Place-based interventions at scale: The direct and spillover effects of policing and city services on crime

Chris Blattman
U Chicago & NBER

Donald Green
Columbia U

Daniel Ortega
CAF

Santiago Tobón
U Chicago & IPA

This is a study of a city's use of everyday state presence—police patrols, municipal services—to tackle crime on moderate to high-crime streets

2016: Bogota undertook a large-scale experiment

Newly elected Mayor pledges to tackle the 750 highest-crime streets and intensify two state services:

1. Increase police patrol time from ~90 minutes a day to nearly 3 hours
2. Instruct city contractors to deliver additional garbage cleanup, light repair, and tree pruning services

- A reallocation of existing state resources to moderate and high-crime street segments
- No new police or contractors added

EL TIEMPO

BOGOTÁ

El plan de Enrique Peñalosa para los primeros cien días de mandato

El candidato a la alcaldía anunció su plan de choque que ejecutaría en su Gobierno.

Por: BOGOTÁ |
9:07 p.m. | 1 de octubre de 2015

Enrique Peñalosa, de Equipo por Bogotá, presentó sus 10 puntos por la seguridad y la movilidad que ejecutará si gana las elecciones.

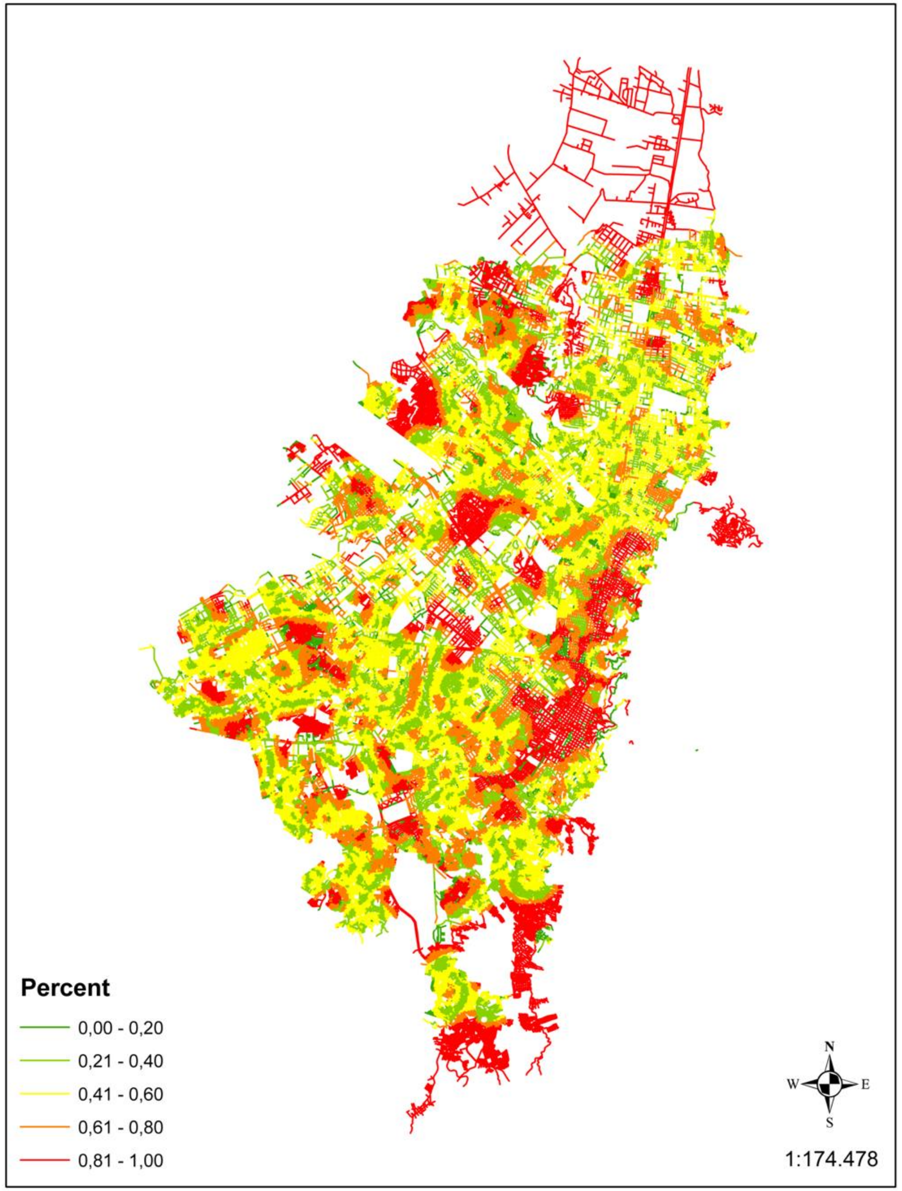
Seguridad

1. Intervenir de manera inmediata 750 puntos calientes del crimen, con más presencia de Policía, operativos, iluminación y limpieza de calles de basuras y grafitis.

Scale brings both opportunities and challenges

- Increased statistical power to estimate:
  - Subtle spillovers
  - Differential effects on types of crime
  - Differential impacts by crime level
- But spillovers in a dense network can bias estimation:
  - Bias treatment effects if not properly accounted for
  - Understate standard errors because of difficult-to-models patterns of “fuzzy clustering” of control and spillover regions

P(segments within 500m have the same treatment assignment as you)



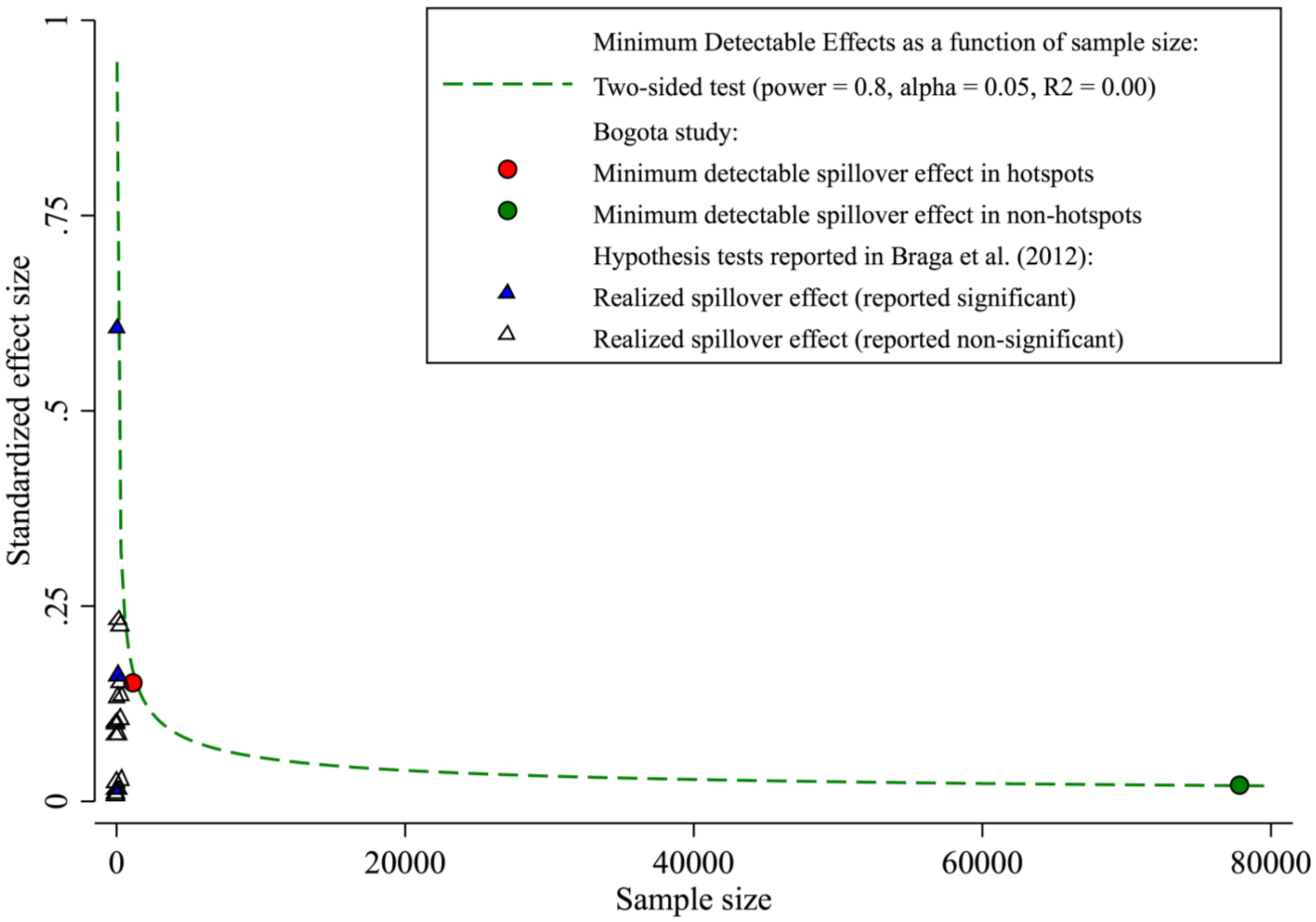
Whether crime is displaced or deterred has both policy and theoretical ramifications

- Policy-wise, the degree of deterrence is central to any cost-benefit analysis
- Theoretically, if place-based interventions have no or beneficial spillovers, it implies at least one of the following:
  A. Non-motivated, non-economic roots of many offenses
  B. For crimes with a sustained motive (e.g. professional theft):
     1. Criminal rents are concentrated, immobile, and unequally distributed within cities
     2. Supply of crime is highly elastic to the probability of detection and apprehension in a small number of high-profit areas
     3. Some offenders are resistant to moving crime locations
- The balance of evidence from U.S. studies tilts towards no or beneficial spillovers

But spillovers are not a matter of average effects, but of the aggregation of those very small (and most likely undetectable) average effects

- Previous literature on hot spots policing illustrates this point
  - There are often thousands of nearby segments outside the experimental sample
  - Very small adverse spillovers will be hard to detect with precision
  - If aggregate effects are important, insignificant results cannot be disregarded so easily

Realized spillover effects for previous studies, minimum detectable effects for the Bogotá experiment

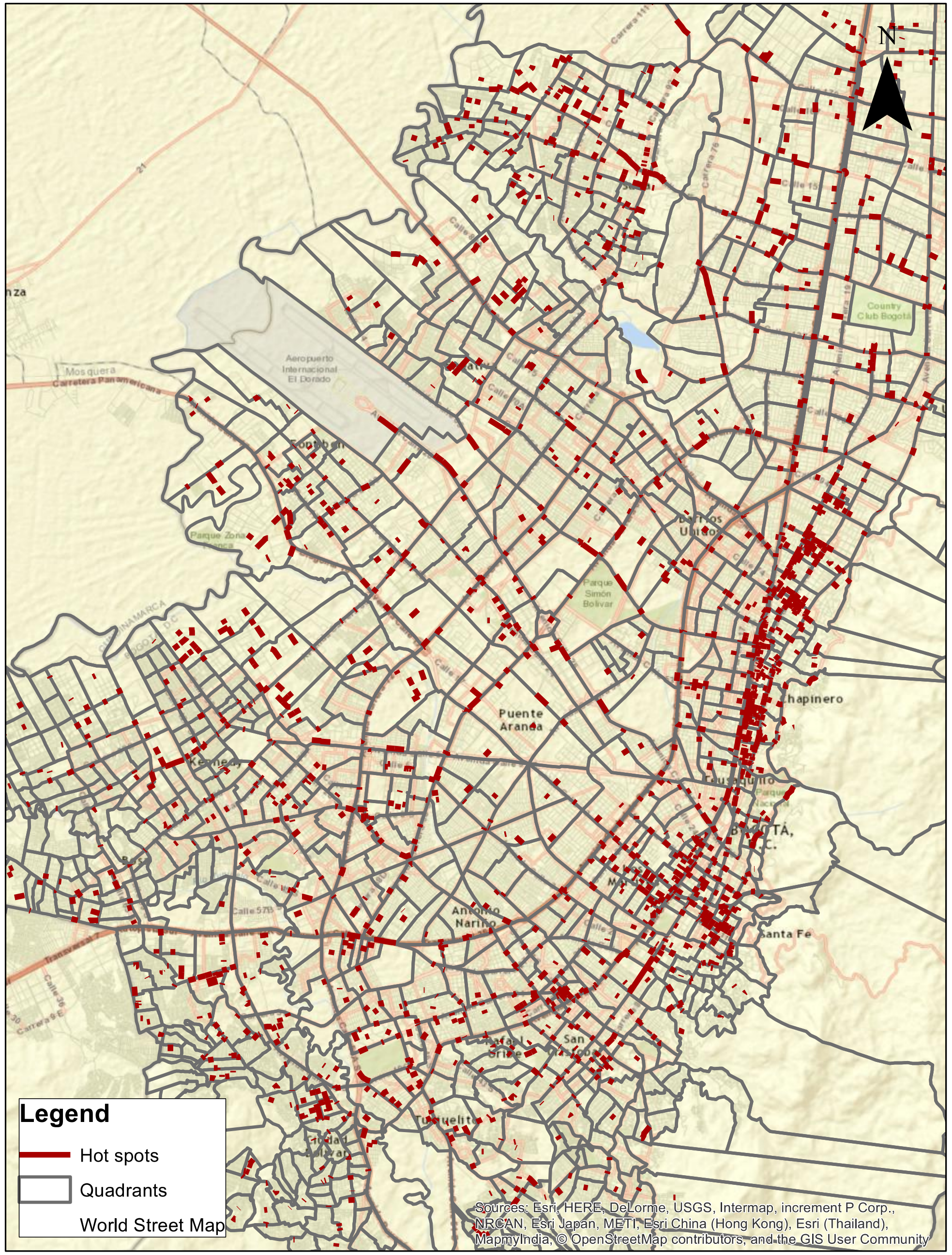


Preview of results

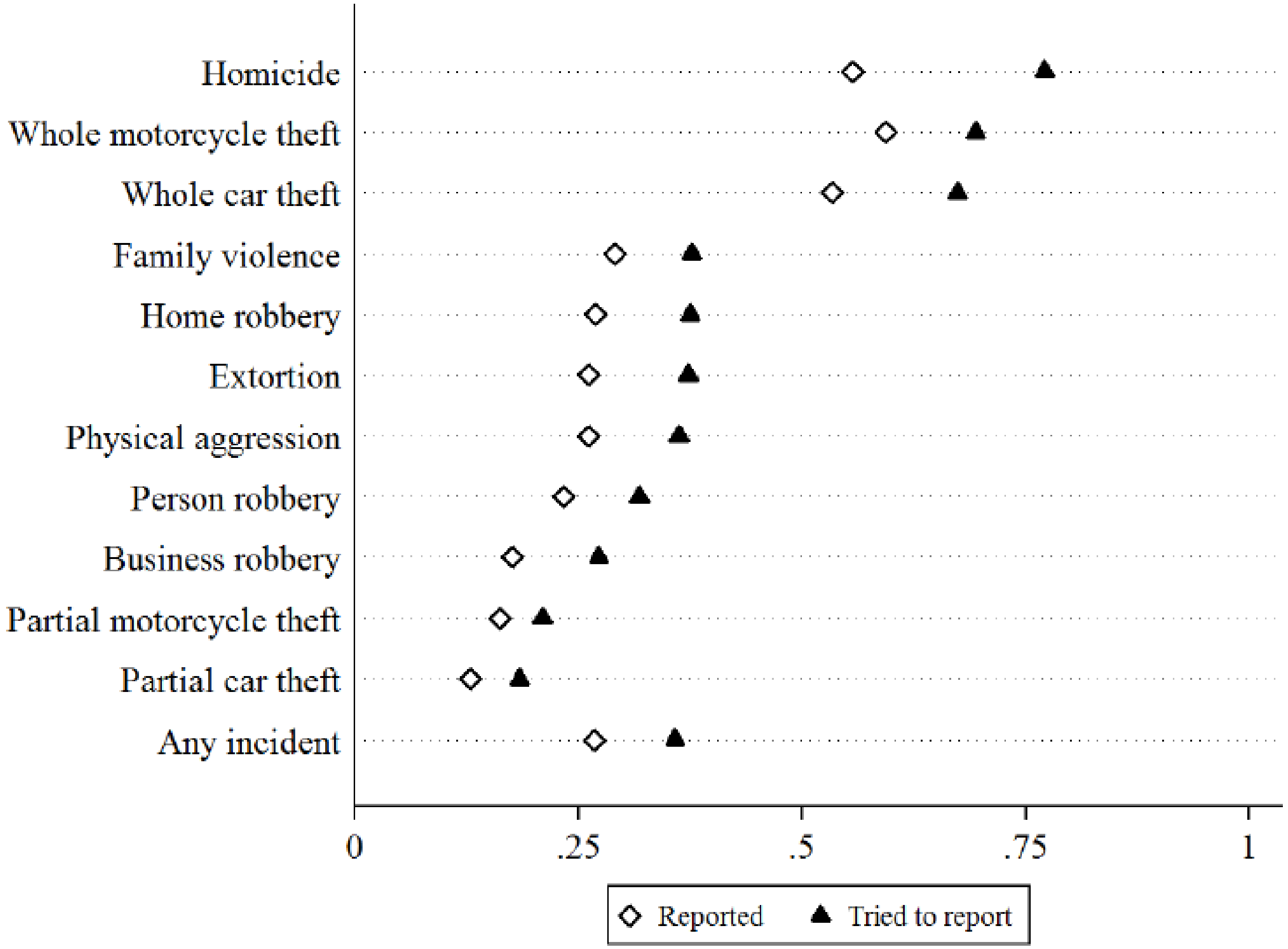
1. At scale, standard designs and inference lead to biased and misleadingly precise results
2. Demonstrate how a design-based approach and randomization inference can correct for bias and hard-to-model patterns of clustering
3. Increasing state presence has at best modest and imprecise direct impacts
4. Both interventions lead to more substantial declines in crime, especially in highest crime streets
5. Adverse spillovers: crime appears to rise in neighboring streets
6. In aggregate, we can rule out a citywide reduction of more than 2-3% in total crimes
7. More promising, adverse spillovers are driven by property crime and the evidence suggests homicides and rapes may have decreased by about 5% citywide

Selecting the experimental sample

- We used 2012-15 data to identify top 2% (2,720) segments
- Main issues:
  1. Most petty crimes and many major crimes not reported
  2. Some crimes assigned to wrong street



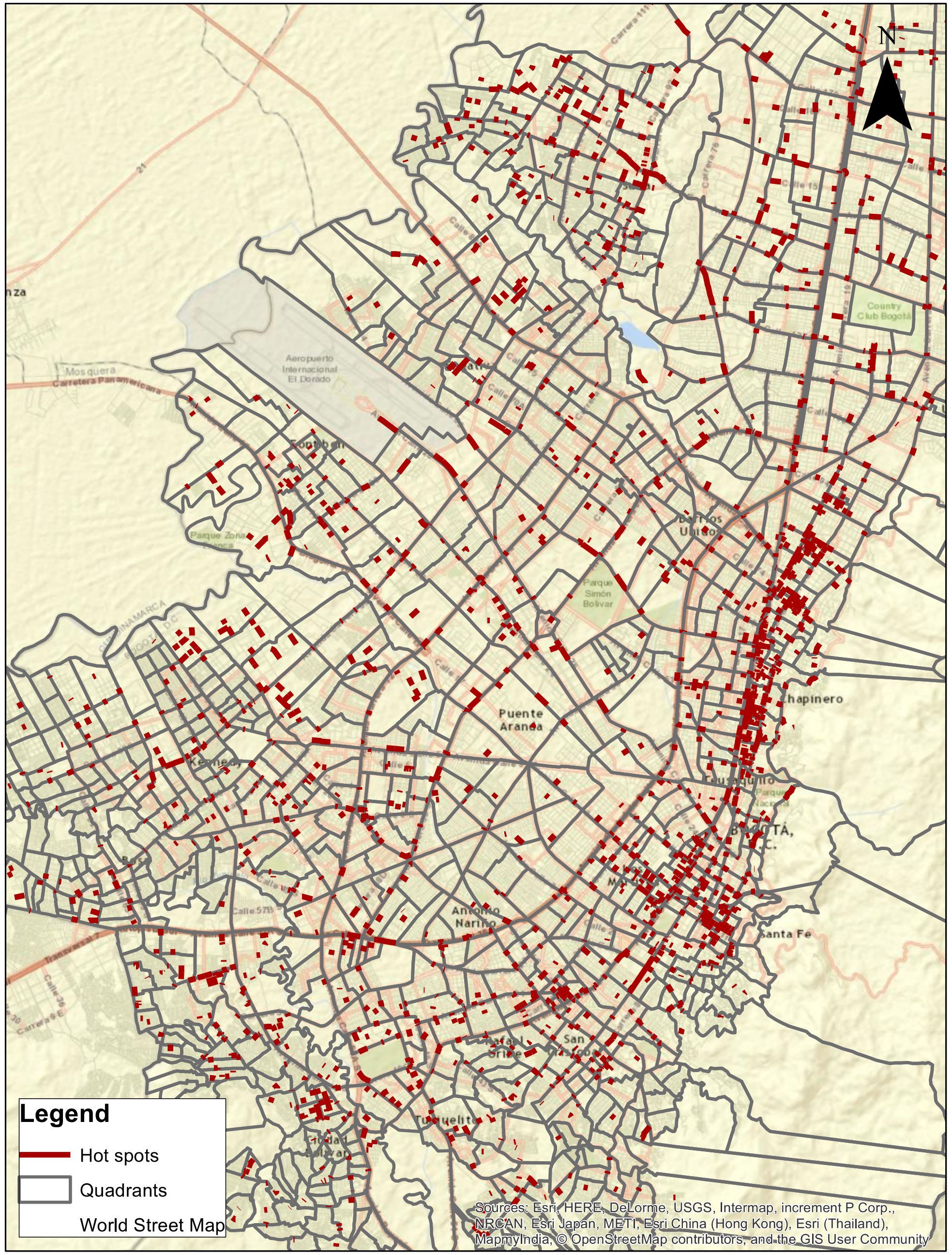
Under-reporting of crime, based on a survey of 24,000 residents of Bogota



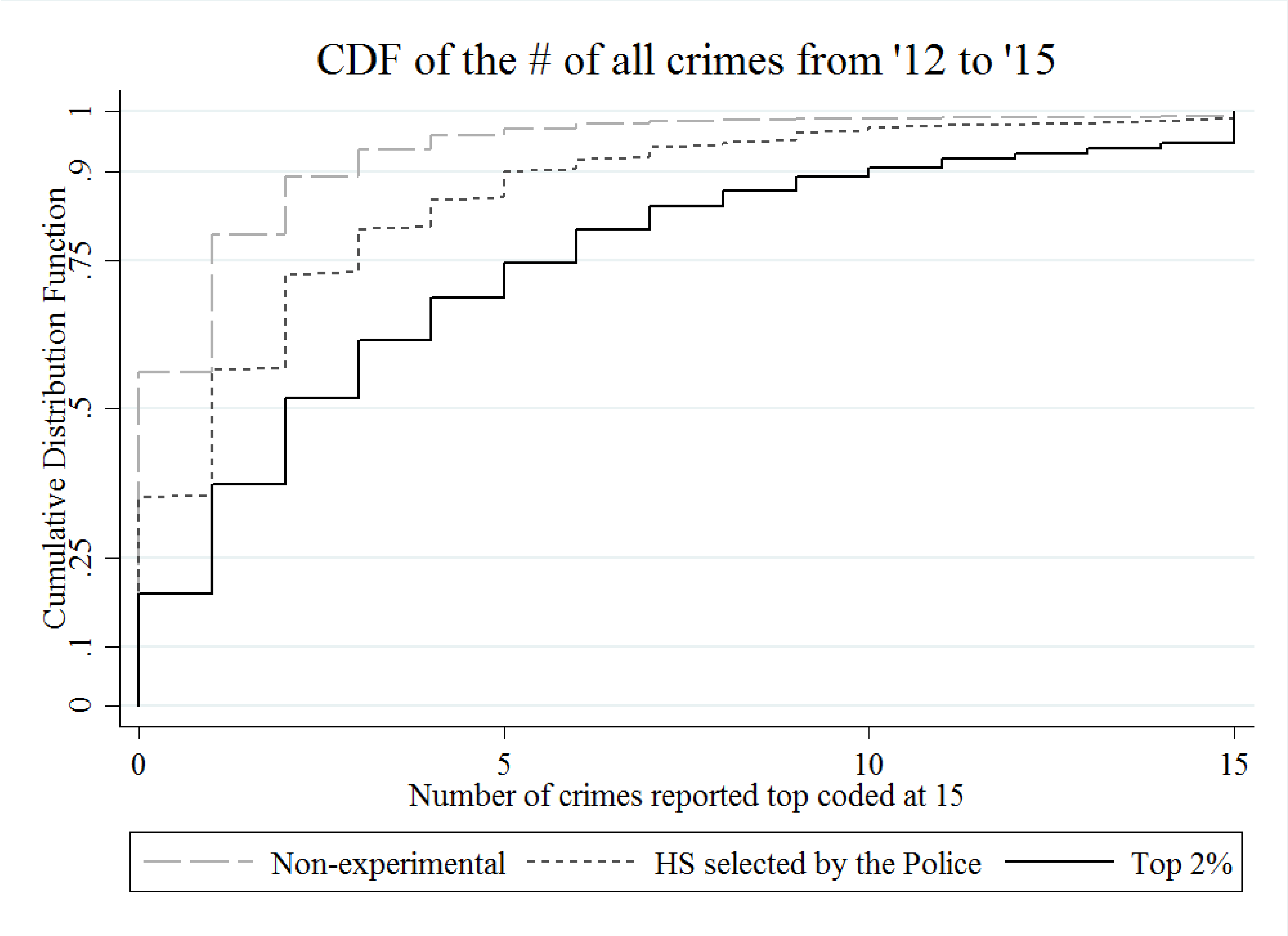
Note: Post-treatment survey, where treatment uncorrelated with underreporting

Selecting the experimental sample

- We used 2012-15 data to identify top 2% (2,720) segments
- Main issues:
  1. Most petty crimes and many major crimes not reported
  2. Some crimes assigned to wrong street
- Thus validated with police patrols
  - They discarded some streets
  - Added some, low on reported crime
- We arrived at 1,919 segments with >60,000 ”non-experimental” segments within 250m

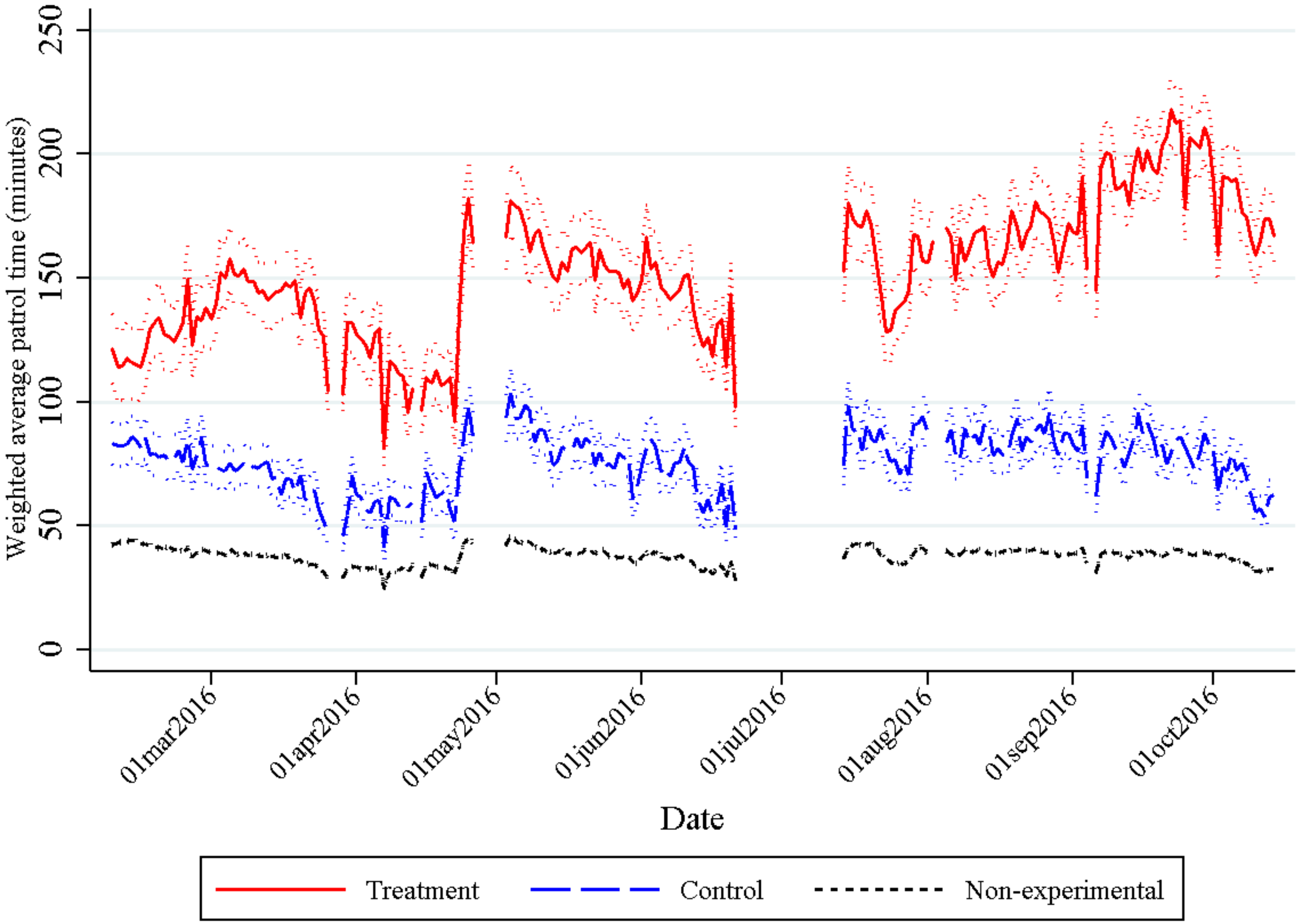


Leads to a sample that likely includes both moderate and high-crime streets



Intervention 1: Increase normal patrolling duties from 92 to 169 mins/day

Average patrolling minutes by treatment status, measured with GPS devices reporting patrol locations every 30 seconds



Intervention 2:
Deliver municipal clean-up and maintenance services to up to 201 hot spots

Ex ante we were less optimistic about this intervention because:

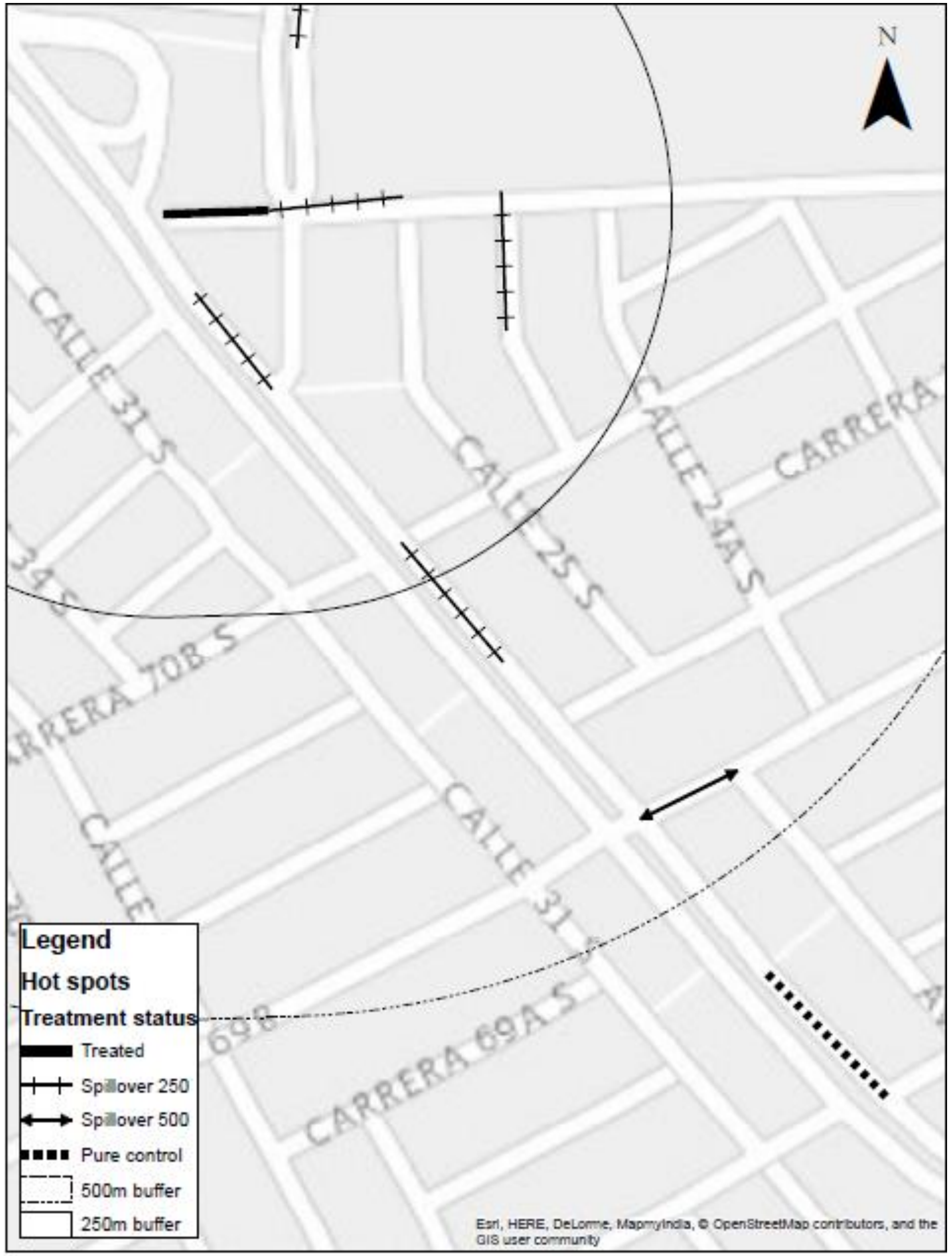
- Sample was smaller
- Not all experimental streets appeared to need much maintenance
- Compliance by city contractors was moderate

Spillovers in dense networks complicate treatment effects estimation

1. No longer possible to examine distant, unrelated treatment and control segments
   - Several possible violations of the assumption of no interference between units
   - Failing to account for spillovers will bias treatment effects
2. Differential probabilities of assignment to treatment arms
   - Some streets have a higher probability of assignment to treatment, spillover or control status
   - These differences are correlated with unobservables
3. Differential probabilities of spillover and control status also lead to hard-to-model patterns of “fuzzy clustering”
   - Will generally lead us to understate standard errors and can lead to bias in small to moderate samples as well

We first take a design-based approach to flexibly estimate spillovers

- 2-stage randomization to smooth probabilities of spillovers and ensure a control group
    1. Assign quadrants to treatment or control
    2. Assign segments to police treatment in treatment quadrants
- Then, assign municipal services treatment blocking on police treatment and eligibility
- Partition control segments according to distance from treated segments:
    - <250 meters, 250–500 meters, >500 meters
- Estimate treatment and spillover effects by comparing means across experimental conditions
- Follow a pre-specified rule for determining whether the spillover region is 250 or 500m

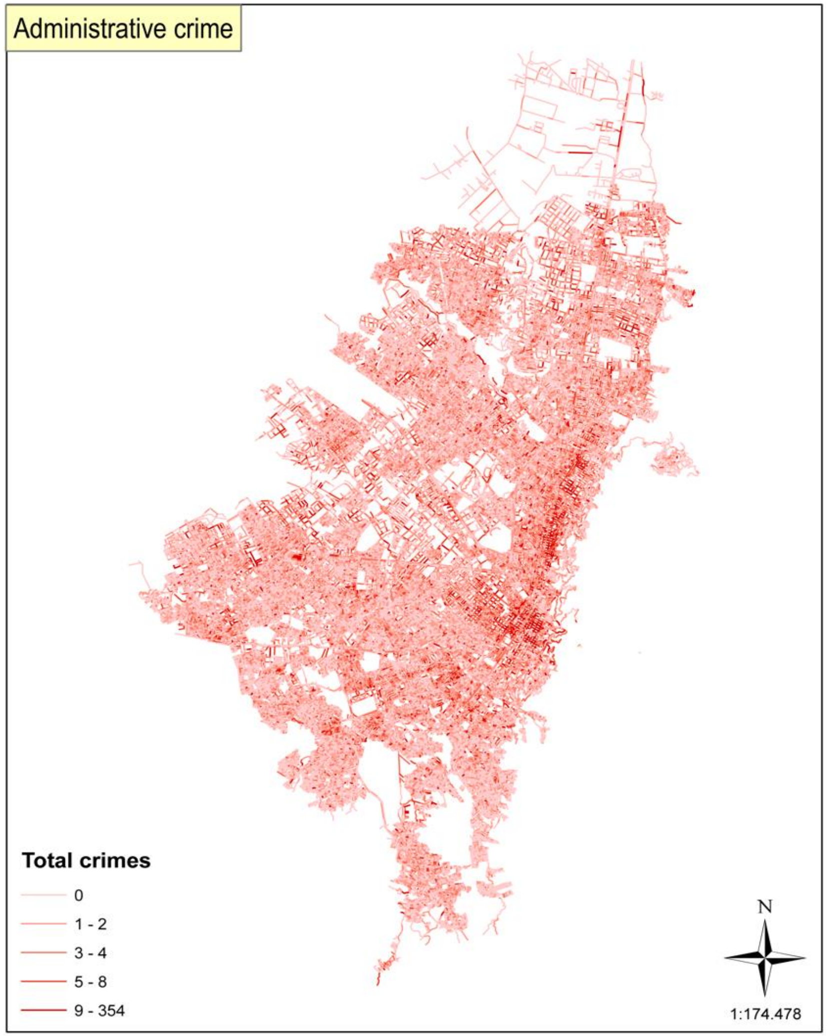


Random assignment produced the expected degree of balance along covariates

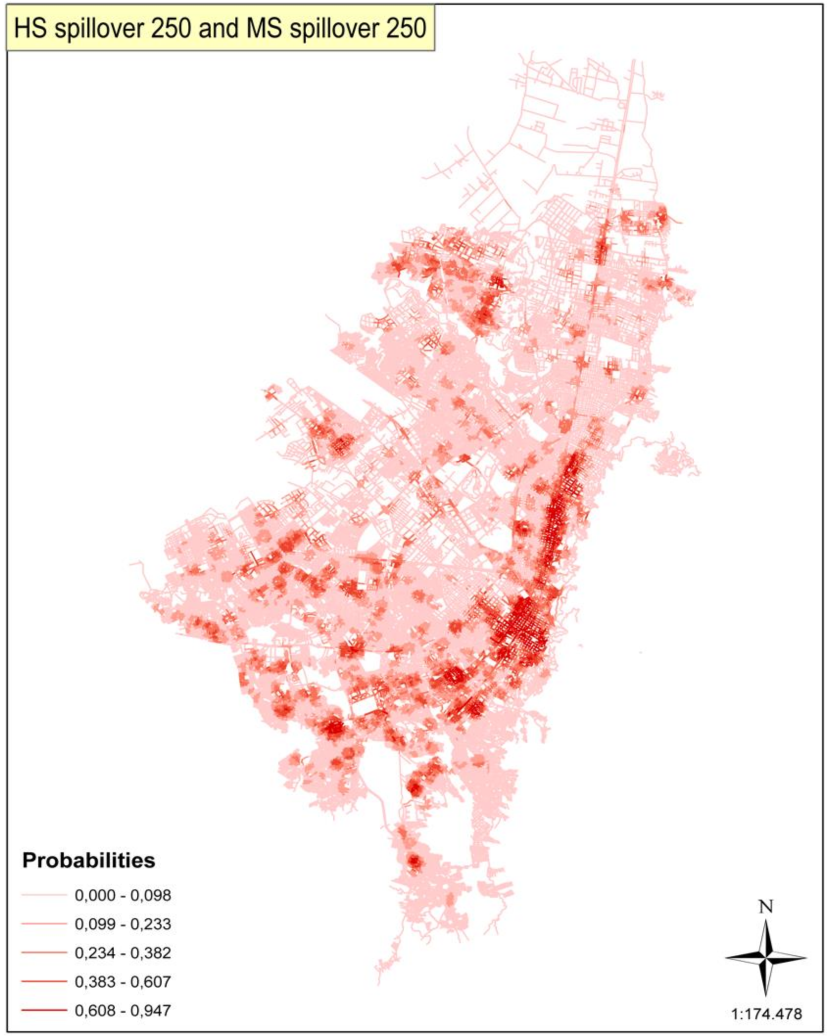
	Summary statistics			Balance test			
				Intensive policing		Municipal services	
	Mean	Std. Dev.	Max.	Coeff	p-val	Coeff	p-val
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Crimes reported per segment, 2012–15 (original)	4.53	5.72	82	-0.18	0.68	-0.14	0.89
# of violent crimes	1.88	2.94	56	-0.20	0.37	-0.06	0.89
# of property crimes	2.66	3.97	50	0.03	0.96	-0.08	0.91
Crimes reported per segment, 2012–15 (updated)	5.18	18.24	461	-0.35	0.86	-0.18	0.91
# of violent crimes	1.40	5.38	78	0.33	0.28	0.33	0.23
# of property crimes	3.78	14.09	407	-0.69	0.61	-0.50	0.81
Patrol minutes per day (11/2015–01/2016)	38.03	70.27	1029	-2.61	0.69	3.02	0.56
Rating of baseline disorder (0–5)	1.18	0.74	5	-0.02	0.48	0.08	0.11
Meters from police station or CAI	551.37	351.46	2805	-25.70	0.24	-1.80	0.79
Zoned for industry/commerce	0.38	0.49	1	-0.10	0.01	0.04	0.23
Zoned for service sector	0.13	0.34	1	0.02	0.31	0.03	0.24
High income street segment	0.07	0.25	1	0.00	0.91	0.01	0.49
Medium-income street segment	0.55	0.50	1	-0.05	0.11	0.01	0.83
Segments in quadrant	127.21	86.99	672	2.59	0.72	-3.95	0.47
Hot spots in quadrant	3.67	2.68	14	-0.05	0.39	-0.24	0.24

Spatial distribution of crime, potential spillovers and treatment restrictions
Lead to differential probabilities of assignment to treatment, spillover and control status

Crimes per segment, 2012-15



P(<250m from treated segment)



We estimate mean differences across experimental conditions, pooling two samples: experimental (N=1,919) and nonexperimental (N=77,848)

$$Y_{sqp} = \beta_1^P P_{sqp} + \beta_2^P M_{sqp} + \beta_3^P (P \times M)_{sqp}$$

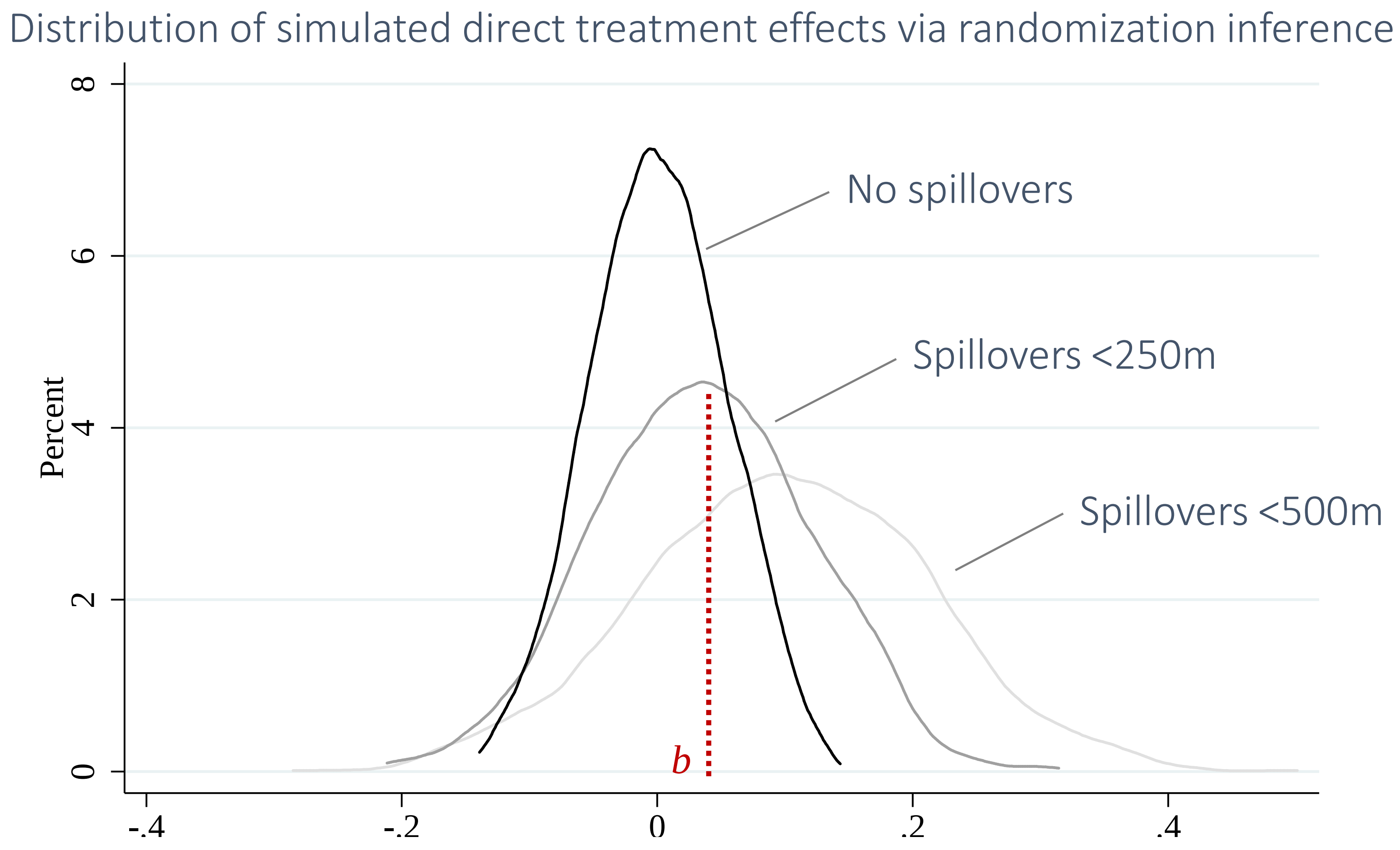
$$+\lambda_1^P S_{sqp}^P + \lambda_2^P S_{sqp}^M + \lambda_3^P (S^P \times S^M)_{sqp}$$

$$+\tau E_{sqp} + \gamma_p^P + \Theta^P X_{sqp} + \delta^P (E \times X)_{sqp} + \epsilon_{sqp}^P$$

- P, M and their interaction estimate direct treatment effects
- S can indicate spillovers <250m or also 250–500m (or be an empty set)
  - We pre-specified a test of p<0.1 to determine relevant spillover region
- E indicates the experimental sample and X a vector of controls
- Use inverse probability weights (IPWs) to account for the different probabilities of treatment assignment

But signs of identification problems:

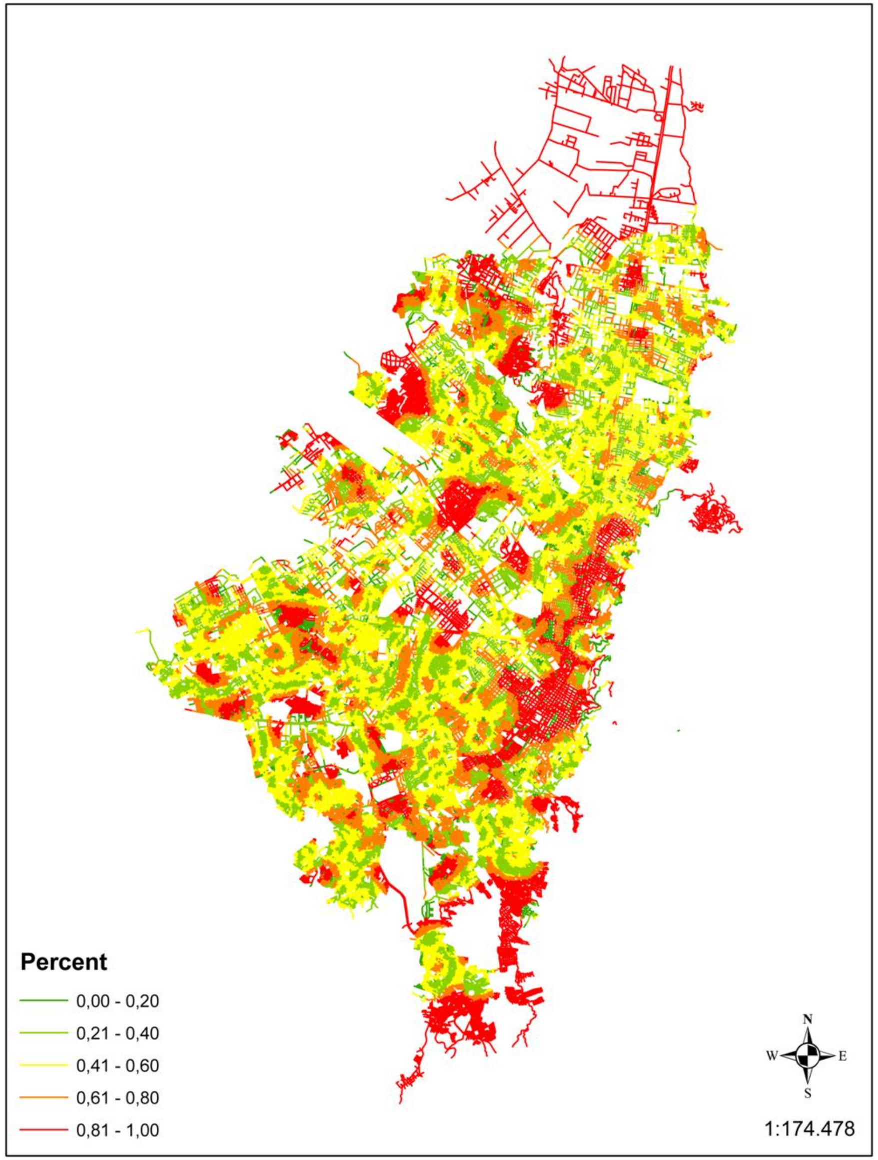
Randomization inference: 10,000 permutations of experiment reveal (1) a wider spread in the likelihood of rejecting the null of no treatment effect, and (2) an upward bias (b), as we account for larger spillover regions



“Fuzzy clustering” (Abadie et al. 2016)

- Spillover streets cluster together in most randomizations because of spatial distribution of crime
  - In most randomizations, streets that are close have a high chance of being in the same condition
  - No easy-to-model unit of analysis
- Widening of the sampling distributions (with spillovers) follows from:
  1. Losing data as we pare off spillover rings
  2. The control region shrinks and begins to exclude high-crime regions of the city
- Thus we use RI *p*-values in place of usual standard errors

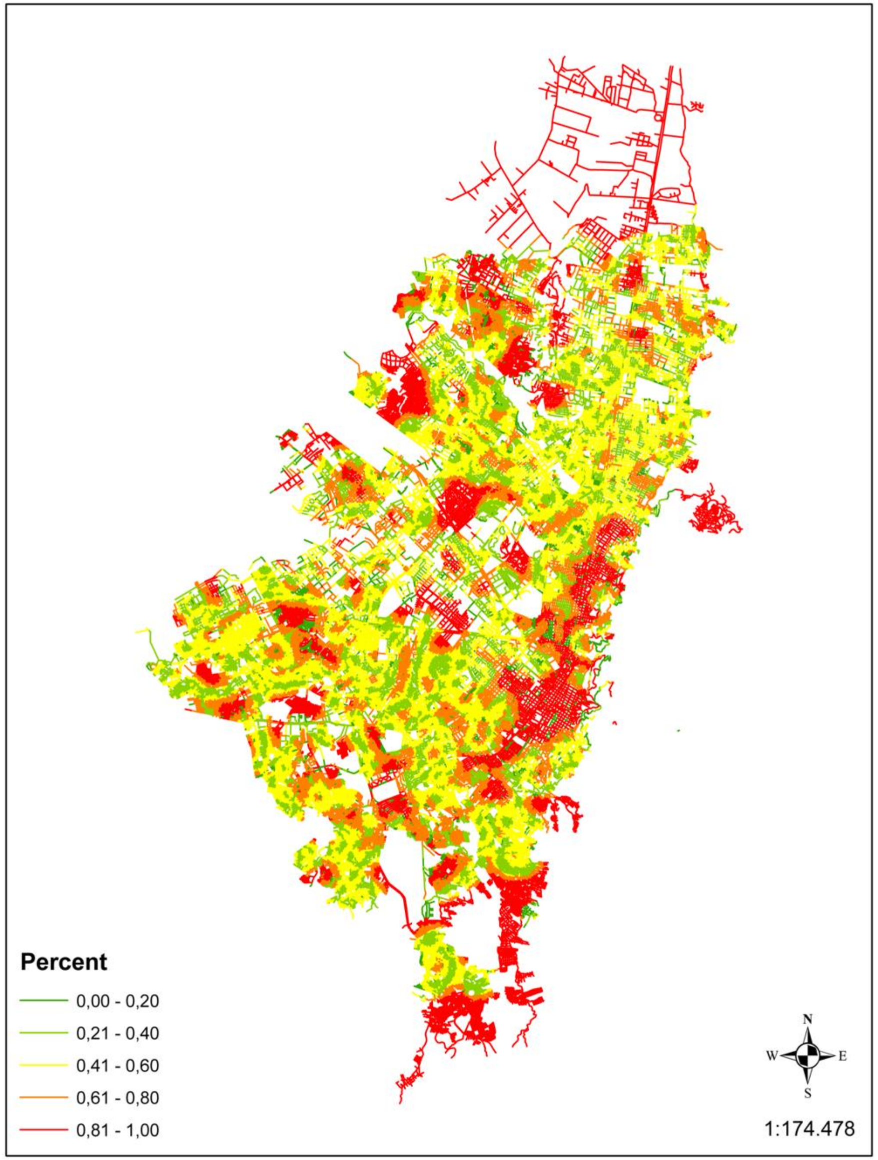
P(segments within 500m have the same treatment assignment as you)



But why is there bias, *b*?

- When we ignore spillovers, we stipulate that there is no such clustering, which is why that distribution is centered at zero
- Clustered assignment introduces bias when there are:
  1. Spillover and control clusters of unequal size, and
  2. When cluster size is correlated with potential outcomes
- Large clusters of control streets (those lying farther away from the downtown) have lower crime, leading to an upward bias
- Bias goes away as the number of clusters increases
  - Hence tiny when we account for non-experimental spillovers
- We subtract bias *b* from WLS estimates and test statistical significance using RI *p*-values

P(segments within 500m have the same treatment assignment as you)



We examine results during the 8 months of the intervention

1. Using reported crime data on all 136,984 city streets
   - Estimate direct and spillover coefficients
   - Estimate aggregate effect city-wide
   - Examine effects on property vs violent crimes
   - Examine impact heterogeneity: Moderate vs high-crime streets
2. Survey data on 1,919 experimental streets and 400 non-experimental streets
   - Check whether direct and spillover coefficients are different for perceived security and all crime (including crimes not officially reported)

Notes:

- No evidence of spillovers beyond 250m, so all spillover regions are 0-250m
- For simplicity, we pool the experimental and non-experimental samples and present spillover coefficients assuming that spillovers in the experimental and non-experimental segments are equal

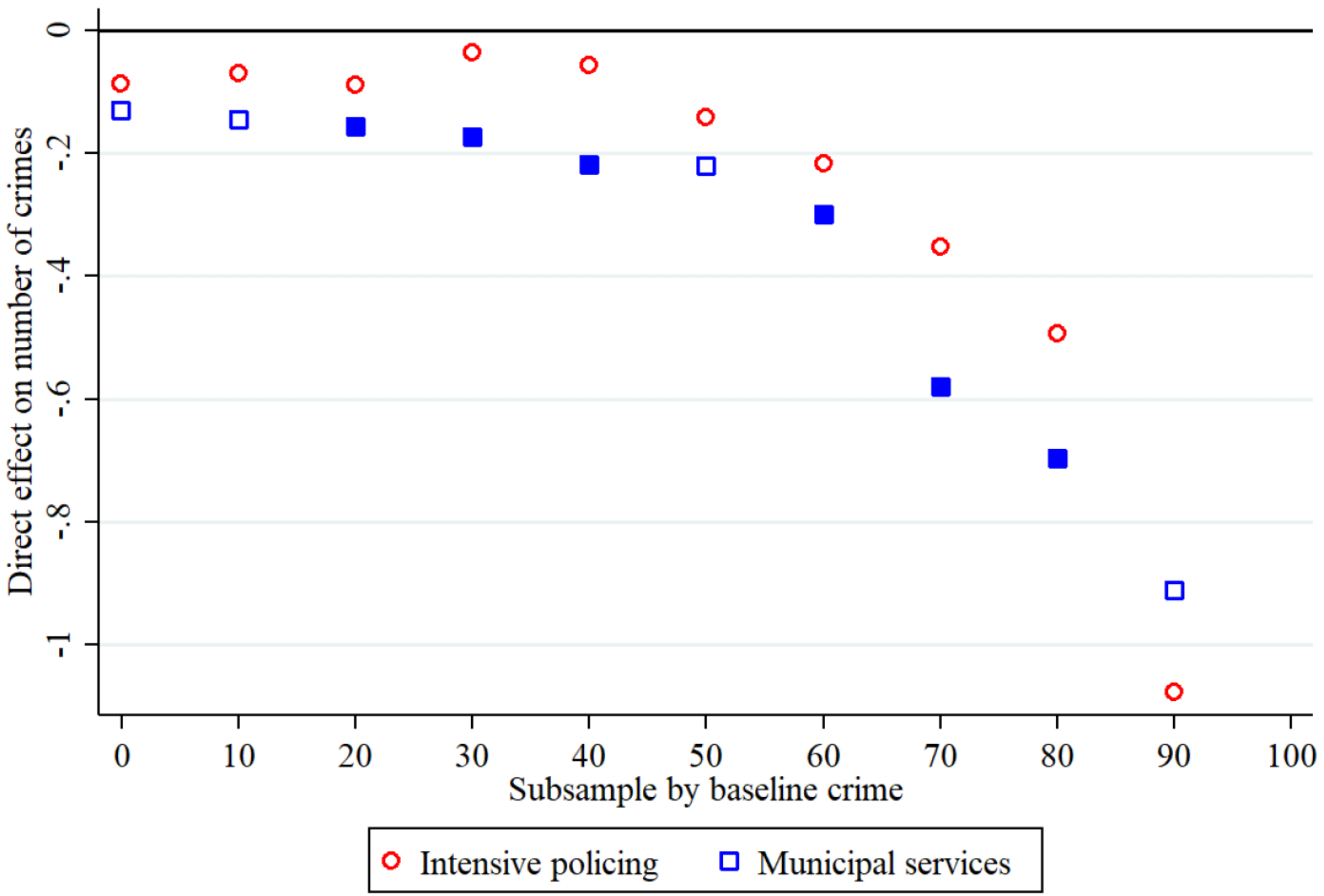
Aggregate impacts on reported crime

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

Aggregate impacts by property vs. violent crime

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

Heterogeneity in direct effects: Impacts in top X% by baseline crime



Filled in figures have p<.10

Results are generally consistent across specifications

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

Why might our conclusions differ from the US literature?
(other than the obvious fact that this is not the US)

- Of the 9 rigorous evaluations with >10 treated units, 6 of the 9 find no evidence of adverse spillovers (Braga et al. 2014, Weisburd & Telep 2016)
- Some reasons for caution
    1. A highly varied set of interventions
        - From drug house invasions to speed traps to problem-oriented policing to round-the-clock policing
        - We may not expect stable treatment effects
    2. Standard errors may be understated
        - In a recent meta-analysis, 9 of 14 component studies had p=0.000 on their spillover estimate despite a median study size of ~30 treated units
- Hence we urge more caution in the interpretation of existing results, and encourage more experiments at scale, with more transparent and replicable analysis

How do our results compare to the Medellín hot spots policing experiment?

(See Collazos et al. 2020)

- Broadly speaking, results on aggregate crime are similar
  - Small direct impacts and wide confidence intervals for aggregate effects (including the possibility of adverse spillovers)
  - Larger effects in the least secure places
- However, different types of crimes seem to respond differently
  - In Medellín, we saw large impacts on property crime with benefits diffusing to neighboring streets
  - Also, there was no effect on violent crimes
- Why?
  - This is not just a matter of internal vs external validity
  - Local crime patterns matter, and these differences have implications for our understanding of criminal incentives and behavior

The answer is important because it speaks to the economic organization of crime

- If most offenses in a city do not have a sustained motive, then place-based interventions might have a large deterrent effect with a minimum of spillovers
  - e.g. momentary crimes of passion
- If crimes with a sustained motive (e.g. professional theft) do not displace, this has important implications for our understanding of criminal markets
  - Criminal rents need to be concentrated, immobile, and unequally distributed within cities
  - Supply of crime is highly elastic to the probability of detection and apprehension in a small number of high-profit areas
  - Or some offenders are resistant to moving crime locations
- The evidence in Bogota is consistent with the first but not the second proposition, while the evidence in Medellín is consistent with the second but not the first proposition

Finally, these econometric issues and solutions will become more common with more urban experimentation

- Many urban programs are place-based and vulnerable to subtle spillovers
  e.g. improve traffic flows, beautify blighted streets and properties, foster community mobilization, rezone land use
- Economists have tended to impose a fair degree of structure on spillovers
  - In situations where the nature of spillovers is unknown, a more flexible approach might be more appropriate
- We show how spillovers threaten identification when the probability of exposure to spillovers varies
  - Follows Gerber and Green (2012), Aronow and Samii (2013), and Vazquez-Bare (2017).
- We also show how variance is underestimated when there is “fuzzy clustering” (Abadie et al 2016)
- Our proposed solution involves
  - Design-based approach to flexibly estimate spillovers and minimize fuzzy clustering
  - Use of randomization inference