman et al. (2018) also show that—when assuming the presence of spillovers of any range—both the clustering of hot spot streets and the different probabilities of treatment assignment across streets in the experimental sample, lead to a positive bias in estimated treatment effects. To account for this problem, we estimate the bias using randomization inference and subtract it from our final estimates.

---

<sup>36</sup>See also Gerber and Green (2012).

## 5 Results

### 5.1 Compliance

Police patrols complied with the instructions to intensify the dosage of patrolling time in targeted streets. Table 4 presents first stage effects on two different compliance measures: daily patrolling time estimated with location devices, and police presence as reported by citizens in the endline survey. *Panel a* presents the results assuming no spillovers (comparing targeted with non-targeted experimental streets), *Panel b* presents the results assuming spillovers up to 125 meters (comparing targeted with non-targeted streets located at more than 125 meters from treated hot spots), and *Panel c* presents the results assuming spillovers up to 250 meters (comparing targeted with non-targeted streets located at more than 250 meters from treated hot spots). In *Panel a* we estimate standard errors clustered at the unit of randomization, while in *Panel b* and *Panel c* we estimate p-values using randomization inference. Targeted streets received between 50 and roughly 80 percent more patrolling time—depending on the specification—and the difference in patrolling time between treated and control hot spots is always statistically significant at the 1 percent level. We also see that surveyed citizens reported an increase in police presence from 0.05 to 0.27 standard deviations—depending on the specification. The reported increase is imprecise for the no spillovers case, but it is statistically significant at the 5 percent level when we assume the presence of short-range spillovers, and when we assume the presence of short and long-range spillovers.

The increase in patrolling time was generally sustained throughout the intervention period. Figure 3 depicts daily average patrolling time for the pre-intervention, intervention and post-intervention periods. The vertical lines denote the beginning and the end of the intervention, and the empty spaces correspond to periods of data instability where measurement was imprecise due to updates in the tracking software. We see that patrolling times were about the same before and after the intervention, while treatment streets received more police time consistently during the intervention period. Since police patrols are generally required to intensify their activities at crime hot spots—and this requirement goes beyond this specific intervention—the sharp drop in compliance at the end was unexpected. We discuss issues of police incentives in section 6.

### 5.2 Results assuming no spillovers

We study the effects of the intervention on two sets of outcomes, reported in Table 5. The first set is in *Panel a*, and consists of police crime data on reported crimes. We focus

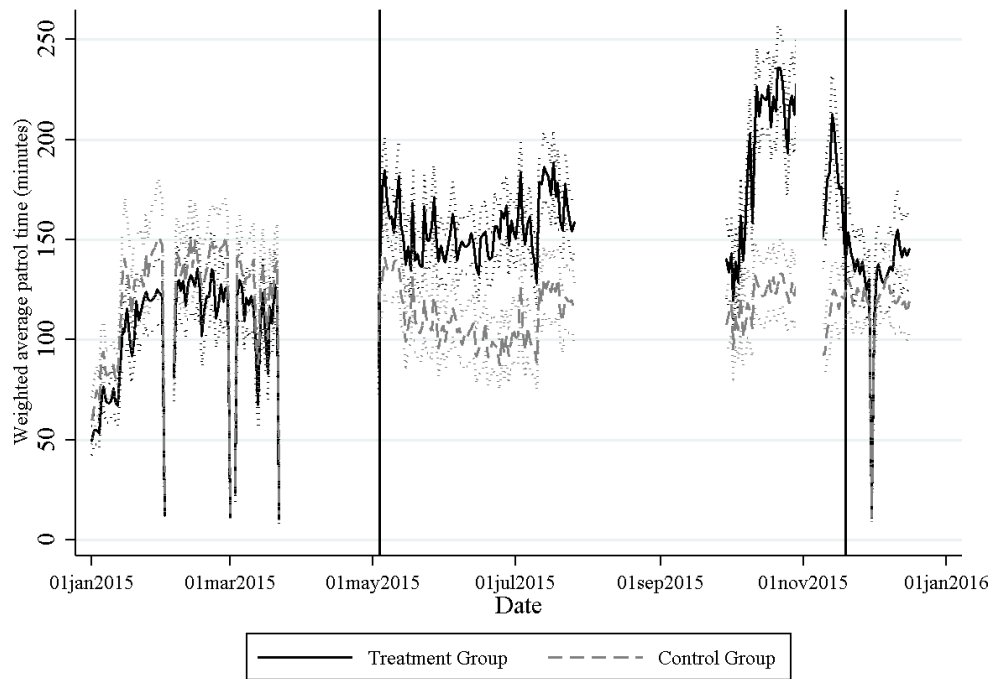


Table 4: First stage effects on compliance measures (N=960)

	Control mean (1)	Intent to treat (2)
<i>a. Assuming no spillovers</i>		
Average patrolling time, minutes	103.410	51.469 [4.406]***
Police presence increased in the street, z-score	-0.006	0.048 [0.071]
<i>b. Assuming short-range spillovers</i>		
Average patrolling time, minutes	98.474	49.179 <b>0.000</b>
Police presence increased in the street, z-score	-0.108	0.172 <b>0.040</b>
<i>c. Assuming short and long-range spillovers</i>		
Average patrolling time, minutes	82.721	70.170 <b>0.000</b>
Police presence increased in the street, z-score	-0.105	0.270 <b>0.035</b>

*Notes:* Column (1) reports control means and column (2) the coefficient on first stage results for compliance. *Panel a* assumes no spillovers and compares all treatment vs all control hot spots. In *Panel a* we cluster standard errors at the randomization level: streets assigned together to treatment are a cluster (those we joined before randomizing), while streets assigned individually have their own cluster (reported in brackets with \* for p-values <0.10, \*\* for p-values <0.05, \*\*\* for p-values <0.01). *Panel b* assumes short-range spillovers and compares all treatment vs control hot spots at more than 125 meters from any treatment hot spot. *Panel c* assumes short and long-range spillovers and compares all treatment vs control hot spots at more than 250 meters from any treatment hot spot. In *Panel b* and *Panel c* we estimate p-values using randomization inference (reported in italics with values <0.1 in bold). All regressions include controls for baseline crime, baseline survey measures and other street characteristics listed in Table 1. Each observation is weighted by the inverse of the probability of being observed in its experimental condition. We exclude 7 streets from the experimental sample that had a probability of treatment equal to 1, resulting from a minor bug in the randomization code.

Figure 3: Daily average patrolling time in the treatment and control groups



*Notes:* The figure depicts daily average patrolling time for the treatment and control groups (including streets in short and long-range spillover radii). The two vertical lines denote the beginning and the end of the intervention, and the empty spaces correspond to periods of data instability where measurement was imprecise due to updates in the tracking software.

our attention on the individual crimes that were included in the crime index to make the selection of hot spot street segments, and for which we have data available: car thefts, motorbike thefts, personal robberies, homicides and assaults. When we assume there is no crime displacement—or contamination of treatment to control hot spots that are close to targeted streets—we see no effect on the number of reported crimes. The coefficients for car thefts, motorbike thefts, personal robberies and assaults are negative but rather imprecise. The coefficient for homicides is positive but negative values are also well within the confidence intervals.<sup>37</sup>

The second set is in *Panel b*, and consists of survey measures on perceptions of security, direct or indirect victimization, and perceptions of police service. When we assume there is no contamination of un-treated hot spots close to targeted streets, we see an increase of about 0.228 standard deviations in perceptions of security for the intervention period of 6 months. The coefficient is statistically significant at the 1 percent level. For all other survey outcomes on perceptions of security, direct or indirect victimization, and perceptions of police service the coefficients are generally far from conventional levels of statistical significance.

### 5.3 Results assuming short-range spillovers

As we mention in section 4.3, if there is crime displacement from treatment to control streets that are located nearby, the estimates from Table 5 would be biased. For instance, crime displacement to control streets would lead us to overstate the effects of the intervention or, on the other hand, a diffusion of the benefits of hot spots policing to control streets would lead us to understate the effects. We report the results of the intervention assuming short-range spillovers (within 125 meters) in Table 6. Since we are now assuming the presence of spillovers, we use randomization inference to estimate exact p-values for the null hypothesis of no effects. These p-values are reported in italics with values below 0.10 in bold. Column (2) presents results for the intention to treat effects and column (3) presents results for short-range spillovers.

When we look at the effects on police crime data (*Panel a*), we find evidence of a decrease of 0.046 reported car thefts in treated hot spots. Relative to the average number of car thefts in control streets of 0.082, this effect is equivalent to a decrease of about 56 percent. This

---

<sup>37</sup>One obvious concern with our approach is that, while we are evaluating program effects using multiple outcomes, we are not correcting p-values for multiple comparisons. We acknowledge this limitation and note that common corrections as the Bonferroni adjustment will render mostly non-significant estimates, as sometimes they are extreme (Dunn, 1961). We leave the interpretation to the reader, but note that the main effects are generally stable across specifications. We also note that the outcome measures using police crime data correspond to the original set of outcomes we used to identify the experimental hot spots in the first place.

Table 5: Intention to treat effects in the experimental sample, assuming no spillovers (N=960)

	Control mean (1)	Intent to treat (2)
<i>a. Police crime data</i>		
# of car thefts	0.059	-0.016 [.018]
# of motorbike thefts	0.202	-0.003 [.034]
# of personal robberies	0.766	-0.070 [.099]
# of homicides	0.021	0.012 [.014]
# of assaults	0.146	-0.010 [.029]
<i>b. Survey data</i>		
Perception - 6 months, z-score	-0.067	0.228 [.067]***
Perception - 12 months, z-score	0.011	0.040 [.065]
Perception - general, z-score	0.029	-0.002 [.066]
Direct or indirect victimization, z-score	-0.001	0.018 [.073]
Police - labor, z-score	-0.003	0.053 [.072]
Police - satisfaction, z-score	0.007	0.028 [.071]

*Notes:* Column (1) reports control means and column (2) the coefficient on the intention to treat. We cluster standard errors at the randomization level: streets assigned together to treatment are a cluster (those we joined before randomizing), while streets assigned individually have their own cluster (reported in brackets with \* for p-values <0.10, \*\* for p-values <0.05, \*\*\* for p-values <0.01). All regressions include controls for baseline crime, baseline survey measures and other street characteristics listed in Table 1. Each observation is weighted by the inverse of the probability of being observed in its experimental condition. We exclude 7 streets from the experimental sample that had a probability of treatment equal to 1, resulting from a minor bug in the randomization code.

result is statistically significant at the 10 percent level. We also see a decrease of 0.057 reported car thefts in un-treated hot spots within 125 meters of targeted streets (almost 70 percent relative to the average number of car thefts in control streets). This positive spillover effect is statistically significant at the 5 percent level. This result is consistent with the estimates reported in Table 5, when we assume no-spillovers. Specifically, since there are positive spillovers to hot spots close to targeted streets for car thefts, when we estimate equation (1) assuming no spillovers we underestimate the direct effects of the program. Moreover, we find no statistically significant direct or spillover effects on the counts of other crimes. Indeed, all p-values are above 0.35.

In *Panel b* we report effects on endline survey measures. We find evidence of an increase of 0.211 standard deviations in perceptions of security for the intervention period of six months. This effect is statistically significant at the 5 percent level. It is also consistent with the results in Table 5, when we assume no-spillovers. In this case, we see no evidence of positive spillovers on perceptions of security, hence the magnitude of the estimates and the statistical significance for intention to treat effects are generally similar in both cases: when we assume no interference between experimental streets and when we assume the presence of short-range spillovers up to 125 meters. We do not see any statistically significant direct or spillover effects on other survey measures on perceptions of security, nor we see any changes in reported victimization or in citizen’s satisfaction with policing services.

## 5.4 Results assuming short and long-range spillovers

Table 7 presents the results of the intervention assuming short-range spillovers (within 125 meters) and long-range spillovers (between 125 and 250 meters). Column (2) presents results for the intention to treat effects, column (3) presents results for short-range spillovers and column (4) presents results for long-range spillovers. Results using police crime data as outcomes are in *Panel a*. Generally, we see no statistically significant direct, short-range spillover or long-range spillover effects. The coefficients for car thefts, however imprecisely estimated, are consistent with the previous results reported in Table 5 and Table 6. In particular, the percentage changes for the direct and short-range spillover effects remain relatively similar to the case when we assume short-range spillovers. The decrease is about 47 percent in reported car thefts in targeted streets (it was 56 percent in Table 6) and about 69 percent in streets within 125 meters from treated hot spots (it was 70 percent in Table 6). However, since the size of the control group decreases we would expect the effects assuming both levels of spillovers to be less precise.<sup>38</sup>

---

<sup>38</sup>Note the pure control group for the case of short-range spillovers includes control streets located at more than 125 meters from targeted streets: 189 that are between 125 and 250 meters plus 123 that are at more

Table 6: Intention to treat effects in the experimental sample, assuming short-range spillovers (N=960)

	Control mean (1)	Intent to treat (2)	Short-range spillovers (3)
<i>a. Police crime data</i>			
# of car thefts	0.082	-0.046 <b>0.051</b>	-0.057 <b>0.038</b>
# of motorbike thefts	0.229	-0.020 <i>0.540</i>	-0.012 <i>0.719</i>
# of personal robberies	0.385	-0.023 <i>0.700</i>	0.057 <i>0.386</i>
# of homicides	0.016	0.011 <i>0.353</i>	-0.013 <i>0.914</i>
# of assaults	0.124	-0.028 <i>0.838</i>	-0.018 <i>0.812</i>
<i>b. Survey data</i>			
Perception - 6 months, z-score	-0.028	0.211 <b>0.029</b>	0.016 <i>0.818</i>
Perception - 12 months, z-score	0.097	-0.031 <i>0.769</i>	-0.106 <i>0.227</i>
Perception - general, z-score	0.043	-0.038 <i>0.649</i>	-0.030 <i>0.632</i>
Direct or indirect victimization, z-score	-0.012	0.067 <i>0.558</i>	0.047 <i>0.806</i>
Police - labor, z-score	-0.012	0.069 <i>0.498</i>	0.114 <i>0.311</i>
Police - satisfaction, z-score	-0.032	0.014 <i>0.736</i>	-0.020 <i>0.826</i>

*Notes:* Column (1) reports control means, column (2) the coefficient on the intention to treat, and column (3) the coefficient on short-range spillovers. We estimate p-values using randomization inference (reported in italics with values <0.1 in bold). All regressions include controls for baseline crime, baseline survey measures and other street characteristics listed in Table 1. Each observation is weighted by the inverse of the probability of being observed in its experimental condition. We exclude 7 streets from the experimental sample that had a probability of treatment equal to 1, resulting from a minor bug in the randomization code.

As for the previous cases, we present the results for endline survey outcomes in *Panel b*. Generally, we see no statistically significant direct, short-range spillover or long-range spillover effects when we assume the presence of spillovers up to 250 meters. However, and similar to the case of reported car thefts, the coefficients on perceptions of security for the intervention period of 6 months are consistent with our previous findings. In particular, we see an increase of 0.2 standard deviations, which is similar to the case of no-spillovers (reported in Table 5, where we see an effect of 0.228 standard deviations), and the case of short-range spillovers (reported in Table 6, where we see an effect of 0.211 standard deviations). As expected, the estimates are less precise for the short and long-range spillovers case—as we lose sample size—and the p-value rises to 0.166. Importantly, the coefficients on direct or indirect victimization point in the direction of adverse effects both in targeted hot spots as well as short and long-range spillover streets. These coefficients are not statistically significant, but they are generally close to the conventional 10 percent level.

## 5.5 Heterogeneous treatment effects

It is evident from Table 1 that there is wide variation in pre-treatment crime levels within our experimental sample of hot spots. For instance, one street had as many as 52 personal robberies while some other had none during 2014. In this section, we explore if this level of variation leads to different responses to the hot spots policing intervention. Figure 4 presents the results for outcomes using police crime data. Each circle corresponds to the intent to treat effect in a regression with a restricted sample of hot spots, measured in standard deviations. We estimate the effects using equation (1) assuming short-range spillovers. Filled in circles denote a randomization inference p-value below 0.1.<sup>39</sup> We see larger and statistically significant effects for car thefts at higher crime hot spots (sub-figure a). Indeed, the effect grows from 0.05 standard deviations when we exclude only streets in the bottom 10<sup>th</sup> percentile to about 0.25 standard deviations when we consider only those streets at or above the 90<sup>th</sup> percentile. As for the average treatment effects, the heterogeneous effects are rather imprecise for motorbike thefts, personal robberies, homicides and assaults. We do observe, however, a pattern of larger effects in highest crime hot spots for motorbikes and assaults (sub-figures b and e), with a statistically significant decrease of almost 0.5 standard deviations in assault cases for the highest crime hot spots. The pattern for personal robberies suggest an increase in reported cases in the least secure places, while the effects on

---

than 250 meters. On the other hand, the pure control group for the case when we assume both short and long-range spillovers includes only the 123 streets located at more than 250 meters from treated streets. See Table 2.

<sup>39</sup>We report these heterogeneous effects as do Blattman et al. (2018).

Table 7: Intention to treat effects in the experimental sample, assuming short and long-range spillovers (N=960)

	Control mean (1)	Intent to treat (2)	Short-range spillovers (3)	Long-range spillovers (4)
<i>a. Police crime data</i>				
# of car thefts	0.043	-0.020	-0.030	0.032
		<i>0.499</i>	<i>0.296</i>	<i>0.261</i>
# of motorbike thefts	0.146	0.019	0.024	0.056
		<i>0.861</i>	<i>0.821</i>	<i>0.406</i>
# of personal robberies	0.238	0.035	0.120	0.056
		<i>0.564</i>	<i>0.303</i>	<i>0.926</i>
# of homicides	0.012	0.009	-0.016	-0.005
		<i>0.678</i>	<i>0.881</i>	<i>0.935</i>
# of assaults	0.112	-0.051	-0.041	-0.040
		<i>0.504</i>	<i>0.854</i>	<i>0.487</i>
<i>b. Survey data</i>				
Perception - 6 months, z-score	0.006	0.200	-0.004	-0.070
		<i>0.166</i>	<i>0.973</i>	<i>0.714</i>
Perception - 12 months, z-score	0.139	0.014	-0.062	0.028
		<i>0.935</i>	<i>0.485</i>	<i>0.973</i>
Perception - general, z-score	0.269	-0.105	-0.099	-0.117
		<i>0.259</i>	<i>0.264</i>	<i>0.278</i>
Direct or indirect victimization, z-score	-0.309	0.175	0.146	0.195
		<i>0.107</i>	<i>0.168</i>	<i>0.135</i>
Police - labor, z-score	0.126	0.033	0.014	-0.050
		<i>0.947</i>	<i>0.767</i>	<i>0.601</i>
Police - satisfaction, z-score	-0.038	0.141	0.184	0.094
		<i>0.346</i>	<i>0.270</i>	<i>0.515</i>

*Notes:* Column (1) reports control means, column (2) the coefficient on the intention to treat, column (3) the coefficient on short-range spillovers and column (4) the coefficient on long-range spillovers. We estimate p-values using randomization inference (reported in italics with values <0.1 in bold). All regressions include controls for baseline crime, baseline survey measures and other street characteristics listed in Table 1. Each observation is weighted by the inverse of the probability of being observed in its experimental condition. We exclude 7 streets from the experimental sample that had a probability of treatment equal to 1, resulting from a minor bug in the randomization code.



homicides are noisy, generally close to zero and do not point in a specific direction.

Figure 5 presents the results for outcomes using survey measures. Generally, we see imprecise effects except for the case of security perceptions over the intervention period (six months). In this case, the effect grows from about 0.2 standard deviations when we consider the whole sample of streets except for those in the bottom 10<sup>th</sup> percentile to more than 0.7 standard deviations when we consider only the highest crime hot spots (sub-figure a). The patterns for general perception and perception over a 12 months period point in the direction of a decrease in security in the least secure places, although these effects are not statistically significant at conventional levels (sub-figures b and c). The heterogeneous effects for direct and indirect victimization, and perceptions of police service do not point in any specific direction.

## 5.6 Spillover effects onto non-experimental streets

In Table 8 and Table 9 we explore the presence of spillovers into the non-experimental sample of 36,088 streets. Recall from section 4.2, that we restrict the sample to all street segments that have a positive probability of being exposed to spillovers (at one or two levels, depending on the case) and a positive probability of not being exposed to spillovers (also at one or two levels).<sup>40</sup> We look at spillover effects into the non-experimental sample using police crime data on reported crimes. Table 8 presents the results assuming short-range spillovers within 125 meters, reported in Column (2). Generally, we see no statistically significant spillover effect for any type of crime, with all p-values above 0.35.

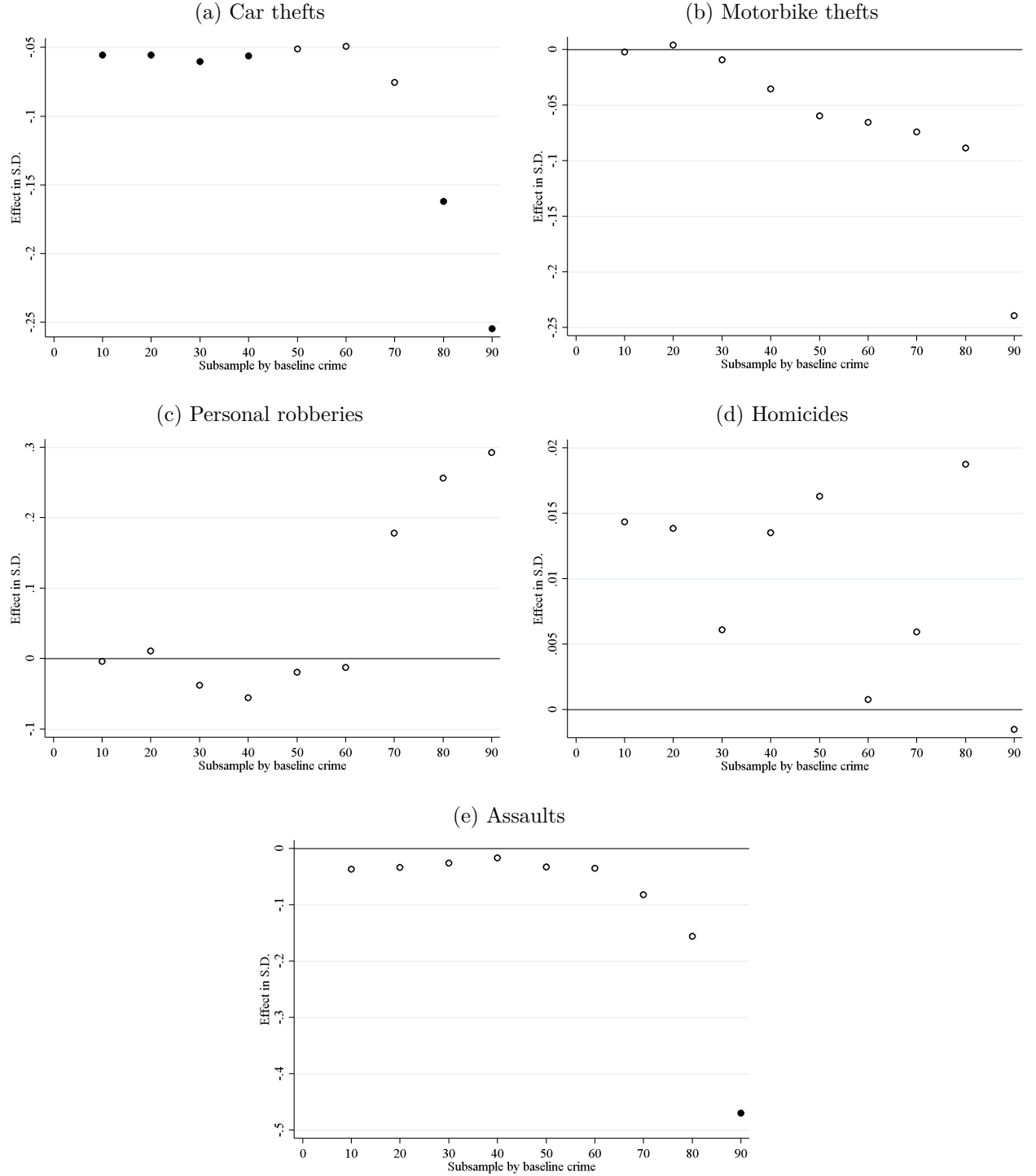
Table 9 presents the results assuming both short and long-range spillover effects. Short-range spillovers are reported in Column (2) while long-range spillovers are reported in Column (3). We find evidence of a decrease of 0.022 reported assaults in non-hot spot streets located within 125 meters of targeted hot spots. This spillover effect is statistically significant at the 10 percent level. Relative to the average number of reported assault cases in the control group, the effect is equivalent to a decrease of about 60 percent. Indeed, when we assume both levels of spillovers, we also see a decrease in reported assault cases in streets located between 125 and 250 meters. This effect is imprecise and does not meet conventional levels of statistical significance. The presence of long-range spillover effects, however, make the results on spillovers within 125 meters in Table 8 and Table 9 consistent.<sup>41</sup> We find no

---

<sup>40</sup>See the notes in each table for details.

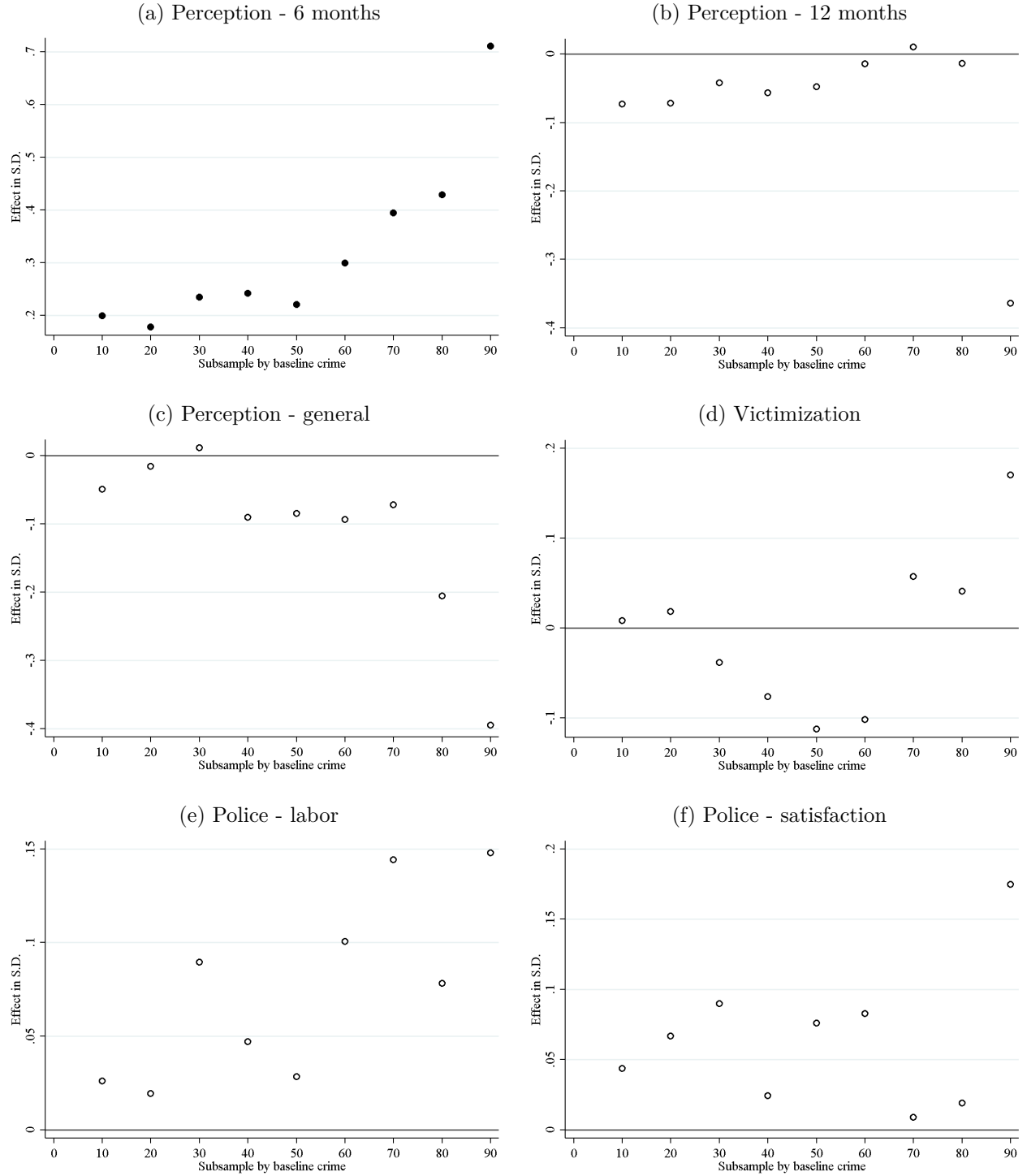
<sup>41</sup>Note that we observe some evidence of long-range positive spillover effects when we assume both levels of spillovers. As a result, the decrease in reported assault cases in streets located 125 and 250 meters lead us to under-estimate the short-range spillover effects when we only assume short-range spillovers. Recall that when we assume only short-range spillover effects, streets located between 125 and 250 meters are in the control group.

Figure 4: Heterogeneous treatment effects using police crime data



*Notes:* The figure depicts point estimates for heterogeneous treatment effects using police crime data. Each circle corresponds to the intent to treat effect in a regression with a restricted sample of hot spots, measured in standard deviations. In each case, the sample includes hot spots at the corresponding percentile or higher based on baseline crime levels. For instance, the circle at 50 for car thefts includes hot spots that are at the 50<sup>th</sup> percentile or above based on car theft levels in 2014. We estimate the effects using equation (1) assuming short-range spillovers. Filled in figures have randomization inference p-values < 0.1.

Figure 5: Heterogeneous treatment effects using survey data



*Notes:* The figure depicts point estimates for heterogeneous treatment effects using survey data. Each circle corresponds to the intent to treat effect in a regression with a restricted sample of hot spots, measured in standard deviations. In each case, the sample includes hot spots at the corresponding percentile or higher based on baseline levels. For instance, the circle at 50 for security perceptions within six months includes hot spots that are at the 50<sup>th</sup> percentile or above based on the same question in the baseline survey. We estimate the effects using equation (1) assuming short-range spillovers. Filled in figures have randomization inference p-values  $< 0.1$ .

Table 8: Spillover effects in the non-experimental sample, assuming short-range spillovers (N=14,695)

	Control mean (1)	Short-range spillovers (2)
# of car thefts	0.014	-0.003 <i>0.433</i>
# of motorbike thefts	0.059	0.000 <i>0.979</i>
# of personal robberies	0.107	-0.015 <i>0.772</i>
# of homicides	0.006	0.001 <i>0.454</i>
# of assaults	0.041	-0.007 <i>0.353</i>

*Notes:* Column (1) reports control means and column (2) the coefficient on short-range spillovers. We estimate p-values using randomization inference (reported in italics with values <0.1 in bold). All regressions include controls for baseline crime and other street characteristics listed in Table 1. Each observation is weighted by the inverse of the probability of being observed in its experimental condition. The sample includes streets with a positive probability of being within 125 meters of targeted streets, and a positive probability of not being within 125 meters of targeted streets.

statistically significant short or long-range spillover effects for car and motorbike thefts, personal robberies and homicides.

## 5.7 Aggregate effects

Even if average spillover effects are minuscule, when a large number of streets is exposed to spillovers aggregate effects add up. In this Section, we follow Blattman et al. (2018) and perform a back-of-the-envelope estimate of aggregate effects citywide. We focus on direct and spillover effects on the experimental and non-experimental samples, as well as the number of streets falling in each of these conditions.

Table 10 presents the results assuming short-range spillovers (assuming short and long-range spillovers renders similar results). Each panel presents results for a different type of crime. In each case, we multiply the estimated coefficient by the number of streets falling in each condition, and add direct and spillover effects to get an estimate of aggregate effects citywide. Even if each independent coefficient is not statistically significant, it represents

Table 9: Spillover effects in the non-experimental sample, assuming short and long-range spillovers (N=11,501)

	Control mean (1)	Short-range spillovers (2)	Long-range spillovers (3)
# of car thefts	0.013	-0.004 <i>0.497</i>	-0.002 <i>0.778</i>
# of motorbike thefts	0.051	0.008 <i>0.559</i>	0.002 <i>0.834</i>
# of personal robberies	0.059	-0.009 <i>0.778</i>	0.012 <i>0.455</i>
# of homicides	0.004	0.000 <i>0.937</i>	0.002 <i>0.654</i>
# of assaults	0.036	-0.022 <b><i>0.053</i></b>	-0.019 <i>0.108</i>

*Notes:* Column (1) reports control means, column (2) the coefficient on short-range spillovers and column (3) the coefficient on long-range spillovers. We estimate p-values using randomization inference (reported in italics with values <0.1 in bold). All regressions include controls for baseline crime and other street characteristics listed in Table 1. Each observation is weighted by the inverse of the probability of being observed in its experimental condition. The sample includes streets with a positive probability of being within 125 meters of targeted streets, a positive probability of not being within 125 meters of targeted streets, a positive probability of being between 125 and 250 meters from targeted streets, and a positive probability of not being between 125 and 250 meters of targeted streets.

our best guess of what really happened. We also estimate a 90 percent confidence interval using randomization inference.

*Panel a* presents estimates for car thefts. We estimate that the intervention led to a decrease of about 55 cases citywide. This is a reduction of 11 percent, relative to the total number of reported car thefts in the city during the intervention period. The 90 percent confidence interval includes zero. Nonetheless, the upper limit is marginally above zero, which suggests that most likely there was a reduction in car thefts in the city resulting from the intervention. Indeed, we cannot rule out a decrease as large as 23 percent citywide, given the lower bound of the confidence interval. This effect is explained by both direct treatment effects and beneficial spillovers onto experimental and non-experimental streets. In particular, note that very small average beneficial spillovers onto non-experimental streets (0.003 crimes per street) led to a reduction of about 24 car thefts in total.

*Panels b* through *e* present estimates for motorbike thefts, personal robberies, homicides and assaults, respectively. The aggregate estimates suggest there was a decrease in reported cases citywide, except for homicides, where we observe an increase of about 14 cases (5 percent relative to the total number of homicides in the city during the intervention period). For motorbike thefts the reduction equals 1 percent relative to the total number of cases, for personal robberies it equals 3 percent and for assault cases it equals 5 percent. In all these cases, the confidence intervals are wide enough to include zero, leaving a larger level of uncertainty. Perhaps, the exception is on assault cases, where we can confidently rule out an increase as small as 1 percent in citywide cases, and the results generally point to a citywide decrease.

## 6 Discussion and conclusions

Regarding the direct treatment effects of the program, our study adds nuisance to the U.S. based research, as we only find positive results for a specific type of crime: car thefts. We can only speculate why this is the case. Medellín is disproportionately affected by this crime relative to other Colombian cities, hence this type of crime can be somewhat more responsive to an increase in police presence. Compared to the closer study conducted by Blattman et al. (2018) in Bogotá, our direct treatment effects show generally larger impacts for property crimes (specifically for car thefts), but we are less optimistic on the program's results regarding violent crimes, for which they do find effects both in targeted streets and surrounding areas.<sup>42</sup> We note, however, that there are large contextual differences between Medellín

---

<sup>42</sup>They find smaller effects for property crime in general, so the comparison is not direct for the exact same type of crime.

Table 10: Back-of-the-envelope estimation of aggregate effects

	Coeff. (1)	RI p-value (2)	# of segments (3)	Total impact = (1) x (3) (4)	90% CI (5)
<i>a. Car thefts</i>					
Direct treatment effects	-0.044	0.060	377	-16.724	
Spillovers (experimental)	-0.055	0.043	271	-14.815	
Spillovers (non-experimental)	-0.003	0.433	8,630	-23.670	
				-55.208	(-114.6, 2.7)
<i>b. Motorbike thefts</i>					
Direct treatment effects	-0.027	0.453	377	-10.182	
Spillovers (experimental)	-0.022	0.564	271	-6.074	
Spillovers (non-experimental)	-0.000	0.979	8,630	-3.527	
				-19.784	(-149.4, 107.4)
<i>c. Personal robberies</i>					
Direct treatment effects	-0.029	0.729	377	-10.795	
Spillovers (experimental)	0.066	0.376	271	17.825	
Spillovers (non-experimental)	-0.015	0.772	8,630	-130.833	
				-123.804	(-552.7, 175.0)
<i>d. Homicides</i>					
Direct treatment effects	0.013	0.306	377	4.762	
Spillovers (experimental)	-0.012	0.949	271	-3.277	
Spillovers (non-experimental)	0.001	0.454	8,630	12.610	
				14.094	(-20.5, 53.0)
<i>e. Assaults</i>					
Direct treatment effects	-0.025	0.889	377	-9.409	
Spillovers (experimental)	-0.013	0.765	271	-3.619	
Spillovers (non-experimental)	-0.007	0.353	8,630	-59.860	
				-72.887	(-183.3, 20.4)

*Notes:* Column (1) reports estimated coefficients, column (2) the corresponding RI p-value, column (3) the number of segments that fall under each condition, column (4) the estimated total impact and column (5) the 90% confidence interval. We estimate the confidence intervals using randomization inference. We simulate 1,000 randomizations to get the distribution of the estimated aggregate effects. The 5 and 95 percentiles of this distribution give us the 90% confidence interval.

and Bogotá. For instance, the extent of control by criminal organizations in Medellín is much larger, and thus crime is more planned and instrumental than it is in Bogotá. Also, the police to population ratio is 60 percent larger in Medellín than it is in Bogotá.

The spillover effects are also concentrated mainly on car thefts, so our interpretation is similar. Our findings are less optimistic than U.S. based hot spots policing research, that generally points to positive spillovers for different types of crime. Since we did not find direct program effects for other crimes, we could not expect any diffusion of benefits. These results also differ from the Bogotá study by Blattman et al. (2018), as they find evidence of large negative spillovers for property crime. They also find evidence of positive spillovers on violent crimes. The difference, again, can be driven by contextual differences. In particular, having more police manpower citywide can be crucial to prevent the negative spillovers on thefts.

Our back-of-the-envelope estimation of aggregate effects (as the estimation by Blattman et al. (2018) for the Bogotá study) sheds light on the importance of accounting for the large number of places that are exposed to crime spillovers (and to include non-experimental places in the spillover analysis). Generally, the estimation of spillovers should not be a matter of average effects only, but rather include the large number of streets located in the surroundings of treatment areas. We saw, for instance, that very small beneficial spillovers of 0.003 car thefts, on average (obviously, non statistically significant), led to a decrease of about 24 reported cases citywide. The explanation is simple: more than 8,000 non-experimental streets were exposed to spillovers.

This study, more generally, sheds light on police incentives. Police patrols in Colombia are instructed to intensify their presence in crime hot spots. However, these instructions are generally not met—and indeed that was the case of Medellín before and after the intervention. This can be a result of simply misunderstanding the directives, lack of information and knowledge on where and when crimes occur, lack of personnel to deploy to these places when there is information, or simply lack of willingness because patrols cooperate with criminals. It can also be a result of a problem of incentive compatibility. For instance, since police performance is measured mainly through crime reports and operational results in a given jurisdiction for a given period of time, that police patrols care more about petty crimes outside major crime hot spots simply because when doing so they can meet performance goals easily.

In any case, when the intervention was implemented and police patrols were informed they were going to be closely monitored, they effectively complied and intensified patrolling time in crime hot spots. At the moment the monitoring stopped, compliance dropped sharply regardless of the general directive of intensifying patrolling efforts at crime hot spots. Even



if the intervention can be deemed successful because of the results on car thefts and security perceptions, police patrols deliberately stopped. Generally, we believe there is need for more interventions and experimentation with police incentives so that the compliance with specific instructions does not rely exclusively on monitoring and enforcement, and the public concerns—rather than only those of the police—are generally accounted for. From the broader perspective, such experimentation can contribute to understand better principal-agent problems of control over bureaucrats.

## References

- Abt, T. and C. Winship (2016). *What Works in Reducing Community Violence: A Meta Analysis and Field Study for the Northern Triangle*. USAID.
- Becker, G. (1968). Crime and punishment: An economic approach. *Journal of Political Economy* 76(2), 169–217.
- Blattman, C., D. Green, D. Ortega, and S. Tobon (2018). Place-based interventions at scale: The direct and spillover effects of policing and city services on crime. *NBER Working Paper No. 23941*.
- Bowers, K. J., S. D. Johnson, R. T. Guerette, L. Summers, and S. Poynton (2011). Spatial displacement and diffusion of benefits among geographically focused policing initiatives: A meta-analytical review. *Campbell Systematic Reviews* (3).
- Braga, A. and B. J. Bond (2008). Policing crime and disorder hot spots: A randomized controlled trial. *Criminology* 46, 577–608.
- Braga, A., A. V. Papachristos, and D. M. Hurreau (2012). An ex post factor evaluation framework for place-based police interventions. *Campbell Systematic Reviews* 8, 1–31.
- Braga, A. A., A. V. Papachristos, and D. M. Hureau (2014). The effects of hot spots policing on crime: An updated systematic review and meta-analysis. *Justice Quarterly* 31(4), 633–663.
- Braga, A. A., D. L. Weisburd, E. J. Waring, L. G. Mazerolle, W. Spelman, and F. Gajewski (1999). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology* 37(3), 541–580.
- Chong, A., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2014). Letter grading government efficiency. *Journal of the European Economic Association* 12(2), 277–299.

- Cohen, J., W. Gorr, and P. Singh (2003). Estimating intervention effects in various settings: Do police raids reduce illegal drug dealing at nuisance bars? *Criminology* 41(2), 257–292.
- Di Tella, R. and E. Schargrodsky (2004, March). Do Police Reduce Crime? Estimates Using the Allocation of Police Forces After a Terrorist Attack. *American Economic Review* 94(1), 115–133.
- Draca, M., S. Machin, and R. Witt (2011). Panic on the streets of london: Police, crime, and the july 2005 terror attacks. *The American Economic Review* 101(5), 2157–2181.
- Duncan, G., A. Eslava, J. G. Ramírez, L. F. Dávila, L. Gutiérrez, F. Lopera, J. P. Mesa, J. Toro, and P. Zapata (2015). *Territorio, crimen, comunidad: Heterogeneidad del homicidio en Medellín* (1 ed.). Medellín: Centro de Análisis Político Universidad EAFIT - Open Society Foundations.
- Dunn, O. J. (1961). Multiple Comparisons Among Means. *Journal of the American Statistical Association* 56(293), 52–64.
- Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation. *Journal of Political Economy* 81(3), 521.
- Gerber, A. S. and D. P. Green (2012). *Field experiments: Design, analysis, and interpretation*. New York: WW Norton.
- Giraldo, J., A. J. Rendón, G. Duncan, C. Arango, S. Bohórquez, A. M. Corpas, L. Gallego, and A. F. Preciado (2014). *Nuevas modalidades de captación de rentas ilegales en Medellín* (1 ed.). Medellín: Centro de Análisis Político Universidad EAFIT - Empresa para la Seguridad Urbana - Alcaldía de Medellín.
- Groff, E. R., J. H. Ratcliffe, C. P. Haberman, E. T. Sorg, N. M. Joyce, and R. B. Taylor (2015). Does what police do at hot spots matter? The philadelphia policing tactics experiment. *Criminology* 53(1), 23–53.
- Lawton, B. a., R. B. Taylor, and A. J. Luongo (2005). Police Officers on Drug Corners in Philadelphia, Drug Crime, and Violent Crime: Intended, Diffusion, and Displacement Impacts. *Justice Quarterly* 22(4), 427–451.
- Mazerolle, L. G., J. F. Price, and J. Roehl (2000). Civil Remedies and Drug Control. A Randomized Field Trial in Oakland, California. *Evaluation review* 24(2), 212–241.
- McDermott, J. (2014). El rostro cambiante del crimen organizado colombiano. *Perspectivas* 9, 1–12.

- Mejía, D., D. Ortega, and K. Ortiz (2015). Un análisis de la criminalidad urbana en Colombia. Technical report, CAF - Banco de Desarrollo de America Latina.
- Petrosino, A., C. Turpin-Petrosino, M. Hollis-Peel, and J. Lavenberg (2003). Scared Straight and Other Juvenile Awareness Programs for Preventing Juvenile Delinquency: A Systematic Review of the Randomized Experimental Evidence. *Campbell Systematic Reviews* (5).
- Police Executive Research Forum (2008). Violent crime in America: What we know about hot spots enforcement. Technical report, Police Executive Research Forum, Washington, DC.
- Ratcliffe, J. H., T. Taniguchi, E. R. Groff, and J. D. Wood (2011). The Philadelphia foot patrol experiment: A randomized controlled trial of police patrol effectiveness in violent crime hotspots. *Criminology* 49(3), 795–831.
- Santos, R. B. and R. G. Santos (2016). Offender-focused police intervention in residential burglary and theft from vehicle hot spots: a partially blocked randomized control trial. *Journal of Experimental Criminology*, 1–30.
- Sherman, L., M. Buerger, and P. Gartin (1989). Beyond dial-a-cop: A randomized test of repeat call policing (recap). Washington, D.C.: Crime Control Institute.
- Sherman, L., S. Williams, A. Barak, L. R. Strang, N. Wain, M. Slothower, and A. Norton (2014). An Integrated Theory of Hot Spots Patrol Strategy: Implementing Prevention by Scaling Up and Feeding Back. *Journal of Contemporary Criminal Justice* 30(2), 95–122.
- Sherman, L. W., P. R. Gartin, and M. E. Brueger (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology* 27(1), 27–56.
- Sherman, L. W., D. P. Rogan, T. Edwards, R. Whipple, D. Shreve, D. Witcher, W. Trimble, The Street Narcotics Unit, R. Velke, M. Blumberg, A. Beatty, and C. a. Bridgeforth (1995). Deterrent effects of police raids on crack houses: A randomized, controlled experiment. *Justice Quarterly* 12(4), 755–781.
- Sherman, L. W. and D. Weisburd (1995). General deterrent effects of police patrol in crime hot spots: A randomized, controlled trial. *Justice Quarterly* 12(4), 625–648.
- Sviridoff, M., S. Sadd, R. Curtis, and R. Grinc (1992). The Neighbourhood Effects of Street-Level Drug Enforcement: Tactical Narcotics Teams in New York. Technical report, Vera Institute of Justice, New York.

- Taylor, B., C. S. Koper, and D. J. Woods (2011). A randomized controlled trial of different policing strategies at hot spots of violent crime. *Journal of Experimental Criminology* 7(2), 149–181.
- Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology* 53(2), 133–157.
- Weisburd, D., D. Farrington, C. Gill, M. Ajzenstadt, T. Bennett, K. J. Bowers, M. S. Caudy, K. Holloway, S. Johnson, F. Lösel, J. Mallender, A. Perry, L. L. Tang, F. Taxman, C. Telep, R. Tierney, M. M. Ttofi, C. Watson, D. B. Wilson, and A. Wooditch (2017). What Works in Crime Prevention and Rehabilitation: An Assessment of Systematic Reviews. *Criminology and Public Policy* 16(2), 1–35.
- Weisburd, D. and L. Green (1995). Policing drug hot spots: The Jersey City drug market analysis experiment. *Justice Quarterly* 12(4), 711–735.
- Weisburd, D., D. Groff, and S. Yang (2012). *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*. New York: Oxford University Press.
- Weisburd, D. and C. Telep (2016). Hot Spots Policing: What We Know and What We Need to Know. *Journal of Experimental Criminology* 30(2), 200–220.