Civilian alternatives to policing: Evidence from Medellín's community problem-solving intervention *Operación Convivencia**

Christopher Blattman Gustavo Duncan Benjamin Lessing
Santiago Tobón†

May 20, 2023

Abstract

Cities are searching for unarmed approaches to disorder and crime. Colombia is a leader here. Police are a national institution, and mayors have limited control over police staffing and operations. Thus, many cities have developed large civilian apparatuses for dispute resolution, family problems, and street disorder—an approach commonly called *convivencia* (coexistence). We worked with Medellín's municipal government to identify 80 typical neighborhoods and experimentally evaluate this approach. In half, the city intensified their civilian apparatus for 20 months, assigning full-time liaisons to help communities organize to improve public order, foster communication with the government, and connect residents to dispute resolution and family/social services. A centralized task force ensured city agencies fulfilled these roles. This represented a 60-fold increase in street-level presence, plus increased city agency attention. To our surprise, *Operación Convivencia* had no average treatment effect on reported governance, legitimacy, or crime. Anticipating that impacts could diverge by neighborhood, however, we prespecified heterogeneity analyses by initial state presence. This proved revealing. In neighborhoods where the state began weak, the task force and city agencies worked sporadically, and opinions of governance fell. In neighborhoods where the state began relatively strong, the liaisons and task force delivered, and there are indications this raised state legitimacy and reduced crime and emergency calls. The divergent results suggest the importance of existing state capacity, plus the dangers of over-promising and under-delivering. This may be especially important in Latin America, where cities like Medellín compete with gangs for local problem-solving and legitimacy.

JEL codes: H11, K42, O17, N46, C93

Keywords: Local government, public safety, governance, policing, crime, state building, urban policy, public services, security, gangs, field experiment, Colombia

*For comments we thank Thomas Abt, Oriana Bandiera, Eli Berman, Jennifer Doleac, Sarah Heller, Max Kapustin, Macartan Humphreys, Raul Sánchez de la Sierra, Jacob Shapiro, Carlos Schmidt-Padilla, Paolo Pinotti, Maria Micaela Sviatschi, Juan F. Vargas, and participants at several seminars and conferences. Innovations for Poverty Action coordinated all research activities. For research assistance we thank Verónica Abril, Bruno Aravena, David Cerero, Peter Deffebach, Felipe Fajardo, Sebastián Hernández, Sofía Jaramillo, Juan F. Martínez, Juan Pablo Mesa-Mejía, Angie Mondragón, Helena Montoya, José Miguel Pascual, Andrés Preciado, Arantxa Rodríguez-Uribe, Zachary Tausanovitch, Nelson Matta-Colorado, Martín Vanegas-Arias and México Vergara. We thank the Secretariat of Security of Medellín for their cooperation, especially the former Secretary of Security Andrés Tobón, as well as Lina Calle and Ana María Corpas. For financial support, we thank the Centro de Estudios sobre Seguridad y Drogas (CESED) of Universidad de los Andes; the Peace and Recovery Program (P&R) at Innovations for Poverty Action (IPA); the PROANTIOQUIA foundation; The National Science Foundation (NSF); the UK Foreign, Commonwealth & Development Office through the Crime and Violence Initiative at J-PAL; and the Economic Development and Institutions Programme (EDI) funded with UK aid from the UK Government, working in partnership with Oxford Policy Management Limited, University of Namur, Paris School of Economics and Aide á la Décision Économique.

†Blattman: University of Chicago & NBER, blattman@uchicago.edu; Duncan: Universidad EAFIT, gduncan@eafit.edu.co; Lessing: University of Chicago, blessing@uchicago.edu; Tobon: Universidad EAFIT, stobonz@eafit.edu.co

1 Introduction

In many cities, police are often the first (or only) official responders to a range of everyday disputes and disorder: domestic violence, civil disputes, homelessness, minor offenses, and other issues of neighborhood safety and regulation. Some police departments and unions would prefer to focus their attention on serious crimes.[1] Likewise, a growing number of city governments and local activists would prefer unarmed specialists, social workers, or communities themselves respond to these nonviolent social problems and street disorder.[2] Calls for civilian alternatives have grown in the wake of highly-publicized police killings of innocents or people committing minor infractions.[3] Tackling everyday neighborhood problems with city staff or community organizations could also be more cost-effective for cities.

Unfortunately, there is only modest evidence on police alternatives. A growing number of cities are piloting promising new ways to respond to emergency calls that involve mental health, addiction, and homelessness—sending social and mental health professionals independently of or in concert with the police.[4] Some communities have turned to grassroots neighborhood monitoring to make walking to school safer.[5] There is also a growing body of evidence on efforts by street outreach workers and social service agencies to identify the people at the most serious risk of gun violence and prevent or halt cycles of shootings.[6]

Colombian cities have been experimenting with civilian alternatives to policing for decades. One reason is that mayors have limited control over their metropolitan police forces. The police are a national institution, a branch of the Defense Ministry. Staffing levels, wages, and training decisions are made by the central government, not mayors. Large cities like Bogotá and Medellín have responded by constructing their own civilian security agencies

[1] In 2023, the Los Angeles police union called for more alternatives to policing for minor infractions and social problems in order to spend more time on violent crime (Zahniser, 2023). Consistent with this, substantial research suggests that more police, and more police attention to violent crimes, can indeed reduce serious crime (e.g., Chalfin and McCrary, 2017, 2018; Blattman et al., 2021). That said, the nature and quality of policing, not just the quantity, affects legitimacy and effectiveness (Owens, 2019; Owens and Ba, 2021).

[2] Evidence-based alternatives include social support for youth (e.g., Davis and Heller, 2020; Hjalmarsson et al., 2015), reducing financial stress in vulnerable communities (e.g., Blattman et al., 2017; Carr and Packham, 2019), or mitigating the harmful effects of the criminal justice system (e.g., Agan et al., 2023; Aizer and Doyle, 2015). This search for alternatives is also frequently present in public debates (e.g., Bokat-Lindell, 2023).

[3] See for instance Ang (2021) or DiPasquale and Glaeser (1998).

[4] There are a large number of meta-analyses of largely informal and some quasi-experimental evaluations. Some recent examples include Irwin and Pearl (2020); Seo et al. (2021); Dee and Pyne (2022).

[5] See for instance Gonzalez and Komisarow (2020).

[6] One approach—community-wide violence interruption—involves street outreach workers who intervene in and mediate active disputes (e.g., Butts et al., 2015). Another approach is to identify the men at the highest risk of violence and offer them cognitive behavioral therapy, often in combination with economic assistance (Heller et al., 2016; Blattman et al., 2017, ming; Bhatt et al., 2023).

with thousands of neighborhood-level staff responsible for dispute resolution, family disturbances, managing public spaces, and building capacity of community groups to manage local problems.

We worked with the municipal government of Medellín, the *Alcaldía*, to identify 80 low- and middle-income neighborhoods and experimentally increase the intensity of their civilian public safety apparatus in a random half of the sample. Medellín is Colombia's second-largest city, with a population of roughly 2.5 million. It is one of the nation's industrial and commercial centers, with an annual income of roughly $11,500 per capita in purchasing parity terms. Like most large Colombian cities, Medellín has a well-organized bureaucracy with high tax revenues and public services. Its main civilian public service agency, the Secretariat of Security, has roughly 1 employee per 1,000 residents. By comparison, there are roughly 2.7 police officers per 1,000 residents (similar to U.S. cities like Los Angeles).

The intervention had two main components. First, the Alcaldía created a special task force to ensure that needs in the 40 treatment communities would get priority attention in the city's many service agencies. This included the city's main dispute resolution officers for problems between and within families, as well as the city agencies that respond to requests for basic services (such as street lights and garbage pickup).

Second, each neighborhood received a full-time "liaison"—a street-level bureaucrat whose job was to: rejuvenate community government organizations; advertise and link people to government agencies; identify public service needs for the task force to fix; and, for serious family, business, and inter-household disputes, connect disputants to professional mediators or family-services officials from the city government.

In contrast to programs that pilot new forms of policing and crisis-response, this intervention sought to evaluate the returns to increasing a broad range of "normal" municipal services for basic order and community organization. In many U.S. cities, similar services are performed by a mix of municipal agencies, government-funded social work organizations, and the constituent services offices for city councilors and aldermen. Nonetheless, police often find themselves the first responders. Medellín's intervention is akin to an increase in a city's civilian capacity for both prevention and first response. Recent examples of similar coordinated intensification programs in the United States include the Building Neighborhood Capacity Program (BNCP) and the Neighborhood Revitalization Program (NRI).[7]

Unfortunately we don't know much about the returns to these approaches. There is virtually no evidence on what happens when cities expand their normal civilian prevention

[7]For information on the BNCP, see https://www.ojp.gov/ncjrs/virtual-library/abstracts/building-neighborhood-capacity-program-bncp-fact-sheet. For NRI, see https://obamawhitehouse.archives.gov/administration/eop/oua/initiatives/neighborhood-revitalization.

and response capabilities—either broadly or in targeted neighborhoods. The costs of hiring bureaucrats, extending public services, and expanding fiscal capacity are often immediate and clear. But what should governments expect to receive in terms of efficacy, legitimacy, and public support for expanding street-level city workers and services?

To assess this, we and the city kept the 80 eligible neighborhoods intentionally small: up to about 10 city blocks with about 1,000–3,000 residents each. We selected neighborhoods that were typical of the city's lower-income neighborhoods. Some already received moderate services from the Alcaldía, while others had seen relatively little state presence on the street.

The intervention began in early 2018 and lasted 20 months—until the end of the mayor's term. Because of the neighborhood size and the level of staffing, we estimate these neighborhoods were assigned to receive a 60-fold increase in street-level attention to problems, and perhaps a doubling or tripling of attention from the central municipal government. The intervention did not change police patrolling in the neighborhoods.

We designed the experiment to be this focused and intensive not because this style of intervention necessarily needs to be so localized and targeted. Rather, we did so primarily to maximize statistical power as well as minimize interference between experimental units. While there are a large number of experiments testing community-level policing strategies, community-level experiments of non-policing strategies are exceedingly rare. Most evaluations are observational, where causal identification is difficult due to small sample sizes and the difficulty of finding counterfactual comparison communities (Farrell et al., 2016; Roman et al., 2018). Thus, while an experiment in 80 small neighborhoods has obvious limitations in terms of statistical power and generalizability, Medellín's willingness to experiment on this scale is still unusual in terms of its size and rigor.

We also designed and launched the intervention two years before police violence in the United States galvanized policy debates around allocating more funds to non-police forms of public safety and problem-solving. Colombia also faced massive protests against police violence in 2020 and 2021, which triggered a similar debate and a subsequent police reform (e.g., Abril et al., 2023). Even before these developments, however, the Alcaldía was focused on improving public-service delivery, reducing crime, and generally improving the image and legitimacy of the state (not to mention the reputations of the mayor and political appointees in charge of civilian security).

State effectiveness and legitimacy matter in every society, but they take on special importance in cities where illegal groups compete for citizen loyalty, in part by providing criminal governance. Like many Latin American and U.S. cities, Medellín is home to entrenched street gangs. Most low- and middle-income neighborhoods have a well-organized local gang called a *combo*, which engages in a variety of illicit businesses, especially local drug sales.

In addition, many of Medellín's combos provide some degree of basic order, local security, dispute settlement, and debt collection. Some do so on a fee-for-service basis, and many also charge weekly fees or taxes known as *vacunas* (vaccines) for these services. The state remains the dominant provider of public goods, including security, but in practice there is a duopoly of governance and coercion. Such "criminal governance" is particularly common in Latin America, but gangs also rule civilians in Italy, the United Kingdom, India, South Africa, and the American prison system (Arias, 2006; Lessing et al., 2019; Lessing, 2020; Melnikov et al., 2020).[8] With this intervention, the Alcaldía aspired not only to reduce disorder and crime, but (eventually) to crowd out gang rule.

After conducting hundreds of qualitative interviews with community members, leaders, police, and gang leaders, we identified 17 forms of everyday order and governance that could be provided by the state, the gangs, or both. We also identified 5 dimensions of trust in and satisfaction with ruling authorities. Based on these findings, we prespecified two primary outcomes for the experiment: an index of relative state governance and an index of relative state legitimacy, which we calculate as the difference between the absolute state and combo measures. 20 months after the intervention began, we interviewed roughly 2,400 people in the 80 experimental sectors, collecting data on governance levels and legitimacy for both the state and the combo. Of course, we are also interested in absolute levels of state governance and legitimacy, as well as impacts on crime and emergency calls to police, and we examine these as secondary analyses.

To our surprise, we see no evidence that intensifying civilian personnel and attention improved any of our outcomes on average. If anything, residents in treated neighborhoods reported a roughly 6 percent *lower* opinion of relative state service provision, statistically significant at the 10 percent level.

Furthermore, despite high levels of street presence by the liaisons, we see only weak evidence that the average resident noticed the increased municipal attention, and no evidence that they participated in the events or were assisted by the liaisons. Thus, on average, the intensification of everyday forms of governance yielded no apparent returns, and may even have backfired.

This null finding, if true, would be a surprising and important policy result. It would imply that tripling city agency attention and increasing unarmed street-level presence 60-

[8]States, even strong ones, often face internal competitors. Traditional leaders, influential persons, and community organizations also regulate everyday life. These groups don't necessarily undermine the state, and are often complementary (Cammett and MacLean, 2014; Van der Windt et al., 2019; Blattman et al., 2014; Henn, 2021). When it comes to public security and justice, however, state legitimacy can suffer when other coercive actors—criminals, paramilitaries, or insurgents—govern the population (Berman and Laitin, 2008; Acemoglu et al., 2020; Cammett and MacLean, 2014). This is one reason why states aim to monopolize the legitimate use of force (Weber, 1946).

fold had no effect on order, security, or legitimacy—even in a city with a reputation for one of the strongest and well-funded municipal governments in Latin America.

This is not, however, the conclusion we think we should draw. Rather, a closer look reveals important heterogeneity. Having anticipated that program effects might depend on initial conditions, we prespecified that we would look for differential results by baseline levels of relative state governance. We divide the 80 sectors into those with above- and below-median levels of initial state presence and examine how program impacts varied in these two types of neighborhoods.

The results suggest that program implementation and impacts diverged by initial state presence. First, we find stark differences in people's awareness of the municipal staff and their participation in events. Where the state began initially strong, residents report significantly more interactions with city staff and events. Where the state began initially weak, they report fewer. The difference between the two kinds of sector is large—about 25 percent of the average level of engagement with municipal staff and events in the city.

Post-program interviews with community leaders and the liaisons provide some clues why. Where the state began with capacity and presence, the task force and liaisons were able to execute the intervention effectively. Where initial state presence was weaker, however, the central task force or other officials were sometimes unable to deliver on important promises. This may have raised expectations of what the task force and municipal apparatus could actually deliver, and so failures to follow through were doubly disappointing.

We see similarly divergent impacts on governance and legitimacy. In initially well-governed sectors, the program increased state legitimacy by almost 10 percent (significant at the 5 percent level), and decreased crime by 28 percent. Emergency calls to the police also fall by 41 percent, as municipal workers or (more likely) communities themselves either prevent or deal with the most common forms of neighborhood disorder—street fights, domestic abuse, and drug users creating a public nuisance. Not all indicators improve, however, as the program had no effect on residents' reports of the responsiveness of the police and Alcaldía to more serious incidents. Nonetheless, the improvements in state legitimacy, crime, and police calls are massive. These results hold whether we look at absolute or relative levels of state governance and legitimacy.

Meanwhile, in initially poorly-governed sectors, there is no evidence the program had positive effects. Indeed, residents' reports of state responsiveness fell about 8 percent (significant at the 10 percent level), and there was a weak decrease in state legitimacy and crime prevention, and a weak increase in calls to the police over street disorder (none of which are statistically significant). Because the program impacts move in opposite directions, the differences in performance between well- and poorly-governed sectors are generally statistically

significant.

Finally, we see no evidence of a combo response in either the quantitative or qualitative data. Combos commonly watched liaisons closely at first. In a small handful of cases, combos even impeded liaison activities for the first few weeks, until they realized the bureaucrats were benign. Thus, for most of the intervention, there was no apparent combo reaction at all, and average treatment effects on combo governance and legitimacy bear this out. This suggests that the impacts we see are largely driven by state capacity in initially well-governed sectors.

Altogether, these findings speak to a central question facing many governments: what are the returns to investments in the apparatus of neighborhood governance, and how do these returns depend on initial capacity and legitimacy? Governments ask this about police and non-police interventions alike. Theoretically, the answer is ambiguous. On the one hand, in areas with little history of state services, we might expect the first investments to have out-sized impacts. (This was our initial hypothesis in Medellín, where residents of the least-served areas initially expressed relief at finally seeing municipal bureaucrats in their neighborhoods.)[9] On the other hand, establishing robust state governance and legitimacy might require large and sustained investments, especially from a low starting point.

Our results are consistent with the latter hypothesis. Granted, we should be careful not to generalize from a single experiment in 80 neighborhoods, especially when the results rely on subgroup analysis (even if prespecified). Nonetheless, the experience of Medellín implies that the returns to investments in state-building on the margins are complex. Large impacts on legitimacy and crime in relatively are clearly possible. But the returns may be low or even negative where state capacity is weakest. This could result in perverse political incentives when it comes to public services and security: bureaucrats might seek the path of least resistance, and politicians might try to invest their marginal effort and resources to the areas where visible progress is easiest to achieve.

Finally, Medellín's willingness to randomize a community-level program is a rare opportunity, one that we hope will be repeated. To the best of our knowledge, it is the only randomized evaluation of a community-level non-police intervention, at least at this scale. This exercise demonstrates that experimentation at this level is possible, and that other approaches could to be evaluated in a similar manner.

In doing so, however, our results also suggests that experimental samples should not necessarily be limited to the most disordered communities and other "hot spots". It is true

[9]This hypothesis finds support in recent literature suggesting that the returns to government investments in fostering political participation might be highest in places where the state is weakest, following tax collection efforts (Weigel, 2020).

that crime and disorder are often concentrated in particular places, and that cost-effective security interventions are often highly targeted (e.g. Weisburd et al., 2012; Abt, 2019). But program effectiveness may vary by neighborhood type, and so diversity in the sample may be important to understanding what interventions work and why.

2 Context

2.1 The state and security in Medellín

The Metropolitan Police

Medellín's police force has roughly 2.7 officers per 1,000 people. This level is slightly higher than the U.S. national average, and comparable to U.S. cities like Los Angeles. It is lower, however, than other large U.S. cities, such as New York and Chicago, which have ratios above 4.[10] While low-level corruption and poor responsiveness are common, the police are fairly professionalized, particularly in comparison with other Latin American countries.

The city is divided into 16 comunas and, except for a couple of exceptions, each comuna is a separate police jurisdiction with its own commander and station. Each police jurisdiction is divided into a large number of *cuadrantes* (quadrants)—a sub-unit relevant only for the police, similar to police beats. Each quadrant has 6 assigned officers who patrol on motorbike, in pairs, in 3 shifts per day.

Crucially, although the Colombian Constitution designates mayors as local police authorities, the Metropolitan Police are independent from the city government. They are part of the National Police, which is currently a branch of the Defense Ministry. City governments in Colombia can use local tax revenues to make major investments in policing and security infrastructure (such as buildings) and fund their police force's equipment and vehicles. But wages are paid by the national government, and the number of officers is set by the National Police. Hence, while mayors have some degree of authority over their Metropolitan Police—shaping the allocation and focus of policing or providing more and higher quality equipment and stations—they cannot easily increase the quantity or quality of police personnel.

Civilian public safety organizations

Police autonomy is one reason why large cities in Colombia have built large civilian organizations to provide public safety, dispute resolution, and other services that reduce disorder and tackle social problems, commonly referred to as *"convivencia,"* meaning coexistence.

[10]For U.S. cities and the national average, we use Tables 24 and 26 from the FBI's 2016 Uniform Crime Reporting system, https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016.

Like most cities, Medellín has a Secretariat of Security—an organization of roughly 2,500 civilian staff who provide numerous services to residents, including responding to various emergencies and street disorder, directly resolving community disputes and domestic violence, and regulating the use of public space. Its structure is subdivided into four areas: (i) the Under-Secretariat of Security Planning, which is in charge of deciding budgeting and investments, as well as coordinating security policy with other agencies such as the police; (ii) the Under-Secretariat of Security Operations, which is in charge of directly coordinating actions with the police and other security agencies; (iii) the Under-Secretariat of Local Government and Coexistence, which handles dispute resolution, domestic violence, and family-related services such as mediation; and (iv) the Under-Secretariat of Public Space, which oversees and regulates public space and some forms of disorder across the city.

The mayor or *Alcalde* oversees this Secretariat and appoints all leadership positions (the secretary and under-secretaries). These leaders, along with other permanent and non-politically appointed senior staff, constitute the top-down task force that was part of the intervention we evaluate in this paper (see Section 3).

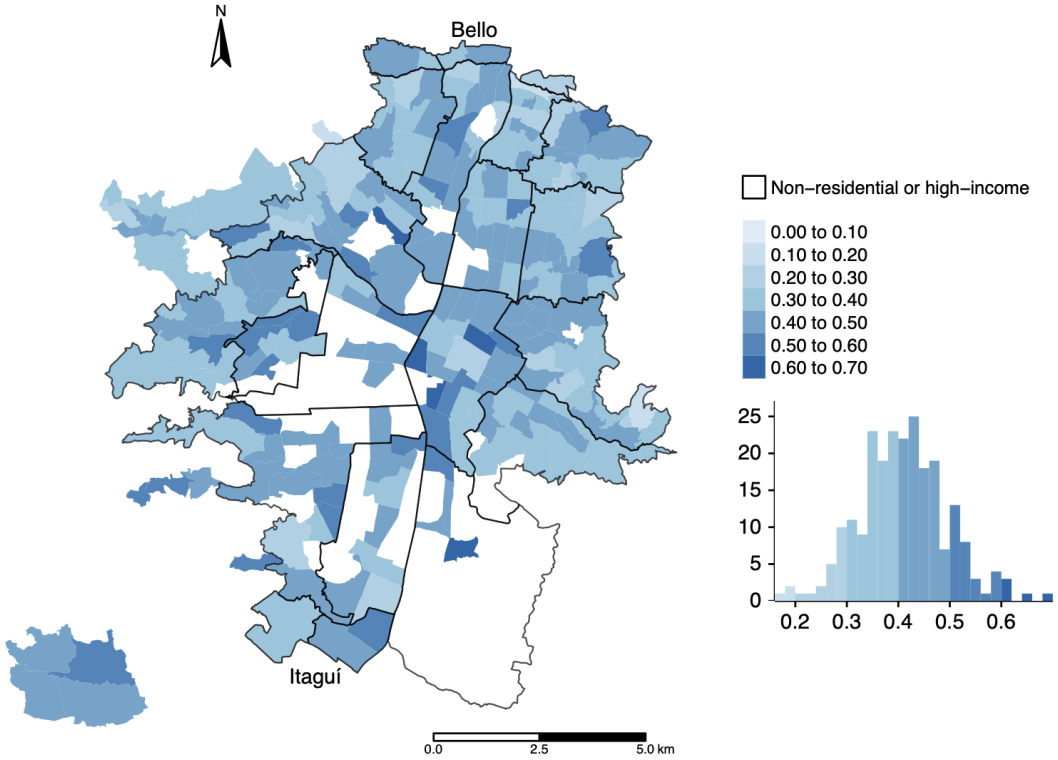
These security- and dispute-related units have several "headquarters" in each comuna, including *inspecciones* who directly resolve community disputes through a formal, fast-track justice service, and *comisarías* who provide a wide range of family services aimed at resolving legal problems, mental health problems, domestic violence, child protection, and family law. Each comuna also has a "liaison" that performs community outreach, in order to identify which neighborhoods or people are in need of these services.

In addition to these comuna-based services, comunas are divided in barrios, and roughly each one has an elected community action board (*Juntas de Acción Comunal*, or JACs) that help local groups regulate and organize their community. They are rarely involved in security, protection, and dispute resolution, however, and so they are not a major focus of the intervention or activities we study in this paper.

Assessing state governance and legitimacy

To understand the quality of these security and dispute resolution services, we and a team of research assistants began by conducting a large number of semi-structured qualitative interviews. We conducted formal interviews with 23 community leaders and 141 community residents and shopkeepers. We also interviewed 19 police officers and officials, 17 city officials, 10 prosecutors, and 18 other crime and security experts. We used these interviews to understand the nature of crime, disorder, and delinquency in different neighborhoods. We also identified the most common forms of governance and taxation by various actors in Medellín—police, the Alcaldía, community organizations, local councils, and criminal actors.

Figure 1: State governance levels by barrio



Notes: The figure displays average levels of state governance reported in each low- and middle-income barrio, using the average of all 17 items from Table 1, averaging across all survey respondents in the barrio. We did not survey high-income residential neighborhoods or non-residential areas, all of which which appear in white.

Table 1: State and combo governance and legitimacy, barrio survey averages, 2019

	Frequency/Rate (0-1 Scale)				Relative State – Combo	
	State		Combo		City-wide survey	Experimental control group
	Estimate	SD	Estimate	SD		
	(1)	(2)	(3)	(4)	(5)	(6)
Governance Index	0.41	0.26	0.34	0.29	0.07	0.07
How often they intervene when:						
HH: Someone is making noise	0.43	0.38	0.19	0.30	0.23	0.26
HH: Home improvements affect neighbors	0.41	0.38	0.25	0.34	0.16	0.14
HH: There is domestic violence	0.51	0.37	0.35	0.37	0.15	0.15
HH: Two drunks fight on the street	0.54	0.36	0.40	0.37	0.13	0.13
Biz: Someone disturbs a business	0.50	0.38	0.36	0.38	0.12	0.16
Biz: You have to react to a robbery	0.52	0.37	0.40	0.39	0.11	0.12
Biz: It is necessary to prevent a theft	0.45	0.37	0.38	0.39	0.07	0.08
Biz: Businesses in this sector are robbed	0.42	0.39	0.35	0.38	0.05	0.07
HH: People smoking marijuana near children	0.29	0.36	0.25	0.36	0.04	0.03
HH: A car or motorbike is stolen	0.46	0.37	0.43	0.38	0.04	-0.01
HH: Someone is threatening someone else	0.42	0.36	0.41	0.37	0.01	-0.01
HH: You have to react to a robbery	0.46	0.36	0.45	0.38	0.01	-0.02
HH: Someone is mugged on the street	0.39	0.36	0.41	0.38	-0.01	-0.05
HH: It is necessary to prevent a theft	0.40	0.36	0.42	0.38	-0.03	-0.04
HH: Kids fight on the street	0.29	0.35	0.32	0.37	-0.04	-0.03
Biz: Someone does not want to pay a debt	0.17	0.31	0.23	0.35	-0.06	-0.05
HH: Someone refuses to pay a big debt	0.22	0.31	0.39	0.38	-0.16	-0.20
Legitimacy Index	0.58	0.21	0.43	0.28	0.13	0.13
When solving problems in the neighborhood:						
How much do you trust the...	0.57	0.30	0.36	0.36	0.19	0.20
How fair is the...	0.55	0.27	0.41	0.35	0.11	0.12
How do you rate the...	0.60	0.22	0.51	0.28	0.09	0.09
How would your neighbors rate the ...	0.59	0.23	0.50	0.29	0.09	0.08
How much do your neighbors trust the...	0.57	0.28	0.47	0.36	0.09	0.06

Notes: The governance and legitimacy indexes are averages of the component questions listed in this table. Columns 1–5 present averages from the city-wide survey, representative of Medellín 's 224 low- and middle-income barrios, with 20–25 respondents per barrio. Column 6 reports averages for the experimental sample of 80 sectors, with roughly 30 respondents per sector. The Relative State measures in Columns 5 and 6 are the differences between columns 1 and 3. All governance scales correspond to: 0 = Never, 0.33 = Occasionally, 0.66 = Frequently, 1 = Always. All legitimacy scales correspond to: 0 = Nothing, 0.33 = A little, 0.66 = Somewhat, 1 = Very. Both households (HH) and businesses (Biz) were surveyed on governance levels, but only households were surveyed on legitimacy (hence there are fewer observations).

Based on these interviews, we designed a survey to measure state legitimacy and its responsiveness to insecurity and disputes. The city is divided into roughly 250 *barrios* (neighborhoods). In 2019 we stratified the city by barrio and randomly sampled roughly 16 percent of the city's 14,600 blocks. We only sampled from the 223 low- and middle-income barrios (excluding high-income neighborhoods, non-residential areas, and the city center). We then sought to interview two households and one business on each sampled block, for a average of about 21 survey respondents per barrio. We call this representative group the *city sample*. This representative city sample of blocks is distinct from the *experimental sample*—80 small neighborhoods about 5–10 blocks each, discussed below.[11] The analysis in this section focuses on the city sample alone.

Our qualitative work identified 17 of the most common and important disputes and forms of disorder. The survey asked residents to report how frequently two state actors—the police and the Alcaldía—responded to each of the 17, where 0 = Never, 0.33 = Occasionally, 0.66 = Frequent, 1 = Always. Using these these responses, we create an average index of *State governance* (0 to 1). Figure 1 reports average levels of state security and dispute resolution responsiveness as reported by residents. Table 1 reports levels of governance and legitimacy for each of the index components plus the overall averages.

Overall, residents score state responsiveness as 0.41 on this 0 to 1 scale—slightly better than "occasionally" responsive to disorder and disputes. State responsiveness is greatest (above 0.5) for robberies, domestic abuse, and adult street fights. It is poorest (below 0.3) for debt collection, teenage disputes, and drugs and smoking near children.

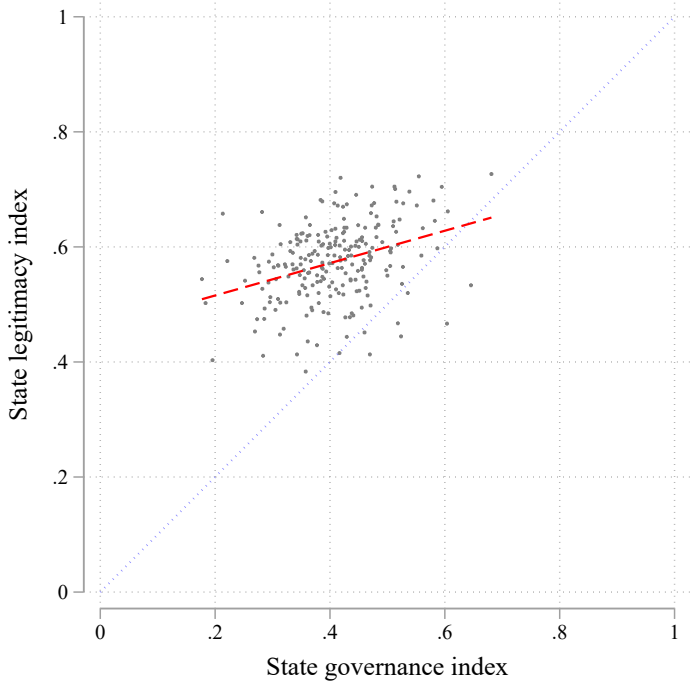
The survey also asked residents about state legitimacy: how much they trust the police and Alcaldia; whether they are fair; whether residents are satisfied with them; and whether residents thought their neighbors trust the state. Table 1 reports these responses as well, including an average index of *State legitimacy* (0 to 1). On average, residents give the state a rating of 0.58.

As Figure 1 illustrates, however, state services are unevenly distributed. Some barrios report levels below 0.2, and others above 0.6. The most significant correlate of state responsiveness is elevation. The city lies in the valley of a river running roughly south-north. Barrios along the central south-north axis in Figure 1 are in the lower slopes and valley, while barrios to the left and right are generally built on steep slopes.

Unsurprisingly, trust in and satisfaction with the state are associated with the quality of security. Figure 2 plots state legitimacy against state governance levels by barrio. We see a

[11] 40 of the neighborhoods received the intervention in the 30 months prior to the 2019 survey. They represent less than 2.5 percent of all city blocks. Could the treatment influence outcomes in the city sample? We test for evidence of such spillovers in Appendix Table A.1 and find no evidence of them.

Figure 2: Relationship between state governance and legitimacy, 2019



Notes: Each dot is a barrio average, and the dashed line indicates fitted values (a correlation of 0.35). We did not survey high-income barrios.

positive correlation of 0.35.

2.2 Street gangs

Virtually every low- and middle-income neighborhood in Medellín also has a local gang called a combo. There are roughly 400 in the metropolitan area. Combo territories—often no more than a few dozen blocks—tend to be long-standing, well-defined, known to locals, and relatively stable over time.

Our qualitative interviews also focused on understanding these gangs and their role in their communities. As part of a larger and ongoing project, we have interviewed 118 leaders and members in 41 gangs and related criminal groups. This is a convenience sample of criminal actors who agreed to speak with us. Half of the interviews took place in one of Medellín's three major prisons, where a large number of high- and middle-ranking leaders continue to direct street operations. We believe they spoke to us for several reasons: pride; respite from boredom; interest in speaking with professors; the fact that they were already prosecuted; the fact that we were not asking about indictable information; and a hope that

this would further their efforts for a peace process with the government.[12]

As in many U.S. and Latin American cities, most of these gangs are small, well-organized illicit firms whose main profits come from local retail drug sales. Medellín's combos typically have 15 to 50 permanent members between the ages of 15 and 35, and each member typically has well-defined positions in one of the combo's illicit business lines. Besides holding a monopoly on drug sales in their neighborhoods, combo members frequently participate in and regulate local informal and sometimes legal markets, including microfinance (loan sharking) as well as consumer goods—especially cooking gas, arepas, milk, and eggs. In Blattman et al. (2023b) we analyze the internal organization and market structure of the combos in more detail, as well as how they manage competition and try to minimize violent conflict.

Criminal governance Combos are also relevant for this paper because many of them offer protection and dispute resolution services in their neighborhoods, often in exchange for security fees. The intervention we study was primarily aimed at improving state governance rather than combating gang rule per se. Nonetheless, this backdrop of criminal governance is important to understand, if only because the state would presumably prefer to have a monopoly on coercive protection and taxation.

Criminal groups have been found to govern civilians in parts of Italy, the United Kingdom, India, South Africa, and the American prison system (Arias, 2006; Lessing, 2020; Melnikov et al., 2020).[13] In Latin America, criminal governance is pervasive. Uribe et al. (2022) estimate the number of people living under some form of gang rule in the tens of millions. This leads to a duopoly of coercion where state and criminal groups' authority overlaps, with both contributing to everyday governance.

Medellín is a well-known case of this widespread phenomenon.[14] After the state, combos are the most common organization that residents turn to in order to settle household and business disputes, collect debts, stop fights, prevent thefts, manage the homeless and drug addicts, and other aspects of neighborhood order.

In return for these services, combos typically collect weekly fees from local businesses and residents, and may also charge on a fee-for-service basis. Residents and businesses typically call the weekly tax a *pago por la vigilancia* ("security" or "surveillance fee") or,

[12]We had several strategies for maintaining confidentiality of criminal group members. Above all, we were transparent about our research aims and work with the government. We made every effort to preserve anonymity and confidentiality, while advising subjects in consent scripts of the potential limits to our ability to do so. Finally, we consulted extensively with the human subjects committees of our institutions, and we obtained written support and assurances of noninterference from several authorities. We discuss these data, ethical considerations, and other information in more detail in Blattman et al. (2023a).

[13]Indeed, historically, warlords and other strongmen have commonly emerged to provide security and adjudication in return for taxes and rents (Tilly, 1985; North et al., 2009; Sánchez De La Sierra, 2020).

[14]e.g. Arias (2017); Cruz and Durán-Martínez (2016); Moncada et al. (2018).

more colloquially, a *vacuna*—literally, a vaccine.[15]

How pervasive are these combo services, and how do residents feel about them? Our 2019 representative survey measured residents' opinions of combo governance and legitimacy (in addition to that of the state). Table 1 reports levels of both. Average reported combo governance is 0.34—about 83% the level of the state. Among the 17 components of this index, the combo is rated as slightly more responsive than the state in eight situations: responses to muggings, vehicle thefts, threats, and robberies; preventing theft; teenage street fights; and business and household debt collection. Average combo legitimacy is 0.43—about 75% of the state"s level.

We construct a measure of *Relative state governance* as the difference between these two measures, where positive values imply more state governance than combo governance.[16] Figure 3 maps this by barrio. The state is present in every neighborhood, but varies in its responsiveness and penetration, as we saw above. Likewise, a combo is almost always present, but combos vary widely in the extent to which they offer governance and security services. Many choose to provide no governance at all. Others provide a wide range of services. As a result, while the state is the dominant provider of protection in most neighborhoods, we observe wide variation. In 31% of neighborhoods, residents report the combo is more responsive to these 17 forms of disorder than the state.[17]

In Blattman et al. (2023a) we study the historical evolution of criminal governance in Medellín. Combos were present in most neighborhoods as early as the 1970s and 80s, but gang rule and taxation is a more recent phenomenon, emerging in the 1990s and early 2000s. This companion paper shows that combos had two main motives for governing. One was a drive for revenues, as gangs could charge taxes and fees in return for settling disputes and providing security. But the combos also ruled to protect their other business lines from state predation, especially drugs. The rationale is simple: if gangs maintain order on the streets,

[15]In most cases, this is not a purely extortionate protection racket in the sense of demanding money from shops or households in exchange for agreeing not to harm them. At the same time, payment and participation is seldom voluntary. If the local combo decides to provide security services on a block, most shopowners will be compelled to pay the vacuna.

[16]Of course, "governance" also includes material public goods such as infrastructure, as well as collective decision-making and coordination. Our interviews and surveys found that combos rarely offer such services. Infrastructure is provided almost solely by the state, while informal leaders and elected neighborhood councils manage most local collective decisions. Instead, combos tend to specialize in services that are at least partially excludable, and those that benefit from coercive power. In the remainder of the paper, we use "governance" as shorthand for this set of protection services in which both gang and state participate.

[17]The state and the combo offer different forms of governance, of course. The state's dispute resolution and court systems tend to be impartial and professional, and city leaders are elected in competitive elections. The combo is unelected and relatively unaccountable, and may provide more "justice" to those who hire them or who are closest to them. At the same time, combos may have more local knowledge and deeper networks than most state bureaucrats. Combos are also available all the time, and act swiftly.

the police may be less likely to be called into the neighborhood, and residents may be less likely to share information or collaborate with the state. In that paper, we use exogenous variation in policing and municipal security services to show how, over 30 years, combo governance tended to emerge where the state was strongest and drug profits were greatest.

For these reasons, this paper will examine program impacts on combo behavior as well as that of the state (hence the pre-specification of relative state governance and legitimacy as our primary outcomes). Nonetheless, we expect that the largest, most direct, short term impacts will be on state governance and legitimacy. Whether combos respond to increased non-police state presence in the short term is a secondary analysis.

3 Intervention

At the outset of the study, the city government was interested in increasing public safety and "coexistence," and thereby improve its legitimacy and popularity. Given that the Metropolitan Police are a national institution, however, the Alcaldía's principal option was to augment a broad array of municipal services related to neighborhood order and problem managament, mainly via the staff and activities of the Secretariat of Security. We worked with the Secretary to identify, implement and evaluate a targeted version.

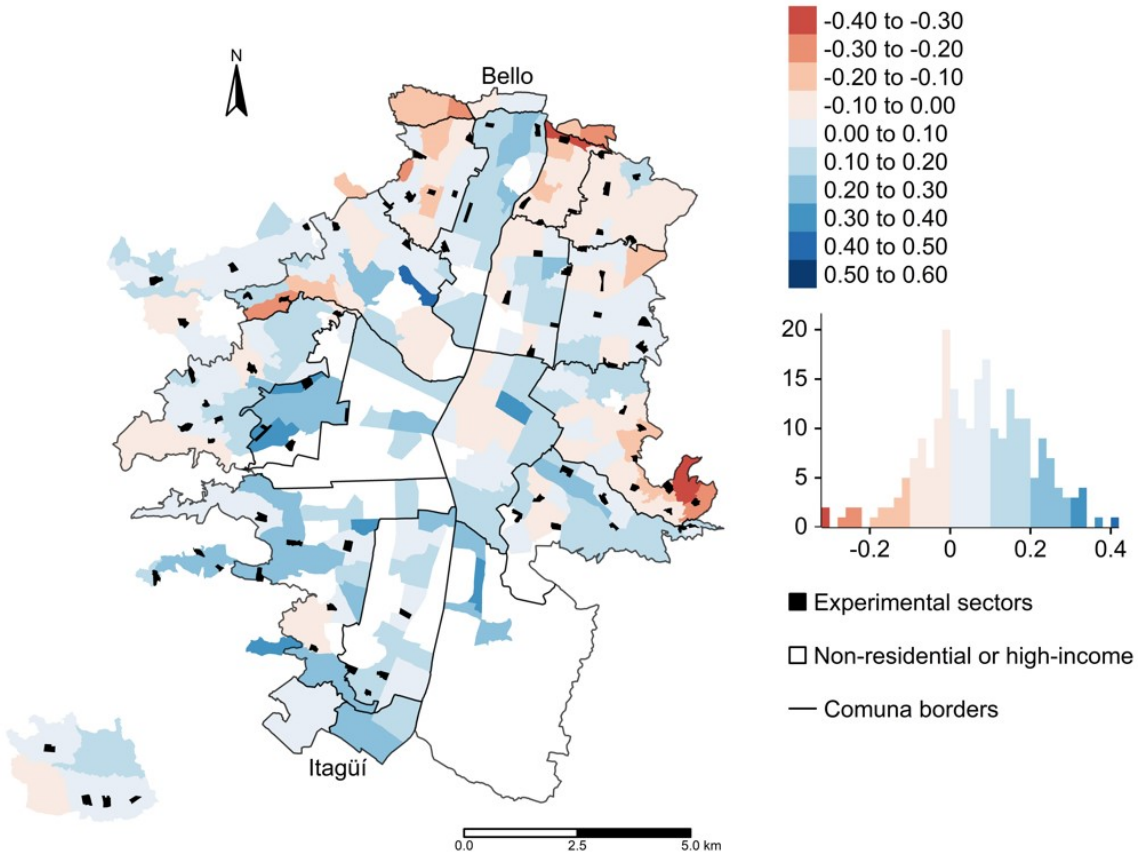
The intervention began in April 2018, midway through the administration of Federico Gutiérrez—a center-right politician who, like many former mayors of Medellín, ran for the Presidency at the end of his term.[18] Generally speaking, this aspiration to higher office is a key motive for single-term mayors to produce broad-based policy successes during their term.

3.1 Preexisting models

The idea for an intensive, localized civilian-led effort came from a small and little-known public safety effort in one of Medellín's poorer and under-served neighborhoods, called La Loma. It had a population of roughly 20,000. A small unit in the Alcaldía had assigned 7 outreach staff to the barrio, called "liaisons." From 2012–17, these staff set out to improve local governance and state legitimacy by: (i) helping existing community organizations better organize themselves to solve local problems (such as neighborhood cleanliness or idle youth); (ii) connecting residents in need to existing city services (such as dispute resolution and mediation, or family and mental health services); (iii) bringing neighborhood problems to

[18]He came in third in the 2022 elections, missing a run-off by a relatively small margin.

Figure 3: Relative state-combo governance levels by barrio, and location of experimental sectors



Notes: The figure displays average levels of relative state-combo governance in every low- and middle-income barrio, using the average of all 17 items from Table 1, averaging across all survey respondents in the barrio. We did not survey high-income residential neighborhoods or non-residential areas, all of which which appear in white. We also display the location of the 80 experimental sectors.

the attention of city agencies (such as garbage collection or broken lights and playgrounds); and (iv) facilitating community-police meetings and communication.

Medellín already had an established city-wide program of community liaisons, but they were few in number—about one per comuna, or roughly 1 per 540 blocks. The La Loma intervention raised this more than 20-fold, to roughly 1 liaison per 23 blocks. Our qualitative interviews and observations in La Loma suggested that the intensification of street-level staff increased community organization, access to municipal services, and legitimacy of the state rose.

We brought this initiative to the attention of the Mayor and Secretary of Security. They decided to expand and evaluate the program to assess the viability and returns to a city-wide expansion of liaisons and municipal agency attention.

3.2 Experimental sample

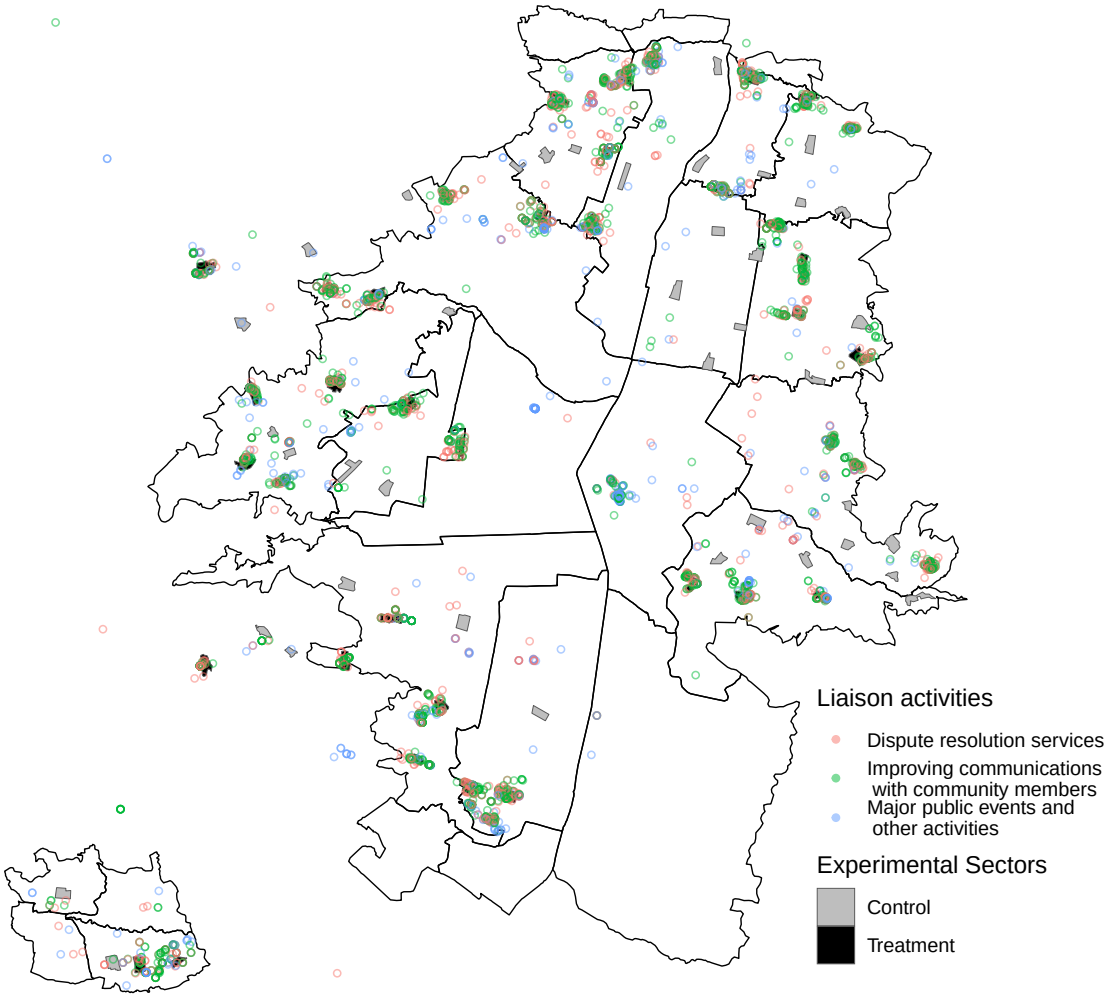
The Alcaldía decided that the appropriate level of intervention would be a "sector"—an informal but well-defined neighborhood, far smaller than a formal barrio, usually with about 200–600 households (about 1,000–3,000 residents), covering 5–10 medium-density city blocks.

The Alcaldía had funds to intervene with the desired level of intensity in 40 sectors. Therefore, they set out to identify 80 sectors for an experimental sample, drawn from the city's low- and middle-income barrios. Figure 3 displays these sectors.

The experimental sectors were not chosen randomly, but they are broadly representative of the city's neighborhoods in terms of their demographics, geographic features, and variation of state and combo governance.[19] State penetration varies widely across these neighborhoods, however. Some had a long tradition of street-level service delivery from the Alcaldía. In others, residents told us this was one of the first times they had seen a city representative other then the police in their neighborhood. All 80 sectors had some degree of combo presence, but this too varied widely (much like the rest of the city). In this respect, our experimental sample is fairly representative of the variation in Medellín. Appendix Figure A.1 illustrates this, showing that the variation in both state and combo governance in the experimental sectors is broadly representative of the city's low- and middle-income barrios.

[19]For instance, Appendix Table **??** compares blocks in the city-wide representative sample to blocks in the control sectors along a variety of pre-treatment characteristics. As for levels of governance and legitimacy, Columns 5 and 6 of Table 1 above show that average levels of relative state governance and legitimacy are extremely similar in the city sample and the control sectors in the experimental sample.

Figure 4: Alcaldía and liaison activities by experimental sector



Notes: Liaisons were instructed to report their major events and referrals, including the location. This figure depicts all major liaison activities conducted, overlaid atop a map of treatment and control sectors. We group the interventions in three subgroups based on the their goals. Treatment sectors are mostly obscured by the density of activities, while control communities receive relatively few.

3.3 Intervention activities

The city attempted intensified normal civilian public safety and coexistence services in 40 of the 80 sectors for 20 months, beginning April 2018 and ending in December 2019. Control sectors received normal services. The intervention had both top-down and bottom-up aspects.

Top-down

The Alcaldía first created an inter-agency task force to respond to local concerns. This could include normal services—e.g., poor trash pickup or broken playground equipment—but the task force also tried to respond to security concerns, including a lack of attention from the city's dispute resolution officers and family services. There was no change in top-down police or other criminal justice attention to the sectors, as both are outside the purview of the Secretariat of Security.

The city also tried to improve communications and relationships with the sector residents. Officials from the Alcaldía and local police commanders were asked to attend twice-annual government-community meetings in the treated sectors, known as *Consejos de Convivencia*, where they and community members would agree on mutual responsibilities and commitments.

Finally, the Mayor's office also organized a large one-time event called *Caravana de la Convivencia*—a weekend-long street festival in each sector where, in addition to music, food, and entertainment, representatives from each agency were on hand to explain their services in detail and identify residents in need of assistance.

Bottom-up

The city also assigned a full-time street-level bureaucrat—a liaison—to each treated sector to facilitate this top down process. Normally, the city has one liaison for each of the 16 comunas—roughly 1 per 540 blocks. The city contracted 40 additional liaisons for this intervention, one per treated sector. Thus treated sectors had 1 liaison per 9 blocks—a 60-fold increase in street-level staffing. Liaisons had several roles, including:

- Collect and formally register community concerns to the inter-agency task force in the Alcaldía, and lobby to see that these concerns are addressed
- Organize community events and meetings, including but not limited to the *Consejos de Convivencia* and *Caravana de la Convivencia* mentioned above

- Help community organizations coordinate local collective action (e.g., coordinating garbage spots and dog excrement norms)
- Provide training to community leaders in dispute resolution and related skills, and encourage them to take an active role in resolving local issues
- Proactively identify individual and neighborhood problems and referred them to the relevant city agency for assistance (e.g. connecting residents with interpersonal conflicts to the comuna's *inspecciones* for dispute resolution or *comisarías* for family disputes)
- Work with police officers to better inform community members of the "police code"—the country's legal guidelines for dealing with and correctly reporting nuisances, misdemeanors, and crimes, what officers were permitted to do, when to call them, and when to approach the Secretariat of Security

Liaisons were not residents of their assigned community. Rather, they were professional staff hired for this position, and were similar to the city's existing cadre of professional liaisons: university-educated men and women ages 25–35.[20]

Limitations and non-compliance

Generally speaking, the bottom-up liaison activities were implemented with high fidelity, while the top-down attention and services was partially successful.

The liaisons had a high level of street presence and visibility for almost two years. We and the city closely monitored liaisons. They had weekly targets and quotas for neighborhood events and resident referrals, and their activities and task force responses were formally logged and geolocated. From these records and our spot visits we confirmed that they spent 3–6 days or evenings per week in their sector, held regular community events, and generally met their referral quotas, all within the few blocks they were assigned to. Qualitatively, our general impression was one of autonomous, enthusiastic, hardworking efforts by skilled young professionals.

Figure 4 displays the formally-logged liaison activities, overlaid atop a map of treatment and control sectors. The figure illustrates the concentration of activities within experimental sectors, which are in most cases obscured by the density of activities reported in or nearby. On average, liaisons logged 36 official events within the borders of treatment sectors and 0.23 within control sectors (see Appendix Table A.2). Many activities were held nearby the

[20]They were employed on a contract basis through a non-governmental organization with extensive experience providing neighborhood outreach. Because of regulations on expenses in full-time personnel, roughly 70% of municipal staff are employed through such agencies.

sector, however, because community centers, meeting spots, and community organization offices were not always located in the 5–10 block sector itself. Liaisons logged an average of 72 major events within a 125 meter buffer of treated sectors, and 2 within a similar buffer around control sectors.

Unfortunately the city does not maintain similar administrative data on the activities of the its regular 16 liaisons covering all comunas, but given their limited reach (1 per 540 blocks compared to 1 per 5–10 blocks in treated sectors) we presume that control sectors received no more than 1–2 percent as many events or activities.

As for top-down performance, the liaisons reported that the central task force met some of the community's requests, but not all. Unfortunately, there is no formal administrative data on the top-down task force's activities or compliance. We interviewed all liaisons after the 20-month intervention, however. On a scale of 0 to 1 (from full compliance to complete failure to deliver) liaisons rated the wider state compliance roughly 0.34, meaning the state "sometimes" failed to deliver on the requested support. We return to this below. Thus, although the liaisons represent a 60-fold intensification of street-level municipal staff, we roughly estimate that that treated sectors received a doubling or tripling of top-down Alcaldía attention and services.

Finally, we monitored gang reactions to and interference in the intervention. Two-thirds of liaisons reported no interference whatsoever. The other third mostly said that the combo was mainly watchful, such as observing public events and meetings from a distance. Liaisons reported that combos rarely interfered with their work or attempted to take credit for services delivered. The exceptions mostly affected the first few weeks of the intervention and afterwards the implementation ran smoothly. For example, in two sectors, the combo initially prevented two liaisons from entering into the community for the first 2–3 weeks, but once the liaisons were able to explain their job and role, they were permitted to enter and perform their jobs without interference. This is consistent with logged liaison events, which if anything were slightly more frequent in sectors with initially low relative state governance (Appendix Table A.2).

3.4 Theoretical motivation and conceptual framework

Our main hypothesis was that by improving public-service delivery, providing non-criminal alternatives for dispute resolution and contract enforcement, and strengthening the ability of formal and informal groups to identify problems and solutions to everyday community problems, the Alcaldía could increase its legitimacy and citizen use of its services—including turning to the state first in the event of a security concern.

In principle, the intervention (if successful) could also reduce citizen dependence on combo governance. Initially, therefore, we viewed the intervention through the lens of duopolistic competition. The state and the combo were offering residents distinct but substitutable governance services. Should one side exogenously increase production, it was possible that its relative share of services should rise, crowding out combo rule. We illustrate this theoretical possibility with a simple model of Cournot competition in Appendix B.

Subsequently, observing the intervention in progress, we moderated the view that increased civilian governance could crowd out gang services, at least in the short run. Citizens have a huge range of everyday disputes, minor forms of neighborhood are commonplace, and neither the state, the combo, nor community organizations respond to all. Both the state and combo governance measures are well below 0.5 in our indexes in Table 1 above. Thus more of one service will not necessarily crowd out the other, given unmet demand. Nonetheless, we continued to expect that state legitimacy would rise by increasing the quality and quantity of services.

We also moderated our views of crowding out combo rule because, as noted above, combos generally regarded the liaisons and regular city services as benign. Their main concern were the police. Indeed, in a longer-run companion study, we discovered that police presence could have the opposite effect on gang rule. In order to protect drug revenues, gangs may decide to provide governance services to reduce police presence and collaboration with civilians. This could reduce the extent to which state and combo governance are substitutes, and raises the possibility that they are strategic complements. A quasi-experimental analysis of a 30-year increase in both policing and municipal survey suggests both mechanisms are at work, and that over several decades the strategic complementarity may dominate (Blattman et al., 2023a). Whether this is also true over a 20-month horizon and this purely civilian intervention is unclear.

4 Data

4.1 Outcomes

We prespecified two primary outcomes: *Relative state legitimacy* and *Relative state governance*.[21] As secondary analyses, we consider variants of each index. Of the 17 governance activities. We separate governance activities that are more or less likely to involve a police

[21]We preregistered the design and outcomes in April 2018, then again prior to final data collection. See the Journal of Development Economics registered report for the the final analysis plan (https://drive.google.com/file/d/1QiEegA-GDdO34-QONMTcxe5bD6MO7nFI/view) and the American Economic Association registry for previous rounds (AEARCTR-0002622).

response (versus a civivlian agency). And we look at residents' perceptions of *absolute* levels of state and combo governance and legitimacy rather than the prespecified relative measures.

Each of these measures come from the December 2019 endline survey, conducted 20 months after the intervention began. The survey was roughly 30 minutes long, and was delivered in person by enumerators on handheld tablets. Enumerators were employed by one of the survey's largest survey firms, and had no affiliation or identification with the intervention.

The experimental sample and the representative city sample are distinct, though all answered the same instrument. The city survey interviewed roughly 21 residents on 7–10 randomly-selected blocks per barrio. None of these blocks are located in the experimental sectors. For the experimental sample, we interviewed 3–5 residents on an average of 9 blocks in each sector (for a total of roughly 30 respondents per experimental sector). Column 6 of Table 1 above summarizes endline relative governance and legitimacy in control sectors.

Naturally, we are concerned that citizens may under-report gang activities, attenuating estimated treatment effects somewhat. Section 6.4 discusses measurement error, and why it is unlikely to influence our results. Briefly, combos are a part of everyday life and not systematically stigmatized. We also designed a survey experiment and find no evidence of response bias.

Finally, in addition to these survey-based outcomes, we also use administrative data on crime reported to the police during the 20-month period of the program, and calls to the police through the city's emergency line. The intervention did not directly address crime, but liaisons did work to improve community-police communications and also helped community organizations organize to improve neighborhood orderliness.

4.2 Baseline data

For baseline data on the sectors, in February 2018 we surveyed three officials per sector for their assessment of: relative governance service provided by the combo and the state; relative street presence of the combo and the state; and their perceptions of local security and drug use. We also have rich, geolocated administrative data including distance to various state and criminal headquarters, crimes committed, and demographics.

5 Experimental procedures

5.1 Randomization and balance

We grouped the 80 sectors into 40 matched pairs using five baseline measures: relative state-combo governance; relative state-combo street presence; insecurity perceptions; administrative crime reports; and distance to public services and infrastructure. We then randomized one in the pair to treatment.[22] This produced the expected degree of balance along baseline covariates, as seen in Table 2.

5.2 Accounting for spillovers

To reduce the chance of interference between units, we selected sectors at least 250 meters distant from one another. A total of 40 intervention sectors also ensured that increased service delivery would minimize any decline in city attention to control sectors, since these represent less than 2.5 percent of all city blocks. In addition, we can use our representative city-wide survey to test for spillover effects into non-treated areas, by comparing blocks close to treatment sectors to those close to control sectors. We see no evidence of such spillovers, as seen in Appendix Table A.1.

5.3 Estimation

We estimate intent-to-treat effects via the simple OLS regression:

$$Y_{isb} = \alpha + \beta T_s + \Theta X_{isb} + \epsilon_{isb}$$

where Y is the outcome from survey respondent i in sector s and matched pair b; T is an indicator for random assignment to treatment; and X is our prespecified vector of controls, including sector-pair fixed effects (the randomization strata) and the five baseline indices we used to match blocks. We cluster standard errors at the sector level.

With this design, we estimated we were powered to detect improvements in state governance and legitimacy of about 12% with a two-tailed test.

Heterogeneity analysis We prespecified one heterogeneity analysis, by initial levels of state presence. We estimate subgroup impacts via the OLS regression:

$$Y_{isb} = \alpha + \beta T_s + \delta(T_s \times Low_s) + \lambda Low_s + \Theta X_{isb} + \epsilon_{isb} \quad (1)$$

[22] See Bai (2022) on the optimality of these matched-pair designs.

Table 2: Baseline summary statistics and test of balance

Covariate	Means		Regression Difference		
	Control	Treated	Coeff	p-value	SE
Baseline indices used for matching , main control vector					
Additive index of combo presence and governance	0.00	-0.02	-0.02	0.92	0.22
Baseline - Combo Governance Index (Relative to State)	0.00	-0.02	-0.02	0.91	0.22
Standardized index of perceived insecurity and drugs	0.06	-0.07	-0.13	0.58	0.22
Index of crime	0.09	-0.12	-0.21	0.35	0.22
Index of distance from public goods and services	-0.14	0.14	0.28	0.22	0.22
Other baseline covariates					
Respondent age between 18 and 25	0.19	0.19	-0.00	0.98	0.02
Respondent age between 26 and 40	0.29	0.31	0.01	0.52	0.02
Respondent age between 41 and 64	0.39	0.37	-0.01	0.58	0.02
Respondent is business owner	0.20	0.20	0.00	0.90	0.00
Multidimentional Poverty Index (2018)	14.34	17.26	2.93	0.20	2.26
Block Longitude	-75.58	-75.58	-0.01	0.37	0.01
Block present in 1970	0.50	0.44	-0.06	0.54	0.10
Median age (2005)	27.20	26.31	-0.90	0.41	1.08
Total women (2005)	133.86	142.04	8.18	0.53	12.90
Total non-mestizo polulation (1993)	0.53	0.61	0.08	0.66	0.18
Median age (1993)	24.16	24.71	0.56	0.60	1.06
Share of women (1993)	0.53	0.52	-0.01	0.68	0.01
Distance to the respective razon headquarters (100 meters)	17.28	19.12	1.84	0.66	4.15

Notes: This table reports treatment and control group means and a test of balance for the covariates used to match treatment and control sectors (the first five variables) and for some of the covariates selected by the lasso method as prognostic of endline absolute state governance. Regression differences come from an OLS regression of each covariate on an indicator for treatment, calculated at the individual survey level, clustering standard errors at the sector level.

Variables are the same as in Equation 1, and *Low* is an indicator for neighborhoods with below-median state governance. In that case, β estimates the program impact on relatively high-state neighborhoods, δ estimates the difference between high and low neighborhoods, and $\beta + \delta$ is the impact on low-state neighborhoods. Appendix Table A.3 shows that treatment-control balance within the subgroups.

Our prespecified measure of state strength comes from baseline interviews with three community and city leaders per sector who were asked for their assessment of the *relative* role of the state in providing everyday governance. These data are described in the next section.

6 Results

6.1 Quality and consistency of implementation

We start with an analysis of treatment compliance—a set of "first-stage" indicators for whether the program was implemented as designed. As we will see, heterogeneity in implementation quality appears to be the key driver of program impacts.

Figure 4 above and Appendix Table A.2 demonstrate that liaison activities were largely concentrated in and nearby the treated sectors, but these come from liaison administrative records alone. To assess whether residents noticed the increase in general municipal government activity, our 2019 endline survey included questions on Alcaldía personnel and activities in their neighborhood. Table 3 reports average treatment effects on six survey questions as well as a family index averaging all six responses (to reduce the number of hypotheses tested). The survey did not ask about police activity.

To our initial surprise, we see no evidence that residents noticed the increase in the Alcaldía's activity, or that they attended more events. The average change in the overall index is 0.002—less than a 1 percent increase over the index average. Only one measure is positive and statistically significant—seeing municipal staff in the sector, which rose roughly 6 percent. Another, however, is actually negative—knowing about community events, which fell roughly 15 percent. This is a striking finding given the 60-fold increase in street-level staff and the closely-monitored liaison compliance.

We anticipated that program impacts could depend on initial state governance and legitimacy. We knew this heterogeneity could go in either direction, but our hypothesis was that there are diminishing marginal returns to state personnel and attention—meaning the intervention would be most noticed and effective in the least-served neighborhoods. Table 3 reports program impacts according to this prespecified heterogeneity. Columns (3) and (4)

Table 3: Did citizens notice increases in state activity? Survey-based measures, average treatment effects, and heterogeneity by initial relative state governance

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Index of first-stage variables (0-1)	0.33	0.002	0.041***	-0.039***	0.080***
	(0.27)	(0.010)	(0.013)	(0.012)	(0.018)
Attended State events	0.21	-0.007	0.025	-0.042**	0.067**
	(0.41)	(0.014)	(0.019)	(0.019)	(0.026)
Knew about State events	0.52	0.020	0.090**	-0.055*	0.145***
	(0.50)	(0.025)	(0.036)	(0.032)	(0.048)
Attended community events	0.10	-0.014	0.005	-0.034**	0.039*
	(0.30)	(0.011)	(0.016)	(0.013)	(0.021)
Knew about community events	0.30	-0.046**	0.004	-0.098***	0.102***
	(0.46)	(0.019)	(0.020)	(0.026)	(0.032)
Saw mayoral employees	0.61	0.036**	0.068***	0.001	0.067**
	(0.49)	(0.016)	(0.024)	(0.020)	(0.030)
Interacted with mayoral employees	0.24	0.022	0.052*	-0.009	0.061*
	(0.43)	(0.015)	(0.027)	(0.015)	(0.032)

Notes: This table reports answers to six Yes/No questions in the survey regarding whether residents and businesses noticed municipal employees and events or attended them. Each row is a different dependent variable. Column (1) reports control sector means. Column (2) intent-to-treat (ITT) estimates of program impacts using Equation 1. Columns (3) to (5) report treatment heterogeneity using Equation 1—treatment effects in sectors above and below the median level of baseline relative state governance, and the difference between the two groups. The unit of observation is the individual survey respondent, and we cluster standard errors at the sector level (the unit of randomization).

report ITT estimates in sectors with above and below the median levels of initial governance, and Column (5) reports the difference between the two.[23]

We see divergent effects depending on the initial levels of state presence, but not in the direction we expected. In initially high state-governed sectors, residents report a 12 percent increase in municipal activities and participation. In sectors where the state governed relatively less, they reported a roughly 12 percent decline. The total divergence in the index between the two kinds of sectors is dramatic— 0.08, equivalent to 24 percent of the index mean.

We see this divergence in every component of the index (Column 5). In sectors with relatively greater initial state rule, residents and businesses were dramatically more likely to notice and interact with municipal staff and be aware of and attend community events.

Post-treatment qualitative and quantitative investigations To better understand this divergence, we collected post-treatment qualitative and quantitative data in early 2020. Altogether, these data suggest that the street-level liaisons logged activities consistently in most sectors. Unfortunately, in the neighborhoods where the central state had the least relative presence, there are indications that the city's central task force and other senior actors had trouble delivering on the promises. Top-down compliance appears to be lowest in the sectors with low initial state presence.

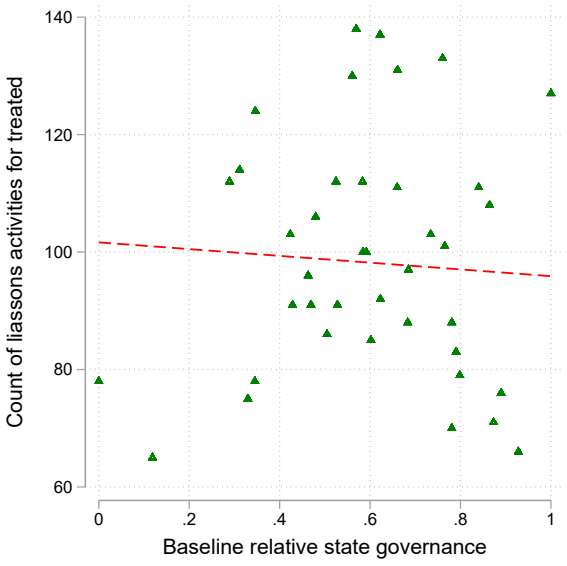
First, the administrative data suggest that liaison compliance and effort were high in all neighborhoods. Panel (a) of Figure 5 uses program data on all events and activities logged by the liaison in the 40 treated sectors. It plots the number of documented activities by baseline relative state governance. Activities are numerous and unrelated to initial government presence. Indeed, if anything, liaisons logged slightly more official events in and around below-median governance sectors (see Appendix Table A.2. We need to regard liaison self-reports with some caution, but these reports are consistent with our qualitative observations that liaisons were highly active in their sectors, physically present 3–6 days a week for 20 months.

The same is not true of the Alcaldía's wider activities, however, including attention from higher-level politicians, bureaucrats, and the task force. Panel (b) of Figure 5 reports the degree to which the central administration failed to deliver on promises in the 40 treated sectors. These come from a post-program survey of liaisons, which asked the liaisons to rate state compliance on a scale of 0 to 1, from full compliance to complete failure to deliver. On average, liaisons rated the Alcaldía's compliance at roughly 0.34, meaning the state

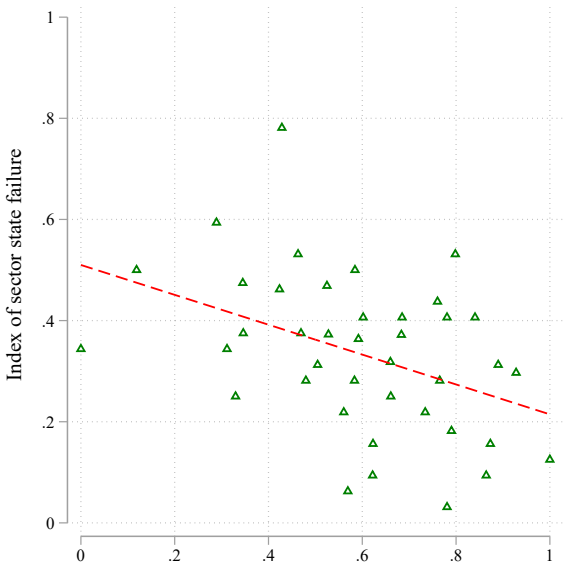
[23]In all tables, above-median treatment effects come from the estimated coefficient on treatment, the difference between sectors comes from the estimated coefficient on the interaction term, and the below-median treatment effects are calculated as the sum of these two estimated coefficients.

Figure 5: How treatment experiences varied by initial levels of relative state governance (treated sectors only)

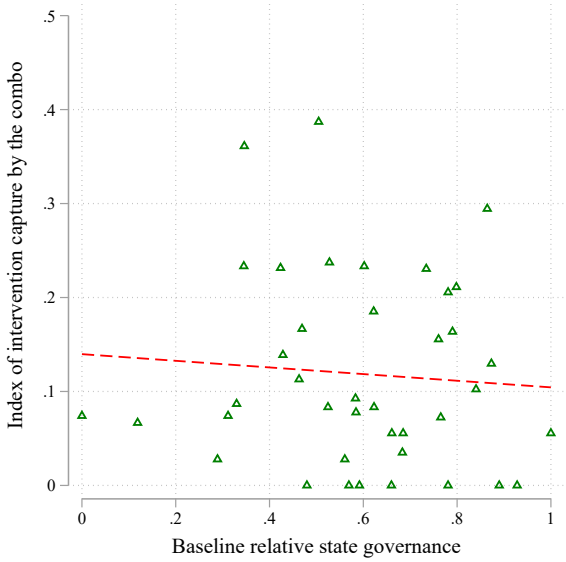
(a) *Count of treatment activities*



(b) *Failed promises of the wider state apparatus*



(c) *Instances of combo interference and capture*



Notes: The city required liaisons to log their activities, and Panel (a) reports the number of activities they logged, by levels of baseline relative state governance. Panels (b) and (c) contain data from a post-program survey of all liaisons. Based on their responses, we created two indexes. Panel (b) reports the frequency of various failures of the liaison or the wider state apparatus to deliver on promises. This includes a scale of the perceived frequency of failures and binary variables for whether specific local state agency failed. Values closer to 1 mean higher state failure. The data in Panel (c) capture the degree with which the combo interfered with liaison activities. This aggregates several measures: a scale for the frequency and difficulties of interaction with local gangs; a set of binary variables on whether local actors (including the gang) took credit for the intervention; and a set of binary variables for activities by which the gang helped the liaison. Values closer to 1 represent higher involvement from locals gang on intervention activities.

"sometimes" failed to deliver on the requested support. But they reported these failures twice as often in the sectors with relatively low initial state presence.

Based on our qualitative interviews and focus groups with all 40 liaisons, the most common were failures of the city to respond to community needs and meetings. Equipment might go unrepaired, for example. Or, as one liaison explained, they organized a meeting between the community and city officials, and the officials never arrived. Another said how they had publicized the new police code—which includes official guidelines for when citizens should call the police versus one of the civilian security and services agencies—but the residents were frustrated because the police did not follow it reliably.

These data suggest that expectations may have been raised by the program but inconsistently met by the central task force. A related possibility is that the presence of liaisons also raised expectations of service delivery beyond feasible levels, perhaps most of all in lower state presence areas. But even if expectations were not raised too high, it there are suggestions that the state apparatus was simply unable to follow through on basic promises in the neighborhoods where historically they did not have a strong presence, even in absolute terms.

Why do we observe adverse effects in poorly-governed neighborhoods? These patterns can help explain why, in initially well-governed sectors, residents were more likely to notice staff and events and attend activities. It is less clear why residents in relatively poorly-governed should report knowing about and attending fewer events compared to control sectors.

One possibility is that, over 20-months, residents in poorly-governed sectors became aware that events were held, but too late to attend. The presence of liaisons may have raised the expectations and reference point for what a state or community event is compared to control sectors. Moreover, the answers to these questions are inherently subjective, and may simply be capturing residents' sentiments towards the liaisons and activities—where expectations went unmet, there may be negative attitudes towards events. Program impacts on reports of state governance (below) are similar and consistent with all these explanations. Ultimately, however, we do not know why we observe this pattern.

Combo response Finally, we see no evidence that combo responses shaped the heterogeneous results. Granted, combos noticed an increased presence of the Alcaldía. In qualitative interviews, for instance, almost all liaisons described having to explain their presence to the combo. Thus, we are confident that combos were generally aware of increased state activity from the beginning.

Most of the evidence, however, suggests that the combos did not react to the presence of these liaisons and the attention of the task force. For example, Panel (c) of Figure 5, captures the degree with which liaisons reported that the combo interfered with their activities.[24] We do not see much evidence that the combo tried to capture the liaison's activities or take credit for their work. The levels are low and there is little relationship with initial state presence. Nor do we see any evidence that combos escalated their governance services or legitimacy in response to the state. The coefficients on the combo indexes in Table 4 are generally close to zero. We do not want to understate the role of combos, since low initial state presence could be endogenous to gang strength. But there is nothing in our results to suggest direct interference. The results on combo governance and legitimacy below reinforce this conclusion.

6.2 Impacts on governance and legitimacy

We see no evidence the intervention improved state performance overall, consistent with our first-stage results. Table 4 reports program impacts and heterogeneity on the primary outcomes—relative state governance and legitimacy—as well as absolute levels of both for the state and combo. Column (2) reports average treatment effects. Three of the four signs are negative, and we actually observe a 0.024 *decrease* in relative state governance, statistically significant at the 10% level.

Nonetheless, we see see signs that program impact varied by initial levels of relative state governance. Columns (3) to (5) report program impacts on sectors with initially high and low relative state governance.[25]

First, relative state legitimacy rises by 0.053 in initially well-governed sectors. This is equal to 41 percent of the state-combo difference in legitimacy (Column 1) and 9 percent of the average level of absolute state legitimacy of 0.57. There is also a small, statistically non-significant decrease in legitimacy below-median sectors. As a result, the difference between the two types of sectors is even larger—0.074, equivalent to 13 percent of the city-wide average. We see the same pattern if we look at absolute state legitimacy.

We break down legitimacy into its five component questions in Appendix Table A.5. The survey asked about the legitimacy of the police and municipal government separately, and we show program impacts on both. On average across all sectors, there is evidence that satisfaction and trust with mayoral staff rise, but overall legitimacy does not significantly improve.

[24] See the foot of the figure for details on how we built the intervention capture measure.

[25] In addition to looking at above- and below-median comparisons, we committed to report treatment effects in the four major quartiles, as seen in Appendix Table A.4. With just 20 sectors per subgroup, however, that analysis is under-powered.

Table 4: Program impacts on governance and legitimacy: Average treatment effects and heterogeneity by baseline governance quality

		ATE	Het. by baseline rel. gov.		
			Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Relative state governance index	0.07	-0.024*	-0.016	-0.033*	0.017
	(0.31)	(0.014)	(0.020)	(0.018)	(0.026)
State governance index	0.41	-0.013	-0.007	-0.019*	0.011
	(0.26)	(0.009)	(0.012)	(0.011)	(0.016)
Combo governance index	0.35	0.009	0.007	0.012	-0.004
	(0.28)	(0.012)	(0.018)	(0.016)	(0.023)
Relative state legitimacy index	0.13	0.017	0.053**	-0.021	0.074**
	(0.32)	(0.018)	(0.023)	(0.029)	(0.037)
State legitimacy index	0.57	0.010	0.029***	-0.010	0.039***
	(0.21)	(0.007)	(0.010)	(0.010)	(0.014)
Combo legitimacy index	0.44	-0.006	-0.022	0.011	-0.033
	(0.28)	(0.014)	(0.017)	(0.023)	(0.029)

Notes: The table reports intent-to-treat (ITT) estimates of program impacts and treatment heterogeneity using Equations 1 and 1. Each row is a different dependent variable. For the ITT estimates (Column 1) we regress each dependent variable on an indicator for treatment and our prespecified control vector: 5 baseline variables and sector-pair fixed effects. The unit of observation is the individual survey respondent, and we cluster standard errors at the sector level (the unit of randomization). Columns (2) to (4) report treatment effects in sectors above and below the median level of baseline relative state governance, and the difference between the two groups. Both households and businesses were surveyed on governance levels (N=2,379), but only households were surveyed on legitimacy and hence there are fewer observations (N=1,910).

In above-median sectors, however, both police and mayoral legitimacy rise to roughly similar degrees. In below-median sectors, police legitimacy falls while mayoral legitimacy does not.

Next we turn to governance, where we see weaker evidence of heterogeneous effects. There is no indication the program increased perceived state responsiveness in above-median sectors. But relative state governance declines by 0.033 in below-median sectors, significant at the 10 percent level. This decline is equal to roughly half of the relative state-combo difference in governance, and about 8 percent of average absolute governance. The difference between initially well- and poorly-governed sectors is not statistically significant, however.

These patterns hold if we break the governance index into its 17 components and into more and less police-related actions, as reported in Appendix Table A.6. We classify the 17 forms of disorder into 8 that are more likely to elicit a call to police or a police officer response, and 9 that are commonly solved by a variety of city and community actors. This is the only way to assess potential differences between police and mayoral staff because, unlike legitimacy, we did not ask all 17 questions for the police and Alcaldía separately (to keep the survey brief). In below-median sectors, the fall in police governance is slightly greater than the fall in non-police governance, but the difference between the two is not statistically significant.[26]

Finally, we see little evidence of a combo response. Table 4 shows no evidence of program impacts on combo governance or legitimacy. This is notable given the results of a companion study of gang reactions to a sustained 30-year increase in police and Alcaldía attention. In Blattman et al. (2023a), we found that combos responded strategically to long-term state presence by increasing their governance over civilians and fostering legitimacy. The results, plus interviews with gang leaders, suggested that the combos were seeking to protect their retail drug businesses. They reduced street disorder and increase civilian loyalty to minimize police presence in the neighborhood. We see no evidence of a strategic combo response in above-median sectors, however. This could be because the combo does not feel threatened by non-police state presence. Alternatively, 20 months of mayoral attention may have been insufficient to provoke a combo response. Qualitatively, our interviews with liaisons are consistent with the combo not feeling threatened by their activities. Indeed, to the extent that the community organization and crime reduction they engendered actually reduced the need for police calls, the treatment may have reduced the threat to combo drug rents. We turn to this next.

[26]The survey also included a number of supplementary measures of efficacy, including the speed of response, ease of accessing services, and the value placed on the actor. We report these in Appendix Table A.8. We see no evidence of that residents perceived an improvement in the speed or ease of contacting the police or mayoral staff, on average or in above-median sectors. There is some evidence that perceived value of the Alcaldía declined in below-median sectors.

6.3 Impacts on crime and calls to police

Finally, we see similar program impacts on crime and calls to police: no effect overall, but in initially well-governed sectors the intervention appears to have reduced property crimes, fights, and associated emergency calls to the police.

Table 5 examines reported crime and Table 6 examines resident calls to the police. Both measures are for the 20-month intervention period, and both collect all crimes and calls geolocated within a 125 meter radius of the sector.[27] Unfortunately, there are no data on the frequency of police patrols by sector, so it is impossible to know the effects of the intervention on normal police presence outside of these demand-driven calls.

To reduce the number of outcomes and hypotheses tested, we focus on two summary outcomes: (1) an index of all crimes reported, with crimes are weighted by their severity (proxied by sentence length guidelines for each crime); and (2) a count of police calls. The tables also list the major components of these summary measures. Note that of this analysis was prespecified, and should be treated as exploratory (especially the many subcomponents).

Before getting to results, it is important to note what each measure captures. In Colombia, crimes can only be reported at a comuna's central police station. Speaking to a police officer or calling the police will not result in a formal crime report (a fact that is widely known). Thus reporting requires traveling up to a kilometer and can take several hours to complete forms. Evidence from other Colombian cities like Bogotá suggests that thefts of vehicles and other high-value items are frequently reported (for insurance purposes), as are crimes that result in serious injury or death. Because of the hassle, however, most petty crime goes unreported (Blattman et al., 2021).

Emergency calls to the police come from local residents and businesses. The vast majority report of callers are reporting a street fight, a case of domestic abuse, or a drug-related complaint—either a concern about a drug seller or (more commonly) drug users causing a public disturbance or loitering. Thus they are a measure of disorder and a perceived need for emergency intervention. All calls are logged and geolocated to an address when police respond.[28] The coding of type of incident is relatively crude, however, and we can primarily distinguish between physical altercations (mixing domestic and street disputes), narcotics-related nuisances, and armed fights.

Starting with crime, the weighted index falls by 0.096 in above-median sectors, significant

[27] We chose 125 meters for both measures because of our requirement that every sector be at least 250 meters from one another. 125 meters is half this distance, ensuring no overlap. Patterns are qualitatively similar for other radii.

[28] Administrative logs say that more than 97 percent of calls receive a police response and are geolocated. We do not know the location of the 3 percent of unresponded calls, and so cannot assess program impacts on police response. Since nonresponse is low, it seems unlikely to qualitatively affect the results.

Table 5: Program impacts on crime index components: Average treatment effects and heterogeneity by baseline governance quality

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Sentence-weighted crime index	0.35	-0.031	-0.096*	0.052	-0.147*
	(0.26)	(0.039)	(0.053)	(0.052)	(0.074)
Homicides	0.04	0.026*	0.026	0.029	-0.003
	(0.05)	(0.014)	(0.020)	(0.020)	(0.028)
Vechicle thefts	0.33	-0.007	-0.074	0.080	-0.154
	(0.31)	(0.051)	(0.071)	(0.071)	(0.100)
Thefts and robbery	1.44	-0.298	-0.645**	0.140	-0.786*
	(1.48)	(0.229)	(0.313)	(0.313)	(0.440)
Assaults	0.64	-0.061	-0.132	0.015	-0.147
	(0.39)	(0.062)	(0.088)	(0.088)	(0.124)

Notes: The table reports summary statistics and treatment effects for the sentence-weighted crime index in Table 4 and its four main components. Each row is a different dependent variable. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4.

at the 10 percent level. This is a large change—a 28 percent decline relative to the 0.35 control group average. The divergence between above- and below-median sectors is even greater, a 0.147 decline—42 percent of the control group average. These declines appear to be largest for thefts and robbery. There is no evidence of a decrease in violent crimes. Indeed, we see weak evidence of a rise in homicides overall. We must treat all index component analyses as suggestive, however, and we have not adjusted standard errors for multiple hypothesis tests.

Note that these reductions in crime are unlikely to arise from differential reporting of crime in treated and control communities. Residents in initially well-governed treated sectors view the state as more legitimate, and so if anything should be more willing to report crimes to the state. Moreover, the intervention explicitly educated communities on the police code and facilitated semi-annual meeting between the community and local police commander, thus making them more familiar with reporting requirements. In principle, these factors could increase crime reporting rates in treated sectors, leading us to understate treatment effects.

Table 6: Impacts of treatment on resident calls to police

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Difference
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Resident calls to police	135.76	-25.026	-55.260**	8.474	-63.733**
	(102.30)	(16.136)	(22.392)	(22.357)	(31.507)
Physical altercations	93.34	-15.179	-32.502**	2.623	-35.125*
	(58.36)	(10.086)	(14.179)	(14.158)	(19.952)
Narcotics related incidents	30.97	-9.105	-21.228	5.328	-26.556
	(60.11)	(9.474)	(13.295)	(13.274)	(18.707)
Armed incidents	11.45	-0.742	-1.529	0.524	-2.053
	(7.97)	(1.548)	(2.183)	(2.180)	(3.072)
Knife related incidents	9.25	-0.942	-2.108	0.437	-2.545
	(6.95)	(1.240)	(1.760)	(1.757)	(2.476)
Firearm related incidents	2.20	0.200	0.579	0.086	0.492
	(2.44)	(0.643)	(0.898)	(0.897)	(1.264)

Notes: This table reports the total number of resident calls to the police emergency line over 20 months, including all calls made within each sector plus a 125 meter buffer zone around the sector. Calls are only geolocated within the city if the police actually respond to the call, meaning we cannot track impacts on unmet calls. But administrative records suggest that more than 97% of calls receive a response, and so are unlikely to affect out results. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4.

The evidence from police calls further suggests that, in initially well-governed sectors, municipal staff or the community itself is either dealing with everyday street disorder without the police, or successfully prevented forms of disorder. There is no change in calls to the police overall, but in above-median treated sectors calls fall by 55 relative to a control mean of 136—a 41 percent decline, significant at the 10 percent level. The divergence between above- and below-median sectors is even larger. We see this decline across every category of call, except for the very small number of firearm-related altercations. The largest decline (and the only statistically significant component) is in calls regarding unarmed street fights and domestic abuse.

6.4 Measurement error correlated with treatment

Our governance and legitimacy measures are self-reported survey data, and hence subject to potential response bias. The fact that we see similar results in administrative police crime and call data reduces this concern somewhat. But there are several additional reasons to

believe that measurement error is low, or at least not correlated with treatment status.

First, we do not believe that the presence of combos—a familiar and historical part of everyday life in our sectors—significantly distorted responses. We refined survey questions after dozens of qualitative interviews, fine-tuning language, questions, and approach to elicit truthful answers. For data collection, we used an independent survey firm that already conducted annual security surveys to avoid any connection with the intervention, and to minimize experimenter demand effects. They conducted all interviews anonymously and in private, typically indoors. In the context of a secret interview, we believe most respondents answered questions freely and truthfully. Three analyses are consistent with this conclusion.

Second, we can compare our approach against prior efforts. The city has run surveys in the past on "security fees" paid to the combo. City-wide, 19% of our business respondents and 7% of residents report making payments, with negligible non-response. A city survey conducted earlier in the same year reported a 3% payment rate, with 80% non-response. This suggests our approach was actually more successful in eliciting honest responses.

Third, for our results to be spurious would require a very specific pattern of misreporting. Residents would need to systematically under-report state governance or overstate combo governance only in the treatment sectors that had low initial government presence—in essence, the reverse of normal experimenter demand.

Fourth, we used a survey experiment to assess under-reporting in security fee payment—a measure our qualitative work deemed as one of the most sensitive questions on gang-related activities. We asked some respondents directly whether they paid (Direct response, or DR); others we used a randomized-response (RR) technique, where they privately flipped a coin and responded to the question honestly or not depending on the flip. In other contexts, this method has detected under-reporting of sensitive behaviors.

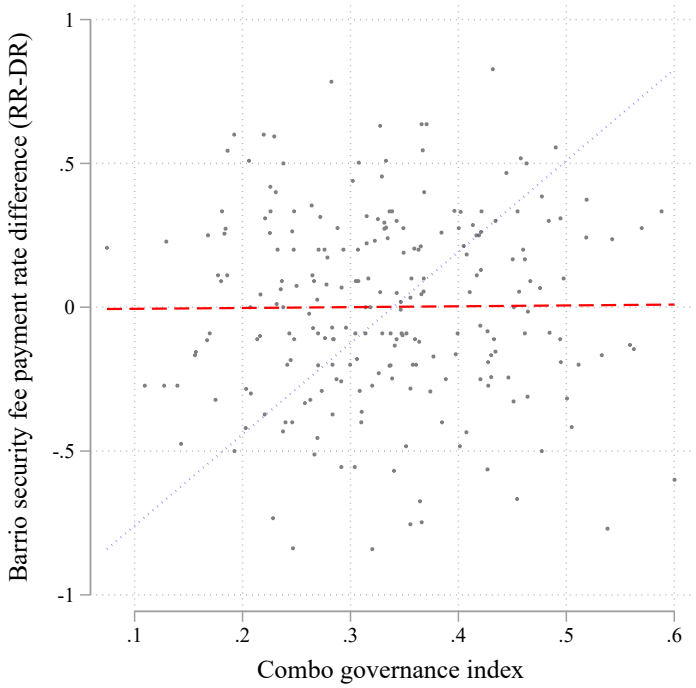
We see little differences in payment rates between the approaches, suggesting people did not misreport this topic. Randomized response elicited an extortion rate of 22.6% from businesses and 6% from households, compared to 19.4% and 7.8% when directly asked. The differences run in opposite directions and are not statistically significant.

We also see no correlation between assignment to treatment and a RR–DR difference. On average, across all respondents, randomized response results in 4 percentage points higher vacuna payments (not statistically significant). The treatment effect on this RR–DR difference is -0.05, with a standard error of 0.063 (p=0.430).

Figure 6 also calculates the difference between the RR and DR methods at the barrio level, and plot this difference against combo governance levels. A simple regression line is relatively flat at zero, indicating that misreporting is no more or less common in areas where the combos are more involved in daily life, and hence where legitimacy or fear could

potentially have influenced under-reporting.

Figure 6: Difference between randomized response (RR) and direct response (DR) to survey questions on combo "security fee" payment



Notes: This figure plots the difference between the RR and DR responses to the survey question on extortion against combo governance. Each point represents a barrio average from the 2019 representative city-wide survey. The figure also plots the 45-degree line and a fitted regression line.

7 Discussion and conclusions

Cities in Latin America and the United States are being pushed to tackle street disorder and insecurity using civilian forces rather than police. Many of these efforts involve unarmed approaches to crisis response. In contrast, in Medellín, we study an attempt to intensify existing "normal" city services and within-community capacity for managing everyday disorder and disputes. This took the form of increased attention from the central Secretariat of Security staff and sub-agencies, as well as dramatically raising the presence of street-level bureaucrats. Many American cities have parallel systems that combine local elected officials, city service agencies, and non-profit organizations contracted to organize the community or provide outreach.

Of course, intensifying services and community organization in 40 roughly 10-block sec-

tors of 1000–3000 people is different than a generalized increase in city staff and state and community capacity across a major metropolis. We must be careful in what we generalize from such a localized experiment.

Nonetheless, it is striking that we see no evidence of a positive impact on average. This runs counter to our prior expectations, especially because we closely monitored compliance with the most visible aspect of the program—the more-or-less daily presence of a liaison on the streets of treated sectors. At the every least, these results suggest civilian alternatives to policing are difficult and complex. The exercise also demonstrates the importance of community-level experimentation, which is rare.

We do not think, however, that the evidence points to a null or inconclusive effect. Interviews and our prespecified heterogeneity analysis suggest that the return to public investment was high in some neighborhoods, and negative in others. We must be cautious when we divide an experimental sample of 80 into subgroups of 40, but still, our results imply that returns depended on initial state capacity—or factors correlated with this trait.

One implication is that there may be increasing returns to state governance, at least at low initial levels of state governance. The returns to investments in state capacity when state capacity is low could be small or even negative. This might help explain a common feature of cities worldwide: high government attention to places where the state and community are already strong, and persistent neglect of regions where the state is weak or contested.

Another implication is that governments should take care manage expectations—their own and the public's. Our measures of governance and legitimacy are highly subjective, and it is possible that the reports of less street presence and lower state governance in below-median sectors is a function of raised expectations that went unmet. But it certainly seems true that both high-level city leaders and street-level staff overestimated their ability to shape community outcomes in the places where they has historically less presence.

A final possibility is that the intervention would have been more effective if policing increased in concert with civilian services. This would be consistent with a literature on counter-insurgency, which argues that a combination of military action followed by state service provision increases state legitimacy and civilian collaboration against the insurgents.[29] We cannot speak to policing alone. But in a related study, we look at quasi-experimental variation that led some city blocks to receive substantially more police patrols and civilian dispute resolution attention over 30 years. Those blocks have slightly lower levels of crime. Unexpectedly, however, they also developed *stronger* combos who governed *more*. The evidence suggests that, to protect their drug profits, gangs tried to reduce street disorder and thus lower the chances that police are called to the street (**?**). We do not see evidence

[29] Albertus and Kaplan (2013); Berman et al. (2011, 2013); Berman and Matanock (2015)

of a combo strategic response over 20 months, but we cannot rule out that a more sustained intervention would elicit a response.

Altogether, these results imply that bringing about order and building state capacity and legitimacy are complicated, and there may be no simple policy solutions. If nothing else, this experiment illustrates the importance of increased experimentation with new strategies alongside rigorous evaluation.

References

Abril, V., E. Norza, S. Perez-Vincent, S. Tobon, and M. Weintraub (2023). Building trust in state actors: A multi-site experiment with the Colombian National Police. Technical report, Working Paper.

Abt, T. (2019). *Bleeding Out: The Devastating Consequences of Urban Violence–and a Bold New Plan for Peace in the Streets*. Basic Books.

Acemoglu, D., A. Cheema, A. I. Khwaja, and J. A. Robinson (2020). Trust in state and nonstate actors: Evidence from dispute resolution in pakistan. *Journal of Political Economy 128*(8), 3090–3147.

Agan, A., J. L. Doleac, and A. Harvey (2023, 01). Misdemeanor Prosecution*. *The Quarterly Journal of Economics*. qjad005.

Aizer, A. and J. Doyle, Joseph J. (2015, 02). Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges *. *The Quarterly Journal of Economics 130*(2), 759–803.

Albertus, M. and O. Kaplan (2013). Land Reform as a Counterinsurgency Policy: Evidence From Colombia. *Journal of Conflict Resolution 57*(2), 198–231.

Ang, D. (2021). The Effects of Police Violence on Inner-City Students. *Quarterly Journal of Economics 136*(1), 115–168.

Arias, E. D. (2006). The Dynamics of Criminal Governance: Networks and Social Order in Rio de Janeiro. *Journal of Latin American Studies 38*(2), 293–325.

Arias, E. D. (2017). *Criminal Enterprises and Governance in Latin America and the Caribbean*. New York: Cambridge University Press.

Bai, Y. (2022). Optimality of matched-pair designs in randomized controlled trials. *American Economic Review 112*(12), 3911–40.

Berman, E., J. H. Felter, J. N. Shapiro, and E. Troland (2013). Modest, Secure, and Informed: Successful Development in Conflict Zones. *American Economic Review 103*(3), 512–17.

Berman, E. and D. D. Laitin (2008). Religion, Terrorism and Public Goods: Testing the Club Model. *Journal of Public Economics 92*(10-11), 1942–1967.

Berman, E. and A. M. Matanock (2015). The Empiricists' Insurgency. *Annual Review of Political Science 18*, 443–464.

Berman, E., J. N. Shapiro, and J. H. Felter (2011). Can Hearts and Minds be Bought? The Economics of Counterinsurgency in Iraq. *Journal of Political Economy 119*(4), 766–819.

Bhatt, M. P., S. B. Heller, M. Kapustin, M. Bertrand, and C. Blattman (2023). Predicting and preventing gun violence: An experimental evaluation of readi chicago. *Quarterly*

Journal of Economics.

Blattman, C., S. Chaskel, J. C. Jamison, and M. Sheridan (forthcoming). Cognitive behavior therapy reduces crime and violence over 10 years: Experimental evidence. *American Economic Review: Insights*.

Blattman, C., G. Duncan, B. Lessing, and S. Tobón (2023a). Gang Rule: Understanding and Countering Criminal Governance. *Working Paper*.

Blattman, C., G. Duncan, B. Lessing, and S. Tobón (2023b). Pax Criminalis. *Working Paper*.

Blattman, C., D. Green, D. Ortega, and S. Tobón (2021). Place-Based Interventions at Scale: The Direct and Spillover Effects of Policing and City Services on Crime. *Journal of the European Economic Association 19*(4), 2022–2051.

Blattman, C., A. C. Hartman, and R. A. Blair (2014). How to Promote Order and Property Rights Under Weak Rule of Law? An Experiment in Changing Dispute Resolution Behavior Through Community Education. *American Political Science Review*, 100–120.

Blattman, C., J. C. Jamison, and M. Sheridan (2017). Reducing Crime and Violence: Experimental Evidence on Adult Noncognitive Investments in Liberia. *American Economic Review 107*(4).

Bokat-Lindell, S. (2023, February). Opinion: How to stop police violence. *The New York Times*.

Butts, J. A., K. T. Wolff, E. Misshula, and S. A. Delgado (2015). Effectiveness of the cure violence model in new york city.

Cammett, M. and L. M. MacLean (2014). *The Politics of Non-State Social Welfare*. Cornell University Press.

Carr, J. B. and A. Packham (2019, 05). SNAP Benefits and Crime: Evidence from Changing Disbursement Schedules. *The Review of Economics and Statistics 101*(2), 310–325.

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

Chalfin, A. and J. McCrary (2018). Are us cities underpoliced? theory and evidence. *Review of Economics and Statistics 100*(1), 167–186.

Cruz, J. M. and A. Durán-Martínez (2016). Hiding violence to deal with the state: Criminal pacts in El Salvador and Medellin. *Journal of Peace Research 53*(2), 197–210.

Davis, J. M. and S. B. Heller (2020, 10). Rethinking the Benefits of Youth Employment Programs: The Heterogeneous Effects of Summer Jobs. *The Review of Economics and Statistics 102*(4), 664–677.

Dee, T. S. and J. Pyne (2022). A community response approach to mental health and substance abuse crises reduced crime. *Science Advances 8*(23), eabm2106.

DiPasquale, D. and E. L. Glaeser (1998). The Los Angeles Riot and the Economics of Urban Unrest. *Journal of Urban Economics 43*(1), 52–78.

Farrell, A. D., D. Henry, C. Bradshaw, and T. Reischl (2016). Designs for evaluating the community-level impact of comprehensive prevention programs: Examples from the cdc centers of excellence in youth violence prevention. *The Journal of Primary Prevention 37*(2), 165–188.

Gonzalez, R. and S. Komisarow (2020). Community monitoring and crime: Evidence from chicago's safe passage program. *Journal of Public Economics 191*, 104250.

Heller, S. B., A. K. Shah, J. Guryan, J. Ludwig, S. Mullainathan, and H. A. Pollack (2016, 10). Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago*. *The Quarterly Journal of Economics 132*(1), 1–54.

Henn, S. J. (2021). Complements or Substitutes? How Institutional Arrangements Bind Chiefs and the State in Africa.

Hjalmarsson, R., H. Holmlund, and M. J. Lindquist (2015). The effect of education on criminal convictions and incarceration: Causal evidence from micro-data. *The Economic Journal 125*(587), 1290–1326.

Irwin, A. and B. Pearl (2020). The community responder model: How cities can send the right responder to every 911 call. *Center for American Progress*.

Lessing, B. (2020). Conceptualizing Criminal Governance. *Perspectives on Politics*, 1–20.

Lessing, B., D. Block, and E. Stecher (2019). Criminal Governance in Latin America: an Empirical Approximation. *Working Paper*.

Melnikov, N., C. Schmidt-Padilla, and M. M. Sviatschi (2020). Gangs, Labor Mobility, and Development: The Role of Extortion in El Salvador.

Moncada, J. J., C. Lopera, N. Maya, C. Cadavid, and L. Zuluaga (2018). *La Extorsión en Medellín Como Fenómeno del Orden Social, Poder Político y Control Territorial*. Medellín: Alcaldía de Medellín.

North, D. C., J. J. Wallis, and B. R. Weingast (2009). *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge University Press.

Owens, E. (2019). Economic approach to de-policing. *Criminology & Pub. Pol'y 18*, 77.

Owens, E. and B. Ba (2021, November). The economics of policing and public safety. *Journal of Economic Perspectives 35*(4), 3–28.

Roman, C. G., H. J. Klein, and K. T. Wolff (2018). Quasi-experimental designs for community-level public health violence reduction interventions: a case study in the challenges of selecting the counterfactual. *Journal of Experimental Criminology 14*(2), 155–185.

Sánchez De La Sierra, R. (2020). On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo. *Journal of Political Economy 128*(1), 000–000.

Seo, C., B. Kim, and N. E. Kruis (2021). A meta-analysis of police response models for handling people with mental illnesses: Cross-country evidence on the effectiveness. *International criminal justice review 31*(2), 182–202.

Tilly, C. (1985). War Making and State Making as Organized Crime. *Violence: A reader*, 35–60.

Uribe, A., B. Lessing, D. Block, E. Stecher, and N. Schouela (2022). Criminal Governance in Latin America: an Initial Assessment of its Extent and Correlates.

Van der Windt, P., M. Humphreys, L. Medina, J. F. Timmons, and M. Voors (2019). Citizen Attitudes Toward Traditional and State Authorities: Substitutes or Complements? *Comparative Political Studies 52*(12), 1810–1840.

Weber, M. (1946). *Politics as a Vocation*, pp. 77–128. Oxford University Press.

Weigel, J. L. (2020). The Participation Dividend of Taxation: How Citizens in Congo Engage More with the State When it Tries to Tax Them. *The Quarterly Journal of Economics 135*(4), 1849–1903.

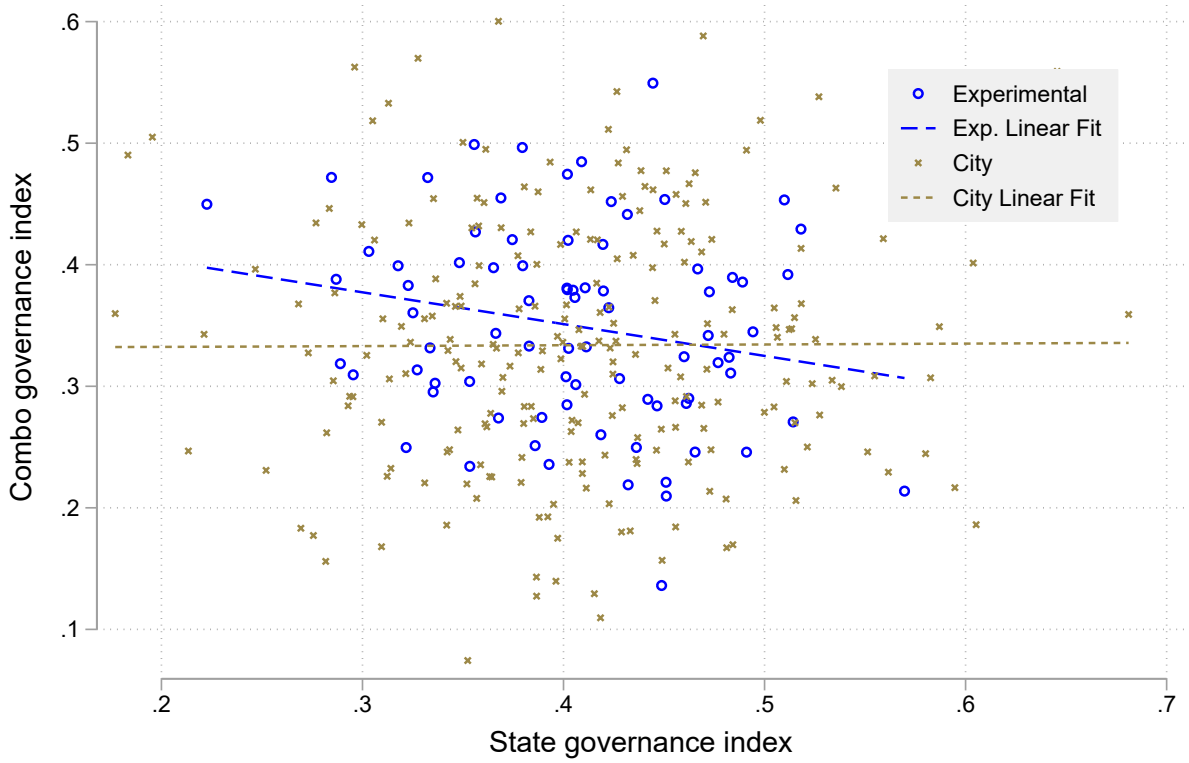
Weisburd, D., E. R. Groff, and S.-M. Yang (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford University Press.

Zahniser, D. (2023, March). Lapd officers want to stop responding to nonviolent calls. *Los Angeles Times*.

Appendix

A Supplemental tables

Figure A.1: Comparison of the experimental and city (representative) sample of blocks in 2019: State and combo governance levels



Notes: The figure plots average 2019 state and combo governance levels in each city barrio as well as the 40 experimental control sectors. We omit treated sectors because the 2019 survey is post-treatment. The dashed lines are lines of best fit for the two samples. The experimental sectors are widely distributed, much like the city barrios, though there are slightly more high combo/low state governance areas in the experimental sample.

Table A.1: Estimating treatment spillovers onto blocks within a 250 meter radius

	Treatment Estimate (1)	P-value (2)	0m-250m Spillover Estimate (3)	P-value (4)
Relative State Governance Index	-0.031	0.121	-0.067	0.919
State Governance Index (0-1)	-0.015	0.232	-0.030	0.946
Combo Governance Index (0-1)	0.014	0.378	0.031	0.869
Relative State Legitimacy Index	0.006	0.889	-0.051	0.706
State Legitimacy Index (0-1)	0.011	0.341	-0.011	0.583
Combo Legitimacy Index (0-1)	0.006	0.776	0.036	0.847

Notes: Our sample includes 6977 survey respondents, including 2,379 in the experimental sectors and 4,598 on blocks from the representative city survey. The tale reports treatment estimates along with an indicator for blocks in the experimental sectors and an indicator for blocks within 250 meters of a treated sector. As Blattman et al. (2021) note, spillovers in a dense network of blocks can lead to fuzzy clustering, where clusters do not conform to defined areas. Hence we use randomization inference to estimate exact p-values under the sharp null of no treatment effect for any unit, correcting estimates for fuzzy clustering. To address systematic exposure to spillovers due to the geographic distribution, we weight each observation by the inverse probability of each treatment category: treated, <250 meters, and >250 meters.

Table A.2: Count of officially-logged liaison activities per sector

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
N. activities in sector area	0.23	35.590***	28.285***	42.235***	-13.950**
	(0.53)	(3.029)	(4.159)	(4.152)	(5.851)
N. activities in 125m sector buffer area	1.88	70.412***	65.718***	75.499***	-9.780*
	(4.75)	(2.564)	(3.576)	(3.570)	(5.031)

Notes: To illustrate treatment compliance, and spillovers of liaison activities into control sectors, this table reports summary statistics and treatment effects for all activities officially logged by liaisons. We examine the count of all activities within the experimental sector itself, as well as within a 125 meter buffer. Note that the activities of other municipal employees are not logged, and so this is an incomplete measure of city staff activities. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4. Note that only households and not businesses were surveyed on legitimacy.

Table A.3: Randomization balance within prespecified subgroups

Covariate	High relative state gov.				Low relative state gov.			
	Control mean	Treatment mean	Coeff	p-value	Control mean	Treatment mean	Coeff	p-value
Baseline indices used for matching , main control vector								
Additive index of combo presence and governance	-0.57	-0.63	-0.06	0.764	0.59	0.61	0.03	0.932
Baseline - Combo Governance Index (Relative to State)	-0.69	-0.75	-0.06	0.718	0.70	0.72	0.02	0.950
Standardized index of perceived insecurity and drugs	0.07	-0.07	-0.14	0.654	0.05	-0.06	-0.11	0.738
Index of crime	0.01	-0.17	-0.18	0.535	0.17	-0.06	-0.23	0.493
Index of distance from public goods and services	-0.21	0.17	0.39	0.268	-0.06	0.10	0.16	0.563
Other baseline covariates								
Respondent age between 18 and 25	0.18	0.18	-0.00	0.852	0.20	0.21	0.00	0.891
Respondent age between 26 and 40	0.26	0.31	0.05	0.103	0.33	0.30	-0.02	0.457
Respondent age between 41 and 64	0.42	0.36	-0.05*	0.083	0.35	0.38	0.03	0.452
Respondent is business owner	0.20	0.20	0.00	0.122	0.20	0.19	-0.00	0.225
Multidimentional Poverty Index (2018)	11.75	12.93	1.19	0.645	16.98	21.73	4.75	0.171
Block Longitude	-75.59	-75.59	-0.00	0.934	-75.57	-75.58	-0.01	0.225
Block present in 1970	0.60	0.51	-0.09	0.479	0.40	0.38	-0.03	0.846
Median age (2005)	29.19	27.61	-1.58	0.315	25.18	24.96	-0.22	0.869
Total women (2005)	135.53	142.93	7.40	0.696	132.15	141.12	8.96	0.619
Total non-mestizo polulation (1993)	0.58	0.20	-0.38*	0.095	0.49	1.04	0.55**	0.032
Median age (1993)	25.80	26.42	0.62	0.702	22.49	22.96	0.47	0.692
Share of women (1993)	0.52	0.51	-0.00	0.855	0.54	0.53	-0.01	0.603
Distance to the respective razon headquarters (100 meters)	17.86	23.49	5.63	0.456	16.69	14.62	-2.07	0.528

Notes: This table reports treatment and control group means and a test of balance for all covariates in Table 2, but does so within the two prespecified subgroups: above and below median baseline relative state governance.

Table A.4: Heterogeneity analysis by quartiles of relative baseline state governance

	(A) Subgroups by: *Relative* baseline governance		(B) Subgroups by: *Absolute* baseline governance	
	Dependent Variable: Relative governance	Dependent Variable: Relative Legitimacy	Dependent Variable: Relative governance	Dependent Variable: Relative Legitimacy
	(1)	(2)	(3)	(4)
Program impacts:				
Q1 (0^{th} - 25^{th} quartile baseline rel. gov)	-0.025	-0.072*	-0.058**	-0.008
	(0.031)	(0.038)	(0.024)	(0.037)
Q2 (25^{th} - 50^{th} quartile baseline rel. gov)	-0.039**	0.020	-0.063***	-0.044*
	(0.018)	(0.037)	(0.022)	(0.026)
Q3 (50^{th} - 75^{th} quartile baseline rel. gov)	0.026	0.051	0.087***	0.127***
	(0.026)	(0.033)	(0.020)	(0.046)
Q4 (75^{th} - 100^{th} quartile baseline rel. gov)	-0.067*	0.052	-0.047	0.015
	(0.034)	(0.040)	(0.037)	(0.041)
Differences relative to Q1:				
Q2	-0.014	0.093*	-0.006	-0.036
	(0.035)	(0.049)	(0.033)	(0.046)
Q3	0.051	0.123**	0.145***	0.134**
	(0.040)	(0.048)	(0.030)	(0.056)
Q4	-0.042	0.124**	0.011	0.023
	(0.046)	(0.058)	(0.046)	(0.058)

Notes: This table replicates the results of Table **??** but partitioning the sample in 4 subgroups (quartiles) as opposed to 2. Columns 1 and 2 replicate the heterogeneity analysis in Table 1 by quartiles of baseline relative state governance. Here we report program effects on each each subgroup in the first 4 rows, while the last 3 report differences with respect to the lowest governance group. Unfortunately this leaves just 20 sectors per quartile subgroup, making this analysis somewhat underpowered. Both households and businesses were surveyed on governance levels, but only households were surveyed on legitimacy (and hence there are fewer observations).

Table A.5: Program impacts on police and mayor's office legitimacy components: Average treatment effects and heterogeneity by baseline legitimacy

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Police legitimacy index	0.57	0.006	0.032***	-0.022**	0.054***
	(0.23)	(0.008)	(0.011)	(0.010)	(0.014)
How much do you trust the police	0.56	0.002	0.034**	-0.032**	0.066***
	(0.34)	(0.011)	(0.013)	(0.015)	(0.019)
How fair is the police	0.57	-0.006	0.007	-0.019	0.026
	(0.30)	(0.009)	(0.012)	(0.015)	(0.019)
How do you rate the police	0.59	0.007	0.032***	-0.019*	0.051***
	(0.24)	(0.009)	(0.012)	(0.011)	(0.016)
How would your neighbors rate the police	0.59	0.016*	0.037***	-0.006	0.043**
	(0.26)	(0.009)	(0.012)	(0.012)	(0.016)
How much do your neighbors trust the police	0.57	0.013	0.057***	-0.034***	0.091***
	(0.32)	(0.011)	(0.014)	(0.013)	(0.019)
Mayor legitimacy index	0.57	0.012	0.026**	-0.003	0.028
	(0.23)	(0.008)	(0.012)	(0.012)	(0.018)
How much do you trust the mayoral staff	0.57	0.004	0.018	-0.011	0.029
	(0.33)	(0.011)	(0.015)	(0.017)	(0.023)
How fair is the mayoral staff	0.53	0.006	0.022	-0.010	0.032
	(0.31)	(0.011)	(0.016)	(0.017)	(0.024)
How do you rate the mayoral staff	0.61	0.003	0.017	-0.012	0.030
	(0.25)	(0.010)	(0.013)	(0.014)	(0.019)
How would your neighbors rate the mayoral staff	0.59	0.019*	0.019	0.018	0.001
	(0.27)	(0.010)	(0.014)	(0.015)	(0.020)
How much do your neighbors trust the mayoral staff	0.55	0.033***	0.047**	0.018	0.029
	(0.32)	(0.011)	(0.019)	(0.012)	(0.023)
Combo legitimacy index	0.44	-0.006	-0.022	0.011	-0.033
	(0.28)	(0.014)	(0.017)	(0.023)	(0.029)
How much do you trust the combo	0.36	0.003	-0.012	0.018	-0.030
	(0.36)	(0.017)	(0.018)	(0.029)	(0.033)
How fair is the combo	0.41	-0.001	-0.033	0.034	-0.067**
	(0.34)	(0.016)	(0.023)	(0.022)	(0.033)
How do you rate the combo	0.50	0.001	-0.015	0.018	-0.033
	(0.27)	(0.013)	(0.015)	(0.021)	(0.027)
How much do your neighbors trust the combo	0.51	-0.010	-0.031*	0.011	-0.042
	(0.30)	(0.014)	(0.017)	(0.022)	(0.028)
How would your neighbors rate the combo	0.48	-0.011	-0.027	0.007	-0.034
	(0.36)	(0.017)	(0.025)	(0.025)	(0.035)

Notes: The table reports summary statistics and treatment effects for 5 survey-based measures of legitimacy per actor, plus a summary index for the 5 questions. Each row is a different dependent variable. Each row is a different dependent variable. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4. Note that only households and not businesses were surveyed on legitimacy.

Table A.6: Program impacts on relative state governance components: Average treatment effects and heterogeneity by baseline governance quality

		ATE	Het. by baseline rel. gov.		
			Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Relative state governance index (less police related)	0.09	-0.020	-0.013	-0.026*	0.013
	(0.31)	(0.012)	(0.016)	(0.016)	(0.021)
HH: Someone is making noise	0.26	-0.025	-0.021	-0.028	0.006
	(0.42)	(0.020)	(0.029)	(0.028)	(0.040)
HH: Home improvements affect neighbors	0.14	-0.007	0.005	-0.017	0.022
	(0.44)	(0.023)	(0.033)	(0.032)	(0.046)
HH: There is domestic violence	0.15	-0.014	0.029	-0.052	0.080*
	(0.45)	(0.023)	(0.033)	(0.032)	(0.046)
HH: Two drunks fight on the street	0.13	-0.018	0.004	-0.037	0.041
	(0.45)	(0.023)	(0.033)	(0.032)	(0.046)
Biz: Someone disturbs a business	0.16	-0.066	-0.030	-0.103	0.072
	(0.50)	(0.050)	(0.070)	(0.070)	(0.099)
HH: People smoking marijuana near children	0.03	0.004	0.029	-0.019	0.048
	(0.40)	(0.020)	(0.028)	(0.028)	(0.039)
HH: Kids fight on the street	-0.03	-0.024	-0.010	-0.037	0.027
	(0.41)	(0.021)	(0.031)	(0.029)	(0.042)
Biz: Someone does not want to pay a debt	-0.05	-0.016	0.013	-0.036	0.049
	(0.33)	(0.037)	(0.052)	(0.051)	(0.073)
HH: Someone refuses to pay a big debt	-0.20	-0.033	-0.007	-0.055	0.048
	(0.45)	(0.024)	(0.035)	(0.034)	(0.048)
Relative state governance index (more police related)	0.02	-0.029*	-0.011	-0.049**	0.038
	(0.39)	(0.018)	(0.026)	(0.023)	(0.034)
Biz: You have to react to a robbery	0.12	-0.083*	-0.121*	-0.045	-0.076
	(0.48)	(0.050)	(0.069)	(0.071)	(0.099)
Biz: It is necessary to prevent a theft	0.08	-0.078	-0.071	-0.086	0.016
	(0.52)	(0.049)	(0.069)	(0.071)	(0.098)
Biz: Businesses in this sector are robbed	0.07	-0.065	-0.071	-0.060	-0.011
	(0.50)	(0.048)	(0.070)	(0.067)	(0.097)
HH: A car or motorbike is stolen	-0.01	0.010	0.049	-0.029	0.077*
	(0.45)	(0.023)	(0.033)	(0.033)	(0.046)
HH: Someone is threatening someone else	-0.01	-0.023	0.020	-0.060*	0.079*
	(0.45)	(0.024)	(0.034)	(0.033)	(0.047)
HH: You have to react to a robbery	-0.02	-0.023	0.018	-0.061*	0.079*
	(0.47)	(0.023)	(0.033)	(0.032)	(0.046)
HH: Someone is mugged on the street	-0.05	0.006	0.037	-0.022	0.058
	(0.42)	(0.022)	(0.031)	(0.031)	(0.044)
HH: It is necessary to prevent a theft	-0.04	-0.018	0.017	-0.053	0.069
	(0.48)	(0.023)	(0.033)	(0.033)	(0.046)

*Notes:*The table reports summary statistics and treatment effects for the 17 components of the governance index in Table 4. We create sub-indexes for what our qualitative work suggests are more and less police-related forms of governance. Each row is a different dependent variable. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4. Note that both households and businesses were surveyed on governance.

Table A.7: Impacts of treatment on survey measures of police, mayoral, and combo efficacy

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Police efficacy index	0.55	-0.009	-0.007	-0.011	0.005
	(0.21)	(0.009)	(0.013)	(0.010)	(0.016)
How easy is it to contact the police	0.54	-0.010	-0.005	-0.015	0.009
	(0.29)	(0.010)	(0.017)	(0.009)	(0.019)
Perceived value of the police	0.71	0.002	-0.005	0.011	-0.016
	(0.25)	(0.010)	(0.016)	(0.010)	(0.019)
How fast is the police	0.42	-0.018	-0.006	-0.030	0.024
	(0.34)	(0.013)	(0.013)	(0.021)	(0.024)
Mayoral staff efficacy index	0.45	-0.010	-0.004	-0.016	0.012
	(0.20)	(0.008)	(0.013)	(0.010)	(0.016)
How easy is it to contact mayoral staff	0.35	0.003	0.007	-0.002	0.009
	(0.30)	(0.011)	(0.017)	(0.013)	(0.021)
Perceived value of the mayoral staff	0.66	-0.013	0.000	-0.028***	0.029*
	(0.26)	(0.009)	(0.014)	(0.008)	(0.016)
How fast is the mayoral staff	0.34	-0.010	0.001	-0.022	0.023
	(0.31)	(0.011)	(0.016)	(0.015)	(0.021)
Combo efficacy index	0.55	-0.004	-0.007	0.000	-0.007
	(0.24)	(0.012)	(0.017)	(0.016)	(0.024)
How easy is it to contact the combo	0.59	0.011	0.019	0.002	0.017
	(0.31)	(0.016)	(0.018)	(0.025)	(0.030)
Perceived value of the combo	0.52	-0.014	-0.034**	0.008	-0.042
	(0.32)	(0.014)	(0.016)	(0.023)	(0.029)
How fast is the combo	0.56	0.002	-0.004	0.009	-0.013
	(0.36)	(0.016)	(0.025)	(0.019)	(0.031)

Notes: This table reports summary statistics and treatment effects on 3 survey-based measures of efficacy per actor, plus a summary index for the three questions. Each row is a different dependent variable. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4. Note that only households and not businesses were surveyed on efficacy.

Table A.8: Impacts of treatment on survey measures of police, mayoral, and combo efficacy

			Het. by baseline rel. gov.		
		ATE	Above median	Below median	Diff.
Dependent variable	Control Mean (SD)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
	(1)	(2)	(3)	(4)	(5)
Police efficacy index	0.55	-0.009	-0.007	-0.011	0.005
	(0.21)	(0.009)	(0.013)	(0.010)	(0.016)
How easy is it to contact the police	0.54	-0.010	-0.005	-0.015	0.009
	(0.29)	(0.010)	(0.017)	(0.009)	(0.019)
Perceived value of the police	0.71	0.002	-0.005	0.011	-0.016
	(0.25)	(0.010)	(0.016)	(0.010)	(0.019)
How fast is the police	0.42	-0.018	-0.006	-0.030	0.024
	(0.34)	(0.013)	(0.013)	(0.021)	(0.024)
Mayoral staff efficacy index	0.45	-0.010	-0.004	-0.016	0.012
	(0.20)	(0.008)	(0.013)	(0.010)	(0.016)
How easy is it to contact mayoral staff	0.35	0.003	0.007	-0.002	0.009
	(0.30)	(0.011)	(0.017)	(0.013)	(0.021)
Perceived value of the mayoral staff	0.66	-0.013	0.000	-0.028***	0.029*
	(0.26)	(0.009)	(0.014)	(0.008)	(0.016)
How fast is the mayoral staff	0.34	-0.010	0.001	-0.022	0.023
	(0.31)	(0.011)	(0.016)	(0.015)	(0.021)
Combo efficacy index	0.55	-0.004	-0.007	0.000	-0.007
	(0.24)	(0.012)	(0.017)	(0.016)	(0.024)
How easy is it to contact the combo	0.59	0.011	0.019	0.002	0.017
	(0.31)	(0.016)	(0.018)	(0.025)	(0.030)
Perceived value of the combo	0.52	-0.014	-0.034**	0.008	-0.042
	(0.32)	(0.014)	(0.016)	(0.023)	(0.029)
How fast is the combo	0.56	0.002	-0.004	0.009	-0.013
	(0.36)	(0.016)	(0.025)	(0.019)	(0.031)

Notes: This table reports summary statistics and treatment effects on 3 survey-based measures of efficacy per actor, plus a summary index for the three questions. Each row is a different dependent variable. Average treatment effects and treatment heterogeneity are calculated using the same approach as in Table 4. Note that only households and not businesses were surveyed on efficacy.

B Conceptual framework

B.1 Cournot competition in local governance

To understand why the intervention could crowd combos out of local governance, we can look at the intervention through the lens of imperfect competition for governance services. Any model of imperfect competition should produce similar comparative statics, but we illustrate with Cournot competition, where each side chooses a fixed quantity of protection services to provide and let prices clear the market.[30]

Of course, states are not necessarily profit-maximizing and have broader objectives. We model this in simple form below. **?** consider a fuller range of models and additional assumptions, focusing on the strategic response of gangs. But that paper also shows how the results here would be similar in other forms of imperfect competition, including a model of stationary bandits competing to provide public goods.

Setup In each neighborhood, a state s and a gang g compete to sell protection in quantities q_g and q_s. Each organization i chooses q_i to maximize their respective pay-off, and each has constant marginal cost c_i. (Here i can either be the state or the gang, and in what follows, j represents a general form of notation for the competing organization.) Products are differentiated, and the price of each one is given by the linear inverse demand function $p_i = a_i - \beta q_i - \gamma q_j$. Here, $\gamma \in (0, 1]$ since the services offered by both organizations are substitutes, and $\beta > 0$ for downward-sloping demand. The pay-off for each organization is $V_i = p_i q_i - c_i q_i$. For simplicity, we assume an interior solution.

Nash Equilibria The best response function for each organization are derived as follows:

$$\begin{aligned} \max_{q_i} V_i &= (a_i - \beta q_i - \gamma q_j) q_i - c_i q_i \\ \frac{\partial V_i}{\partial q_i} &= a_i - 2\beta q_i - \gamma q_j - c_i = 0 \\ q_i^* &= \frac{a_i - c_i}{2\beta} - \frac{\gamma}{2\beta} q_j \end{aligned}$$

Replacing values we obtain (for each organization):

$$q_i^* = \frac{2\beta(a_i - c_i) - \gamma(a_j - c_j)}{(4\beta^2 - \gamma^2)}$$

[30] Note that Cournot fits some of our stylized facts well—especially that governing requires investments and advanced commitments, and that it is hard to adjust output capacity quickly.

Comparative statics We are interested in how the quantity of services supplied by the gang behave in response to any increase in state governance: $\frac{\partial q_i^*}{\partial q_j}$. To obtain this comparative static, we begin by defining:

$$G(q_i, q_j) \equiv \frac{\partial V_i}{\partial q_i} = a_i - 2\beta q_i - \gamma q_j - c_i$$

which is a continuously differentiable function from $\mathbb{R}^2 \to \mathbb{R}$. At the optimum, we know:

$$G(q_i^*, q_j^*) = a_i - 2\beta_i q_i^* - \gamma q_j^* - c_i = 0.$$

Since $-2\beta \neq 0$, we can use the implicit function theorem to obtain our main comparative static:

$$\frac{\partial q_i^*}{\partial q_j} = -\frac{\partial G(q_i, q_j)/\partial q_j}{\partial G(q_i, q_j)/\partial q_i} = -\frac{\gamma}{2\beta}$$

Since the two services are not complements, this comparative static implies that increases in one duopolist's supply of protection will reduce the other's.

B.1.1 Cournot competition with benefits to governing

In the simple model above, increases in the quantity supplied by the state would mainly come from reductions in the state's marginal cost of providing these goods. One way to conceive the experimental intervention is an exogenous investment by the state in lowering the marginal cost of providing governance services. Another way to view the intervention, however, is the result of an exogenous increase in the value the state places on being the market leader in that neighborhood, or even a monopolist. To illustrate this, we introduce a new term to the utility function.

Setup As above, but now the payoff for each organization is $V_i = p_i q_i - c_i q_i + \rho(q_i, q_j)\pi_i$, where $\rho(q_i, q_j)\pi_i$ represents each player's returns to loyalty, legitimacy, and control of the neighborhood.
Set up this way, π_i is the return to full control of the block. For example, π_s includes electoral rewards, achievement of policy aims, or preferences for dominance and citizen loyalty.

Meanwhile, $\rho(\cdot)$ scales each organization's ability to capture, retain, or enjoy these benefits. We can think of it as the share of π_i each player enjoys, one that is increasing in own governance and decreasing in the other's, such that: $\frac{\partial \rho(q_i, q_j)}{\partial q_i} > 0 > \frac{\partial \rho(q_i, q_j)}{\partial q_j}$. Importantly, however, we remain agnostic here about whether $\rho(\cdot)$ exhibits increasing or decreasing re-

turns to own and other's governance provision.

Nash Equilibria For simplicity, we assume an interior solution. We can derive the best response function for each organization:

$$\begin{aligned}
\max_{q_i} V_i &= (a_i - \beta q_i - \gamma q_j)q_i - c_i q_i + \rho(q_i, q_j)\pi_i \\
\frac{\partial V_i}{\partial q_i} &= a_i - 2\beta q_i - \gamma q_j - c_i + \frac{\partial \rho(q_i, q_j)}{\partial q_i}\pi_i = 0 \\
q_i^* &= \frac{a_i - c_i + \frac{\partial \rho(q_i, q_j)}{\partial q_i}\pi_i}{2\beta} - \frac{\gamma}{2\beta} q_j
\end{aligned}$$

We obtain an identical best response function for the other organization analogously, and replacing values we obtain:

$$q_i^* = \frac{2\beta(a_i - c_i) - \gamma(a_j - c_j) + \left(2\beta \frac{\partial \rho(q_i, q_j)}{\partial q_i}\pi_i - \gamma \frac{\partial \rho(q_i, q_j)}{\partial q_j}\pi_j\right)}{(4\beta^2 - \gamma^2)}$$

with an identical function for q_j^*.

The state's equilibrium level of governance services supplied is increasing in the value they place on neighborhood control, π_i, and their expected returns to investment in citizen loyalty and neighborhood control, $\frac{\partial \rho(q_i, q_j)}{\partial q_i}$.