

# Preaching to the Choir: A Problem of Participatory Interventions

## Online Appendix

Rebecca Hanson\*      Dorothy Kronick†      Tara Slough‡

December 7, 2022

## Contents

<b>A1 Design</b>	<b>A-2</b>
<b>A2 Survey</b>	<b>A-2</b>
<b>A3 Supplemental Information on Research Ethics</b>	<b>A-4</b>
<b>A4 Supplemental Information on Reach and Spillovers</b>	<b>A-6</b>
<b>A5 ITT Analysis</b>	<b>A-6</b>
<b>A6 Supporting Information on Selection</b>	<b>A-8</b>
A6.1 Disaggregation of compliance measures . . . . .	A-8
A6.2 Auxiliary results on positive selection into participation . . . . .	A-8
A6.3 Positive selection in other Metaketa Experiments . . . . .	A-12
A6.4 Non-survey evidence of differential selection into community policing . . . . .	A-12
<b>A7 Framework for selection and updating</b>	<b>A-13</b>
<b>A8 Three challenges of positive selection</b>	<b>A-15</b>
A8.1 Censoring . . . . .	A-15
A8.2 Mean reversion . . . . .	A-16
A8.3 Preaching to the choir (updating) . . . . .	A-17
<b>A9 Relationship to Gonzalez and Mayka (2022)</b>	<b>A-17</b>
<b>A10 Anonymized Pre-Analysis Plan</b>	<b>A-21</b>

---

\*Assistant Professor, University of Florida, r.hanson@ufl.edu

†Assistant Professor, University of Pennsylvania, kronick@upenn.edu

‡Assistant Professor, New York University, tara.slough@nyu.edu

## A1 Design

We use a  $2 \times 2$  factorial design that is block- and cluster-randomized at the level of police beat. There are 347 beats in the sample, of which 87 were assigned to pure control, 87 were assigned to only community-police meetings, 87 were assigned to only informational flyers, and 86 were assigned to both meetings and flyers. To evaluate balance across treatment arms, Table A1 reports estimates from the following regression:

$$X_i = \beta_0 + \beta_1 \text{Meetings only}_i + \beta_2 \text{Flyers only}_i + \beta_3 \text{Meetings and Flyers}_i + \epsilon_i$$

We conduct an  $F$ -test of the null hypothesis that  $\beta_1 = \beta_2 = \beta_3 = 0$ . The associated  $p$ -values are reported in Table A1.

	Meetings		Flyers		Meetings & Flyers		Control Mean ( $\beta_0$ )	Control Std. Dev.	$p$ -value for dif.
	$\beta_1$	Std. Error $\beta_1$	$\beta_2$	Std. Error $\beta_2$	$\beta_3$	Std. Error $\beta_3$			
Population	69.701	-(582.57)	-352.08	-(586.49)	-221.56	-(569.71)	6180.57	4716.71	0.87
Households	13.529	-(155.04)	-82.425	-(156.01)	-9.167	-(156.63)	1671.44	1214.78	0.92
People per household	0.06	-(0.055)	-0.034	-(0.050)	-0.032	-(0.050)	3.63	0.48	0.27
Share of rented homes	0.002	-(0.004)	-0.002	-(0.003)	0.002	-(0.004)	0.02	0.02	0.43
Avg. bedrooms per household	-0.027	-(0.052)	0.071	-(0.054)	-0.009	-(0.052)	2.25	0.4	0.28
Share of overcrowding households	0.009	-(0.009)	0	-(0.009)	0.008	-(0.010)	0.05	0.07	0.67
Household shares cooking with electricity	0.001	-(0.004)	-0.003	-(0.003)	0.001	-(0.005)	0.02	0.02	0.49
Household shares cooking with piped gas	0.004	-(0.012)	-0.01	-(0.010)	0.004	-(0.014)	0.07	0.07	0.49
Household shares with fridge or TV	0.004	-(0.005)	-0.001	-(0.004)	0.004	-(0.006)	0.03	0.03	0.67
Household shares with computer	0.004	-(0.008)	-0.006	-(0.006)	0.003	-(0.010)	0.04	0.04	0.51
Household shares with motorcycle	0.005	-(0.008)	-0.006	-(0.007)	0.005	-(0.010)	0.05	0.04	0.39
Household shares with landline	0.004	-(0.014)	-0.002	-(0.013)	0.016	-(0.012)	0.9	0.1	0.43
Household shares with pipeline gas	-0.007	-(0.017)	-0.003	-(0.019)	-0.018	-(0.017)	0.23	0.23	0.72
Employed per household	0.004	-(0.020)	0.005	-(0.018)	-0.013	-(0.020)	1.31	0.13	0.83
Unemployed per household	0.001	-(0.006)	0	-(0.005)	-0.005	-(0.005)	0.11	0.05	0.69
Retirees per household	0.001	-(0.001)	-0.001	-(0.001)	0.002	-(0.001)	0	0.01	0.17
Household shares with family living abroad	0	-(0.004)	0.003	-(0.004)	0.005	-(0.004)	0.04	0.03	0.43
Share of males per household	-0.009*	-(0.005)	-0.001	-(0.004)	-0.008	-(0.006)	0.47	0.03	0.13
Share of females per household	0.002	-(0.006)	0.006	-(0.005)	0	-(0.007)	0.52	0.03	0.55
Share of under-aged per household	0.002	-(0.007)	0.004	-(0.007)	-0.006	-(0.007)	0.24	0.07	0.53
Share of seniors per household	0.008	-(0.009)	-0.005	-(0.007)	0.014	-(0.010)	0.09	0.05	0.15
Household shares with male head	-0.007	-(0.007)	0.005	-(0.006)	-0.008	-(0.008)	0.63	0.04	0.16
Household shares with single-male parent	-0.008	-(0.005)	-0.002	-(0.005)	-0.010**	-(0.005)	0.13	0.05	0.13
Household shares with single-female parent	-0.003	-(0.006)	-0.002	-(0.006)	-0.007	-(0.006)	0.32	0.04	0.69
Household shares w/o children at home	-0.01	-(0.009)	0.001	-(0.008)	-0.007	-(0.009)	0.3	0.1	0.53
Household shares with university students	0.033	-(0.041)	-0.056	-(0.039)	-0.004	-(0.040)	2.8	0.32	0.12
Household head born in Colombia	-0.01	-(0.009)	0.003	-(0.007)	-0.017	-(0.010)	0.96	0.05	0.15
Household head born in Medellín	0.004	-(0.012)	-0.004	-(0.011)	-0.002	-(0.011)	0.39	0.07	0.94
Log of monthly rent	0.1	-(0.075)	0.049	-(0.072)	0.027	-(0.073)	12.41	0.66	0.58
Homicide rate	-5.493	-(12.530)	-16.499	-(11.523)	15.001	-(19.392)	34.81	78.04	0.25
Theft rate	-4.739	-(5.615)	-6.461	-(5.515)	-4.83	-(5.089)	18.38	59.13	0.71
Lagged homicide rate	3.053	-(6.980)	-1.413	-(6.095)	9.098	-(7.425)	29.16	41.24	0.49
Lagged theft rate	-43.301	-(40.045)	-46.651	-(40.743)	-36.574	-(39.085)	57.55	426.86	0.70
Second study treatment condition	0.034	-(0.024)	0.023	-(0.022)	0.039*	-(0.022)	0.1	0.31	0.34
Second study control condition	-0.023	-(0.023)	-0.034	-(0.023)	-0.019	-(0.020)	0.14	0.35	0.52

Table A1: Balance on census attributes and cross randomization with a different study. Robust standard errors in parentheses.

## A2 Survey

We conducted a two-wave panel survey. The baseline was fielded between January and April 2018; the end-line was fielded between September and December 2019. The surveys are representative of the *prioritized*

*neighborhoods* (see main text) within each of the 347 police beats in the experimental sample (see Figure 3). The baseline consisted of 5,205 respondents (15 respondents per beat), the endline of 3,644 respondents (10 to 13 respondents per beat, with 10 or 11 respondents in 340 of the beats). A total of 2,434 baseline respondents were successfully recontacted at endline, for a recontact rate of 47%. The remaining 1,210 respondents were randomly sampled at endline.

Tables A2 and A1 suggest no evidence that attrition in the panel survey was related to treatment assignment.

	(1)	(2)	(3)	(4)
Meetings	0.005 (0.017)	0.005 (0.015)	0.005 (0.025)	0.005 (0.023)
Flyers	0.010 (0.017)	0.010 (0.015)	0.010 (0.024)	0.010 (0.022)
Meetings×Flyers			-0.001 (0.035)	-0.000 (0.030)
Control group mean	0.53	0.53	0.52	0.52
Block FE		✓		✓
Num. obs.	5205	5205	5205	5205
N Clusters	347	347	347	347

\*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$

Table A2: Analysis of panel survey attrition in the baseline sample, as a function of both treatments. The control group means refer to the pooled “no meetings” conditions (flyers and pure control) in Columns 1–2, and to the pure control condition in Columns 3–4. Standard errors are clustered at the beat level, the unit of treatment assignment.

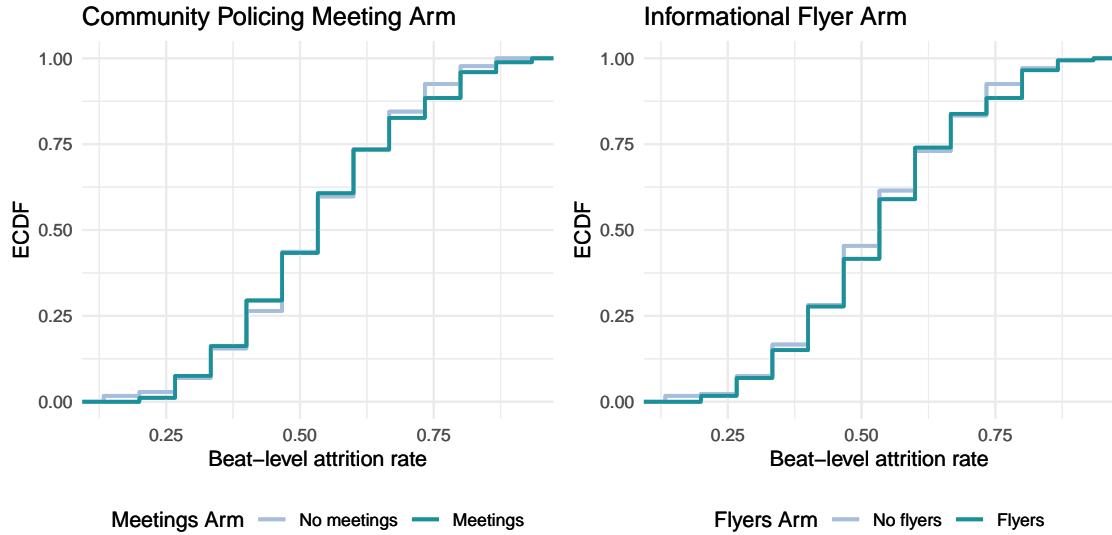


Figure A1: ECDFs of beat-level attrition in each treatment arm. Each subgroup (line) includes 173 or 174 beats, respectively.

### A3 Supplemental Information on Research Ethics

The manuscript describes some of the ethical considerations of an experiment entailing participation with the police. There are many additional ethical considerations relevant to our study, including those stemming from human subjects research with a multi-method research design involving:

- Pre-intervention focus groups with Medellín residents
- Interviews with police officials
- Baseline and endline surveys
- Experiment
- Participant observation in community-police meetings

Table A3 enumerates the twelve APSA Principles for Human Subjects Research. For each component of the broader research design (each column), we indicate whether an exception is necessary.

APSA Principles for Human Subjects Research	Focus groups	Interviews	Surveys	Experiment	Part. Obs.
1 Respect autonomy and consider the wellbeing of participants and other people affected by research					
2 Researcher responsibility to consider the ethics of research.					
3 APSA Principles as standards of conduct.					
4 Power [between researcher and subjects]					
5 Informed and voluntary consent.				✓	✓
6 Avoidance of deceptive or covert research.				✓	
7 Avoidance of harm					
8 Avoidance of trauma					
9 Confidentiality of participant identities		✓	✓		
10 Avoid compromising the integrity of broad political processes					
11 Awareness of relevant laws and regulations governing research and related activities					
12 Shared responsibility					

Table A3: The checkmarks represent exceptions justified below.

We discuss these considerations in the context of each piece of the research design. (We omit IRB protocol information to preserve the blind review process, though there was IRB approval covering all research staff for each component of the research design.)

1. **Focus groups:** The focus groups relied on verbal consent. There was no deception. Recruitment was conducted through contacts in local elected neighborhood councils (*Juntas de Acción Comunales*, or JACs). The focus groups did not collect any identifying information about participants.
2. **Interviews:** All interviews of police were conducted with verbal consent. There was no deception. Interviews with police officials in the development of the experiment were explicit about the design, treatments, outcomes, and intended learning. The identities of these officials were not circulated with intervention materials at any point. While some aspects of policing are controversial in Colombia,

topics related to community policing and beat-level service provision are not especially controversial. As such, sharing information with the research team did not entail substantial risks to the officers.

3. **Survey:** The survey included verbal consent in both waves. Respondents had the option to decline participation or refuse to answer any question. In order to maintain a panel survey, we collected identifying information at baseline. These identifying details were kept separate from the reported survey data and the investigators did not have access to participant identities.
4. **Experiment:** Participants (citizens and police) in community-policing meetings were not aware that meetings were randomly assigned. Note that all participants in these meetings self-selected into attendance, as we discuss in the manuscript. Police were assigned to these meetings by their supervisors as part of their jobs. Neither set of participants gave explicit consent for participation in research. In this sense, the research design was covert but not deceptive. The participant observation in these meetings (see below) was visible as research assistants took notes during meetings. The meetings included a sign-in sheet for officers and citizens. Signing in was optional. As such, some identifying information was collected. It is used in this research only to validate research assistants' assessments of the gender composition of participants.
5. **Participant observation:** Trained research assistants conducted participant observation in each of the community-police meetings. They identified themselves and their role in the meetings so this observation was known to all participants (police and citizens). They did not record names or identifying information about participants.

The subjects varied across components of the project. Because we relied heavily on field research staff, we include both subjects and field research staff in Table A4, which considers both compensation and the characteristics of each population.

Panel A: Subjects					
Activity	Subject	Payment	Diverse population	Marginalized population	Differential harm/benefit
Focus groups	Medellín residents	yes	yes	mixed	no
Interviews	MEVAL police	no	no	no	no
Surveys	Medellín residents	no	yes	mixed	no
Experiment	Citizens	snacks	yes	mixed	no
	Police	(through job)	no	no	no
Participant obs.	Citizens	snacks	yes	mixed	no
	Police	(through job)	no	no	no

Panel B: Field Research staff (excluding authors)				
Activity	Position	Payment	Diverse population	Marginalized population
Focus group	Facilitator			
Survey	Enumerator	yes	no	no
Experiment	RA/facilitator	yes	no	no
Participant Obs.	RA/facilitator	yes	no	no

Table A4: Description of participants in the research process (excluding the PIs). Most ethical considerations apply to subjects in Panel A, though we include research staff for reference in Panel B.



Figure A2: Meeting invitation flyer

## A4 Supplemental Information on Reach and Spillovers

In control neighborhoods, the proportion of residents who reported hearing about police–community meetings increased from 5% to 26%; this was considerably smaller than the increase in treated neighborhoods (5% to 43%), but substantial nonetheless. Given that control neighborhoods were (in some cases) just blocks from treated neighborhoods, and given that signs advertising the meetings were posted outside of meeting locations, it is perhaps unsurprising that there was some spillover in awareness of meetings.<sup>1</sup>

## A5 ITT Analysis

Table A5 includes the operationalization of our pre-specified survey measures which constitute the outcomes of interest in Figure 5. To the security perceptions outcome, we add an index of “*convivencia*” concerns that was not pre-specified as a more concrete assessment of outcomes within the influence of beat police.

Tables A6–A8 report the regression results plotted in Figure 5. Note that all regressions are estimated following Equation 1. Standard errors are clustered by beat, the level of treatment assignment.

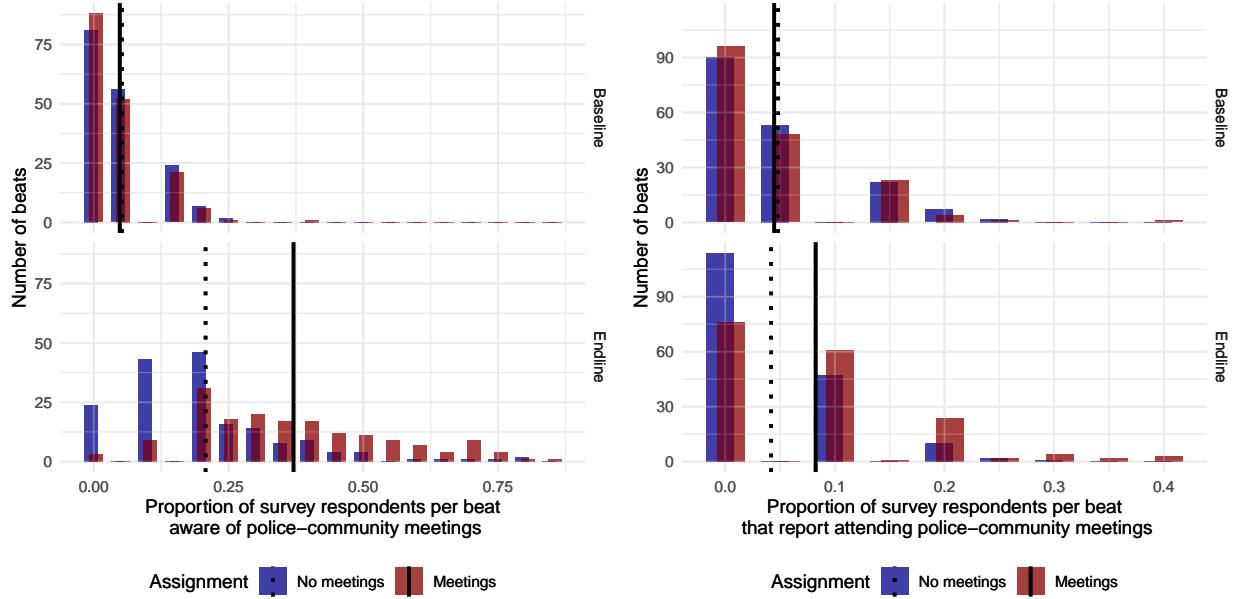
<sup>1</sup>Another possibility is that the increase in awareness of meetings in control neighborhoods reflects a citywide policy change unrelated to our intervention. This strikes us as unlikely, given that neither we nor our research team heard about any such changes, and given that there was no change in meeting *attendance* in control neighborhoods.

Outcome	Details	Range
Trust	<p><i>Responses to the question:</i> How much do you trust the following institutions or groups?</p> <ul style="list-style-type: none"> <li>→ The police</li>   <li>→ The police officers in your neighborhood.</li> </ul>	{1, 2, 3, 4, 5}
Police Quality	<p><i>Index of responses to:</i> To what extent do you agree or disagree with each of the following statements?</p> <ul style="list-style-type: none"> <li>→ The police act upon citizen comments and complaints about security in my community.</li> <li>→ The police take cases seriously and investigate them.</li> <li>→ The police are corrupt.</li> <li>→ The police provide the same quality of service to all citizens.</li> <li>→ The police have the capacity to respond to incidents of crime in a timely manner.</li> <li>→ The police have the capacity to investigate crimes and gather evidence effectively.</li> </ul>	Continuous, $\mu = 0, \sigma = 1$
Other Institutions Quality	<p>Same as above but with respect to</p> <ul style="list-style-type: none"> <li>→ The District Attorney's office.</li> </ul>	Continuous, $\mu = 0, \sigma = 1$
Police relative quality	Difference between Police Quality Index and District Attorney Quality Index	Continuous, $\mu = 0, \sigma = 1^{\dagger}$
Security Perception	<p><i>Index of responses to</i> the following questions:</p> <ul style="list-style-type: none"> <li>→ Overall, how safe do you feel in your neighborhood? (Level).</li> <li>→ With respect to six months ago, how safe do you feel in your neighborhood? (Change).</li> <li>Out of fear, during the last 12 months did you ever ... (Yes/No) <ul style="list-style-type: none"> <li>→ avoid going out alone at night?</li> <li>→ avoid certain streets or roads at night?</li> <li>→ avoid using public transportation?</li> <li>→ avoid new purchases as they could be stolen?</li> <li>→ avoid letting children play on the street?</li> <li>→ purchase any type of firearm?</li> </ul> </li> </ul>	Continuous, $\mu = 0, \sigma = 1$
Convivencia <sup>‡</sup>	<p><i>Index of responses to</i> the following questions: “Compared to 12 months ago, do you think the following situations are better, the same, or worse in your neighborhood?”</p> <ul style="list-style-type: none"> <li>→ Trash in the streets.</li> <li>→ Stray pets.</li> <li>→ Illegally parked cars and motorcycles.</li> <li>→ Open consumption of alcohol and drugs in public space.</li> <li>→ Noise.</li> <li>→ Brawls and fights.</li> <li>→ Lack of lights.</li> <li>→ Lack of respect for authority</li> </ul>	Continuous, $\mu = 0, \sigma = 1$

Table A5: Outcomes measuring citizen beliefs. All outcomes come from the endline citizen survey. The indices are constructed using a  $z$ -score index. <sup>†</sup>: The range of this variable was incorrectly written in the pre-analysis plan, though the operationalization remains the same. <sup>‡</sup>: The convivencia index outcome was not pre-specified.

Figure A3: The Intervention Reached Substantial Fraction of the Population

The left panels plot the distribution (across neighborhoods) of the proportion of residents who report having *heard about* police–community meetings, at baseline and endline, for control and treated neighborhoods. The right panels plot analogous distributions for the proportion who report *attending* police–community meetings. These figures suggest evidence of spillovers for *awareness* of meetings, but not for attendance.



## A6 Supporting Information on Selection

### A6.1 Disaggregation of compliance measures

Our two primary measures of compliance are: (1) *any exposure* to the intervention (i.e., either attended a meeting or “heard about” meetings) and (2) *attended meeting*. The main text documents large and significant differences in these exposure measures across treated and control neighborhoods. Figure A4 shows that these differences stem from: (1) differences in the probability of receiving an invitation and (2) differences in the probability of attending a meeting—not from differences in hearing about the meetings second hand.

### A6.2 Auxiliary results on positive selection into participation

We provide supplementary information about the positive selection described in the manuscript:

- Figure ?? documents positive selection into the intervention in beats assigned to meetings and in beats assigned no meetings. This disaggregates (by treatment condition) the rates reported in Figure 7.
- Figure A5 reports Lasso-selected and regularized coefficient estimates from a model that included the following predictors: both treatment indicators (community meetings and flyers), gender, age (in decades), self-reported socio-economic strata (*estrato*), socio-economic strata of the dwelling at which a respondent was interviewed, occupation, race, income category, *comuna*, and baseline trust in police. All covariates are binary indicators so coefficient magnitudes are comparable. These plots show that baseline trust in police indicators are selected in models that predict both exposure to meetings (left)

Table A6: ITT estimates on “trust in” ...

	Police			Beat Police			Beat Police - Police			Police - Baseline		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Meetings	0.021 (0.034)	0.020 (0.033)	0.011 (0.032)	0.031 (0.036)	0.033 (0.035)	0.021 (0.033)	0.010 (0.039)	0.010 (0.039)	0.007 (0.039)	0.004 (0.048)	-0.008 (0.048)	0.022 (0.038)
Flyers	0.058 (0.034)	0.057 (0.033)	0.052 (0.032)	0.044 (0.036)	0.045 (0.035)	0.039 (0.033)	-0.008 (0.039)	-0.005 (0.039)	-0.008 (0.040)	0.061 (0.048)	0.061 (0.048)	0.081* (0.038)
Block FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline outcome		✓			✓			✓		✓		✓
Control mean	2.264	2.264	2.264	2.600	2.600	2.600	0.334	0.334	0.334	-0.223	-0.223	-0.223
Control std. dev.	1.006	1.006	1.006	1.020	1.020	1.020	1.092	1.092	1.092	1.117	1.117	1.117
R <sup>2</sup>	0.001	0.029	0.108	0.001	0.031	0.145	0.000	0.029	0.033	0.001	0.039	0.333
Adj. R <sup>2</sup>	0.000	0.005	0.085	0.000	0.006	0.123	-0.001	0.003	0.007	-0.000	0.002	0.308
Num. obs.	3597	3597	3597	3476	3476	3476	3441	3441	3441	2399	2399	2399
RMSE	1.021	1.019	0.977	1.034	1.031	0.968	1.094	1.092	1.090	1.162	1.161	0.967
N Clusters	347	347	347	347	347	347	347	347	347	347	347	347

\*\*\*p < 0.001; \*\*p < 0.01; \*p < 0.05

Table A7: ITT estimates on beliefs about quality (index) of ...

	Police			Beat Police			Beat Police - Police			Prosecutors			Police - Prosecutors		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Meetings	0.075* (0.034)	0.074* (0.035)	0.050 (0.031)	0.049 (0.034)	0.051 (0.034)	0.027 (0.031)	-0.027 (0.023)	-0.025 (0.023)	-0.024 (0.023)	0.020 (0.038)	0.018 (0.035)	0.003 (0.032)	0.039 (0.030)	0.041 (0.030)	0.035 (0.029)
Flyers	-0.038 (0.034)	-0.039 (0.035)	-0.021 (0.031)	-0.006 (0.034)	-0.005 (0.034)	0.013 (0.031)	0.032 (0.023)	0.034 (0.023)	0.033 (0.023)	-0.042 (0.023)	-0.041 (0.038)	-0.033 (0.035)	0.015 (0.032)	0.015 (0.030)	0.019 (0.029)
Block FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline outcome		✓			✓			✓		✓		✓		✓	✓
Control mean	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.003	0.003	0.003
Control std. dev.	1.000	1.000	1.000	1.000	1.000	1.000	0.693	0.693	0.693	1.000	1.000	1.000	0.911	0.911	0.911
R <sup>2</sup>	0.002	0.027	0.228	0.001	0.026	0.215	0.001	0.030	0.031	0.001	0.049	0.187	0.001	0.030	0.078
Adj. R <sup>2</sup>	0.001	0.003	0.208	0.000	0.002	0.195	0.000	0.006	0.006	-0.000	0.025	0.166	0.000	0.005	0.054
Num. obs.	3641	3641	3641	3610	3610	3610	3609	3609	3609	3612	3612	3612	3606	3606	3606
RMSE	0.994	0.994	0.885	1.000	0.999	0.897	0.673	0.671	0.671	0.978	0.966	0.893	0.884	0.881	0.860
N Clusters	347	347	347	347	347	347	347	347	347	347	347	347	347	347	347

\*\*\*p < 0.001; \*\*p < 0.01; \*p < 0.05

and attendance at meetings (right). Among all predictors, high baseline trust in police was the second most predictive predictor of exposure (after only assignment to treatment) and was the most predictive predictor of attendance (surpassing assignment to treatment).

- Figure A6 plots selection into the treatment as a function of the baseline police quality index. We observe similarly dramatic positive selection as a function of these prior beliefs about the police.

Table A8: ITT estimates on perceptions of convivencia and personal security (indices are *increasing* in convivencia progress and security):

	Convivencia			Convivencia - Baseline Convivencia			Security			Security - Baseline Security		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Meetings	0.074 (0.039)	0.073* (0.037)	0.043 (0.034)	0.039 (0.041)	0.038 (0.042)	0.088* (0.037)	0.007 (0.038)	0.007 (0.035)	0.002 (0.033)	0.016 (0.046)	0.001 (0.043)	0.011 (0.037)
Flyers	0.007 (0.039)	0.009 (0.037)	-0.002 (0.034)	0.013 (0.041)	0.018 (0.042)	0.041 (0.037)	-0.016 (0.038)	-0.016 (0.035)	-0.039 (0.032)	-0.085 (0.046)	-0.076 (0.044)	-0.039 (0.037)
Block FE		✓	✓		✓	✓		✓	✓		✓	✓
Baseline outcome			✓			✓			✓			✓
Control mean	0.000	0.000	0.000	-0.044	-0.044	-0.044	0.000	0.000	0.000	-0.015	-0.015	-0.015
Control std. dev.	1.000	1.000	1.000	0.942	0.942	0.942	1.000	1.000	1.000	1.034	1.034	1.034
R <sup>2</sup>	0.001	0.044	0.199	0.000	0.039	0.300	0.000	0.051	0.188	0.002	0.053	0.303
Adj. R <sup>2</sup>	0.001	0.020	0.178	-0.000	0.003	0.273	-0.000	0.027	0.167	0.001	0.018	0.277
Num. obs.	3643	3643	3643	2434	2434	2434	3644	3644	3644	2434	2434	2434
RMSE	0.981	0.971	0.889	0.960	0.959	0.819	0.980	0.966	0.894	1.009	1.001	0.859
N Clusters	347	347	347	347	347	347	347	347	347	347	347	347

\*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$

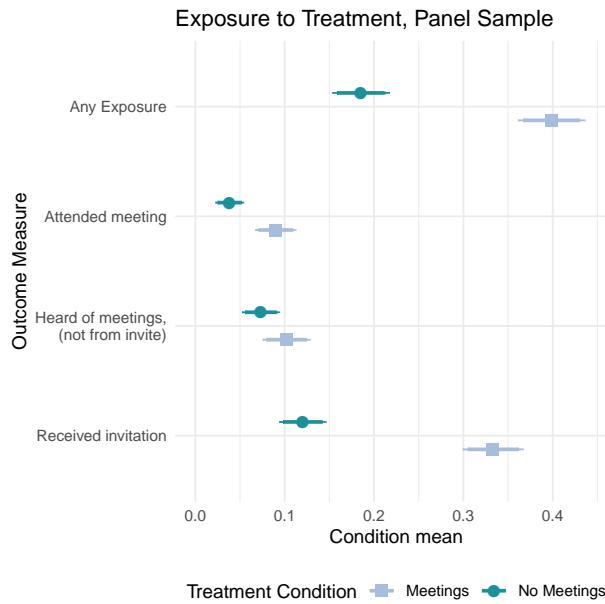


Figure A4: Mean levels of compliance with treatment as a function of assignment to police-community meetings. The bars depict 95% confidence intervals, which are constructed on standard errors clustered at the beat level.

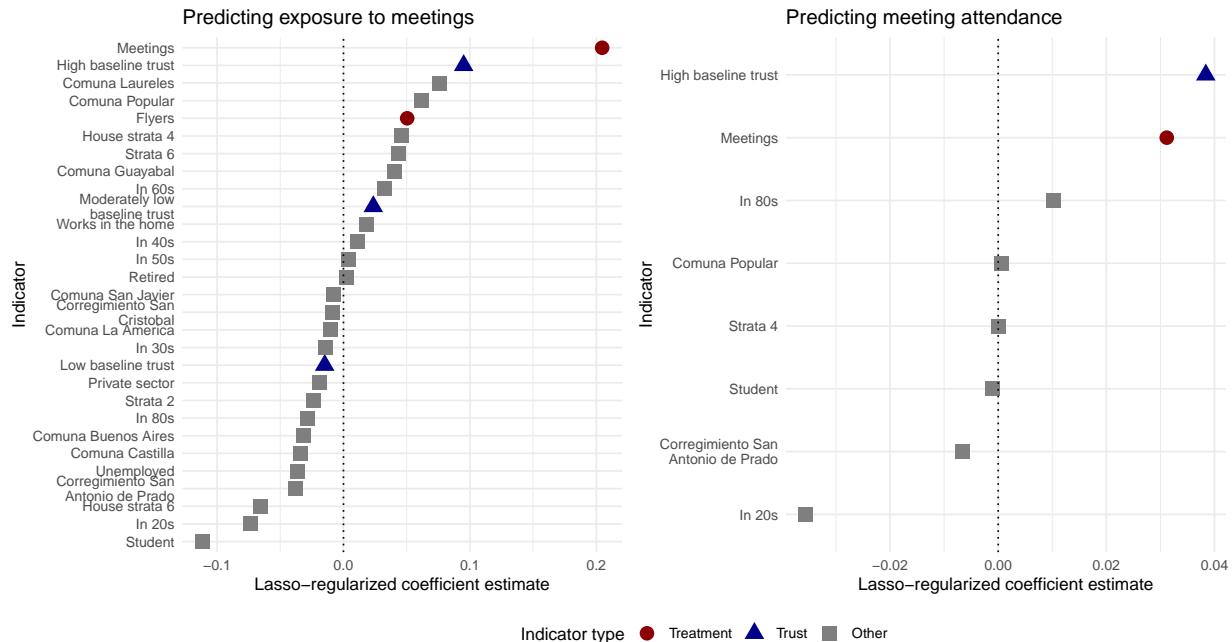


Figure A5: Lasso-selected and regularized coefficient estimates from a model that included individual demographic and community-level predictors in addition to trust in police. All variables are indicators so the plotted coefficients are comparable. High trust in police is the second strongest predictor in the left panel and the strongest predictor in the right panel.

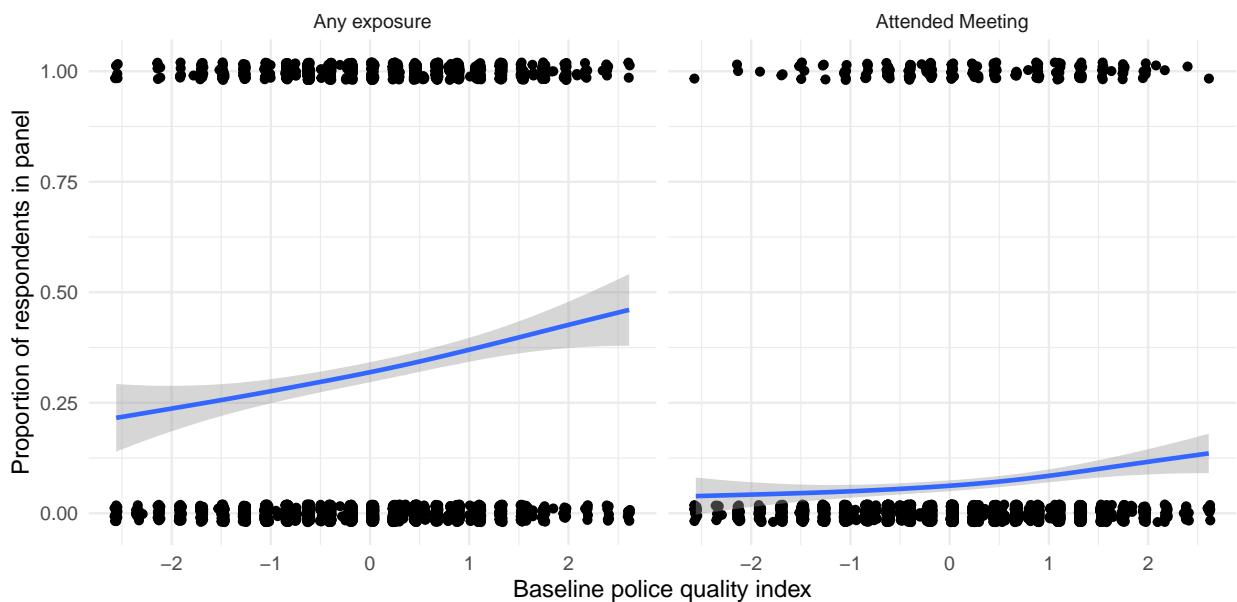


Figure A6: Positive selection as a function of baseline assessments of police quality. Points represent individual respondents in the panel sample and are jittered for easier visibility. The blue lines are estimated by loess regression.

### A6.3 Positive selection in other Metaketa Experiments

This experiment was one of six studies in the Metaketa IV. We document the extent of positive selection in the other five sites in the main text. Table A9, we provide additional details about the surveys conducted in each site as well as estimated treatment effects on trust. We note substantial variation in (1) recontact rates (among panel surveys) and (2) autocorrelation of trust.

Country	Survey design	$N$ respondents		Autocorrelation of trust in police	ITT estimate on...	
		in endline	in panel		Trust in police	$\Delta$ in trust in police
Brazil	Panel	1,513	841	0.14	0.28 [0.05, 0.51]	0.09 [-0.05, 0.23]
Colombia	Panel	3,644	2,434	0.48	0.02 [-0.05, 0.08]	0.02 [-0.05, 0.08]
Liberia	Repeated cross-section	1,850	–	-0.01 <sup>†</sup>	0.03 [-0.09, 0.16]	0.03 [-0.09, 0.16]
Pakistan	Panel	3,449	3,175	-0.02	0.18 [0.03, 0.33]	0.10 [-0.02, 0.22]
Philippines	Endline only	4,470	–	–	-0.03 [-0.12, 0.55]	–
Uganda	Panel	3,456	2,946	0.38	0.02 [-0.09, 0.14]	0.03 [-0.05, 0.12]

Table A9: Features of surveys in studies across the meta-analysis. <sup>†</sup> note that this represents the community-level autocorrelation given the repeated cross-section survey design.

### A6.4 Non-survey evidence of differential selection into community policing

Our measures of selection rely on self-reported attendance data. While we do not believe that citizens had any reason to systematically misreport exposure to or attendance at these community-police meetings, we use a behavioral measure to further interrogate the assertion that meeting attendees did not represent a random subset of their communities. Our attendance data, collected by our RAs in each meeting, provides the (1) count of attendees, and (2) the gender breakdown of attendees (as assessed by the RAs).

Attendance breakdown by gender allows us to estimate  $\Pr(\text{Female}|\text{Attended})$  and  $\Pr(\text{Male}|\text{Attended}) = 1 - \Pr(\text{Female}|\text{Attended})$  at the meeting level.<sup>2</sup> In principle, we would like to show that  $\Pr(\text{Attended}|\text{Female}) - \Pr(\text{Attended}|\text{Male}) \neq 0$ . Using Bayes' rule, we can show how these quantities relate:

$$\Pr(A_c|F_c) - \Pr(A_c|\neg F_c) = \frac{\Pr(A_c)\Pr(F_c|A_c)}{\Pr(F_c)} - \frac{\Pr(A_c)(1 - \Pr(F_c|A_c))}{1 - \Pr(F_c)} \quad (1)$$

$$= \Pr(A_c) \left( \frac{\Pr(F_c|A_c)}{\Pr(F_c)} - \frac{1 - \Pr(F_c|A_c)}{1 - \Pr(F_c)} \right) \quad (2)$$

Note that  $\Pr(A_c) > 0$  for all meetings that were not cancelled, i.e., those where citizens were present. Further, census data give us a sense of the empirical distribution of  $\Pr(F_c)$ . We also know that  $\Pr(F_c) \approx 0.5$  in all beats from census data. Figure A7 plots the empirical distribution of  $\Pr(F_c) - (1 - \Pr(F_c))$ , revealing that women were much more likely to attend.

---

<sup>2</sup>The RAs classified all attendees as male or female.

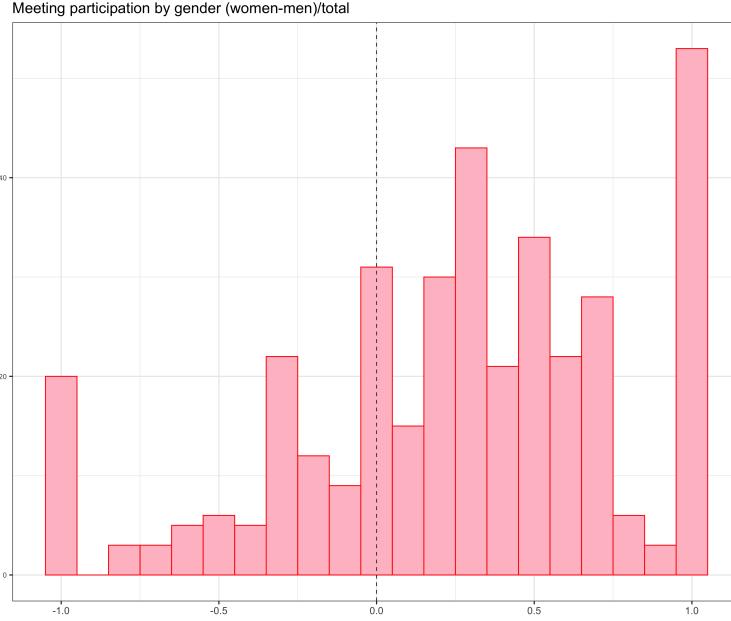


Figure A7: The gender composition of meeting attendees, as classified by RAs. This plot depicts  $\Pr(F_c) - (1 - \Pr(F_c))$  given the notation in Equation 2.

## A7 Framework for selection and updating

To motivate our simulations, define three groups  $G$ : those that attended meetings ( $G = m$ ), those that heard about meetings ( $G = h$ ), and those that neither attended nor heard about meetings ( $G = n$ ). Our goal is to compare posterior beliefs  $E[\mu|G = g]$  for different  $g$ . Recall that the prior belief,  $\pi \in \{1, 2, 3, 4\}$ . The group-specific posterior is a weighted sum of the conditional posteriors for each level of the prior:

$$E[\mu|G = n] = \sum_{x=1}^4 \Pr(\pi = x|G = n)\rho_x$$

where  $\rho_x = E[\mu|\pi = x, G = n]$

$\rho_x$  captures the process generating posterior beliefs. It includes some combination of updating, mean reversion, and anything else that changes beliefs between baseline and endline.

Heterogeneity in the quality of the signal affects the extent of updating. One might think that the tone or quality of the police–community meetings would be determined largely by neighborhood characteristics (such as income) or by police station, given that station chiefs appeared to vary in their commitment to the meetings, and given that station chiefs have considerable influence over officer behavior. In fact, though, our dictionary-based measure of meeting sentiment appears uncorrelated with any of these factors. Figure A9, for example, shows that neighborhood-average trust in police at baseline is a strong predictor of the number of meetings that actually took place (left panel) but does not at all predict meeting sentiment. In other words, it is not the case that the most-positive meetings generally took place in the most pro-police neighborhoods. Rather, the tone of the meetings was highly idiosyncratic.

Figure A8: Baseline Neighborhood Trust Uncorrelated with Meeting Sentiment

Neighborhood-average baseline trust in police strongly predicts the number of held (vs. cancelled) police–community meetings, but does not predict meeting sentiment.

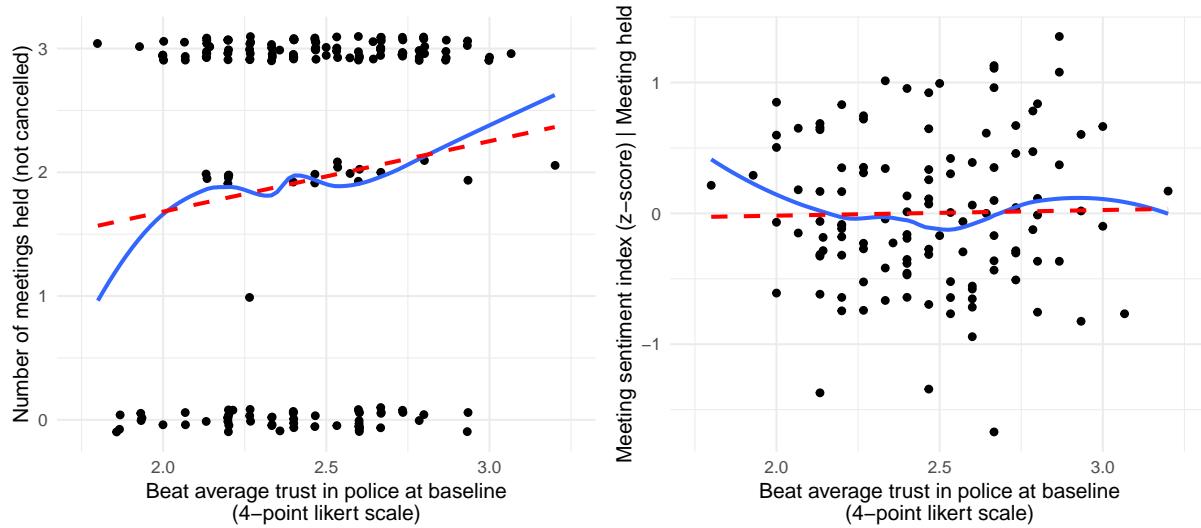


Figure A9: Relationship between baseline trust in police (prior beliefs) at the beat level and number of meetings realized (left) and average meeting sentiment (right). Sentiment is measured from the ethnographic notes documenting meetings.

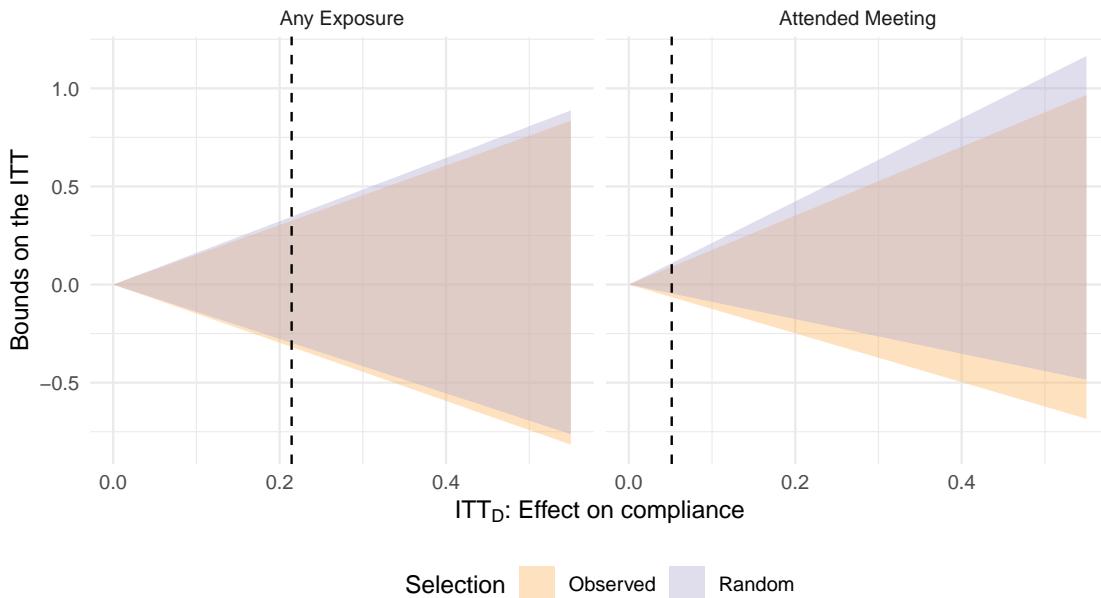


Figure A10: Extreme value bounds on the trust outcome (a 4-point scale). The left panel assumes that treatment effects come through hearing about the intervention (including attending the meetings). The right panel assumes that treatment effects come only through attending the meetings. The purple triangles depict bounds under random selection into meetings; the orange triangles depict bounds under the observed selection into meetings. The orange triangles are shifted down, indicating that positive selection shrinks the range of possible treatment effects. The dashed lines depict the  $ITT_D$  (first stage) for each compliance measure.

## A8 Three challenges of positive selection

### A8.1 Censoring

Positive selection may attenuate treatment effects due to top-censoring. Because our survey outcomes are measured on a Likert scale, trust (or assessments of the police) could increase *within* the top category on the scale. We conduct two exercises to understand how this might affect our estimates. First, we construct modified extreme value bounds. These bounds are modified in the sense that they make assumptions about whose outcomes are affected by treatment. We consider two scenarios:

- Anyone who reported hearing about the meetings (in treatment or control)
- Anyone who reported attending a meeting (in treatment or control)

While we do make assumptions about who may be affected by the intervention, extreme value bounds make no assumptions about the sign or magnitude of those effects. Figure A10 plots these bounds as a function of the level of compliance ( $ITT_D$ ) under random and observed patterns of selection. We see that the observed positive selection, which is stronger for attendance than for hearing about meetings, reduces the range of possible treatment effects.

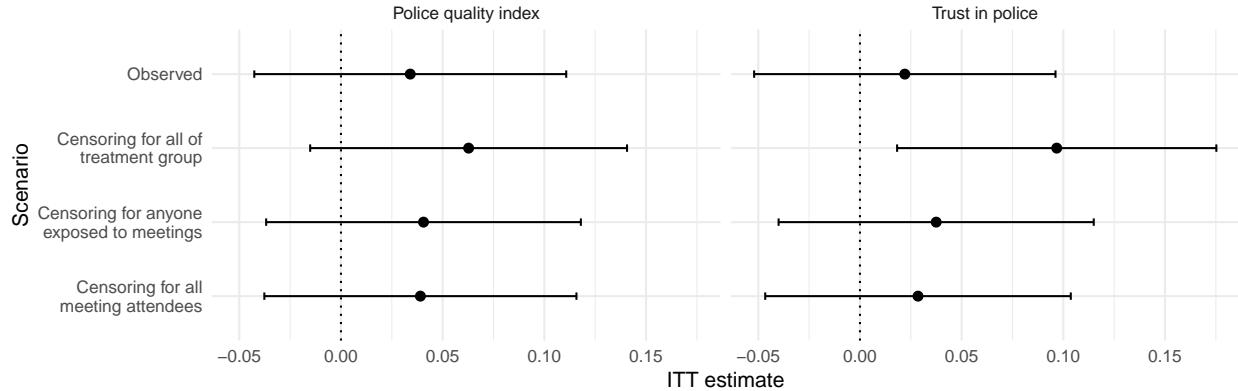


Figure A11: Estimates of observed ITT effects (in the panel sample) alongside those produced by the three censoring scenarios described above. Bars represent 95% confidence intervals constructed on clustered standard errors.

In Figure A11 we explore how large the ITT effects may have been in the absence of top-censoring. To construct these estimates for the panel sample, we assume that a subject with prior  $\pi = \max\{\pi\}$  and posterior  $\mu = \max\{\mu\}$  (expressed on a Likert scale), could have expressed posterior  $\tilde{\mu} = \mu + 1$ . Because the police quality index is constructed from multiple Likert scales, we impute the counterfactual values of each of the constituent indicators before constructing the index. We consider three scenarios:

- The responses of any subject assigned to community meetings,  $Z_i^m = 1$  were top-censored. (This creates a strict upper bound on the ITT.)
- The responses of any subject who heard about community meetings (in treatment or control) were top-censored.
- The responses of any subject who attended community-police meetings were top-censored.

## A8.2 Mean reversion

We investigate whether the observed posterior beliefs could be generated by mean reversion alone. To do so, we compare the observed posterior beliefs to counterfactual posteriors  $\mu_{CF}$  generated by assuming that, conditional on a prior belief  $\pi$ , all subjects (meeting attendees and non-attendees) mean revert at the same rate. We use block bootstrapping to estimate this counterfactual distribution of conditional posteriors, assuming a common mean-reversion process:

1. Create a sample of  $n_j$  observations from the baseline panel data in each cluster,  $j$ , where  $n_j$  is the number of observations in the panel at endline.
2. Using observations without exposure to treatment, i.e., those in the set  $G = n$ , we estimate  $\hat{\rho}_x$  for  $x \in \{1, 2, 3, 4\}$ .
3. Estimate  $E[\hat{\mu}_{CF}, G = g] = \sum_{x=1}^4 \Pr(\pi = X | G = g) \hat{\rho}_x$  for  $g \in \{h, a\}$ .

Bootstrapped predictive distribution of endline trust in police assuming mean regression at rate of mean regression among unexposed citizens

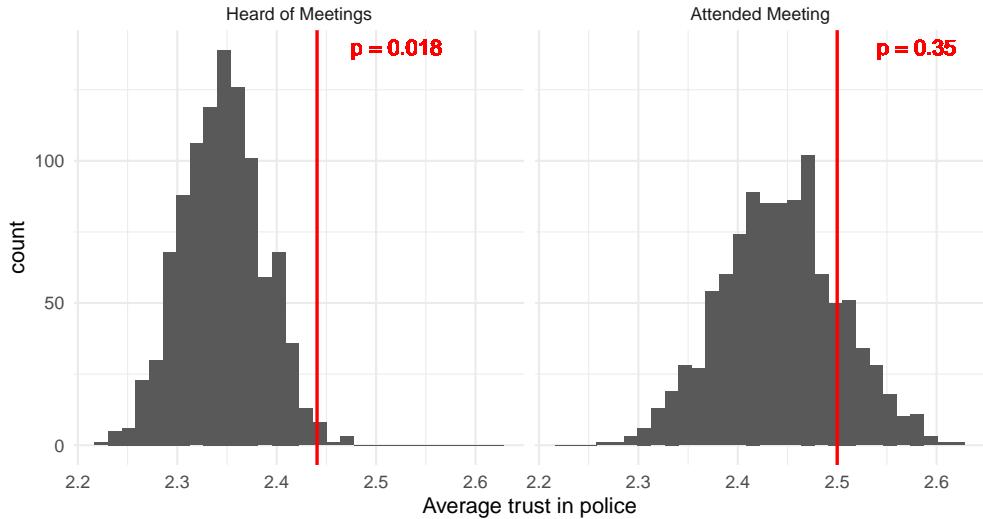


Figure A12: Comparing observed posterior beliefs to counterfactual distributions generated by a common mean-reversion process.

We repeat steps #1–3  $k = 1000$  times to construct a counterfactual distribution of  $E[\hat{\mu}_{CF}|G = g]$ . We can then draw inferences about the likelihood of observed conditional posteriors under the assumption of equivalent mean-reversion across groups. We plot the results of this exercise in Figure A12. We show that posterior trust in police is systematically more *positive* than what we would expect due to mean reversion alone. The difference is significant for those that *heard of* meetings but not for the smaller sample that attended meetings. There are two possible explanations for this finding. First, citizens who hear about police–community meetings are updating positively as a result. Second, this group of citizens who participate is different and therefore mean-reverting at a different rate than non-participants, for reasons unrelated to the signal observed in the meetings. This could occur if, after conditioning on priors, the idiosyncratic error in this survey response were lower for participants than for non-participants.

### A8.3 Preaching to the choir (updating)

In the main text, we provide a simulation showing that both the ITT and the ATT are decreasing in the degree of positive selection. In Figure A13 we show that the *observed* level of positive selection attenuates our estimates of both the ITT and the ATT—no matter what the (assumed) quality of the signal provided by the meetings. (Obviously, higher signals produce more positive updating; the figure shows the *difference* in treatment effects, comparing observed with random selection into meetings.)

## A9 Relationship to Gonzalez and Mayka (2022)

Gonzalez and Mayka (2022) find that community-police meetings result in calls for repression of marginalized groups in São Paulo. We replicate their analysis for one marginalized group—youth/children, the group subject to the most calls for repression in São Paulo—using our data from police–community meetings in

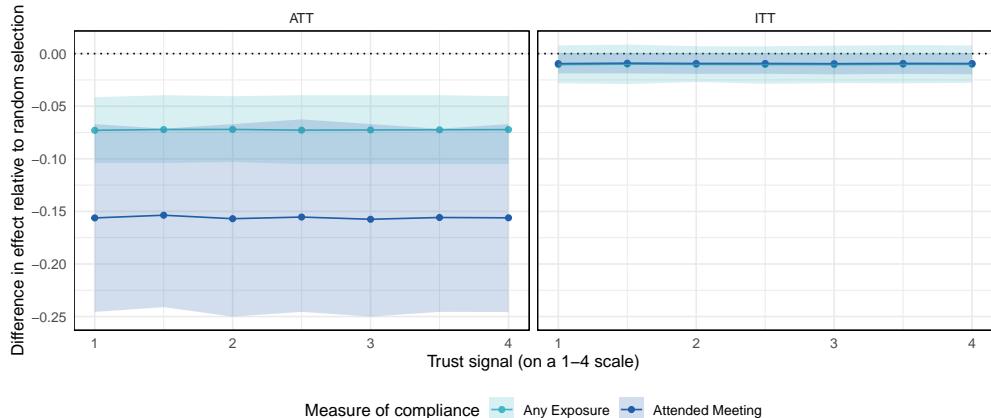


Figure A13: The effect of observed rates of positive selection relative to random selection on the ATT and ITT, under different definitions of compliance and different signals.

Medellín.

Gonzalez and Mayka hand-code all 793 meeting minutes in their collection. We instead (1) use an automated search for youth (“*joven*”) or children (“*nin*” stem) to locate the subset of our meeting minutes that mention youth/children, and then (2) hand-code a random sample of that subset, marking which of them include calls for repression. Specifically, 52% of the 519 meeting notes ( $n = 270$ ) mention *youth* or *children*;<sup>3</sup> we hand-code a random sample of 18.5% of these meeting notes ( $n = 50$ ). The confidence intervals reported below reflect this sampling from the subset of 270 meeting notes that mention youth/children.

- **$n = 50$ :** Random sample of meeting notes that contain youth or children.
  - $n = 6$  (0.12, 95% CI: [0.03, 0.21]) are *irrelevant* mentions. In these six meetings, the only mention of *youth* or *children* was in reference to the physical environment of the meeting. For example, one set of case notes describes a mural depicting children on the wall of the meeting site.
  - $n = 14$  (0.28, 95% CI: [0.16, 0.40]) use these terms only to *describe the traits of attendees or police officers*. For example, they note how many children were present or describe a participant as young (“*joven*”).
  - $n = 30$  (0.6, 95% CI: [0.46, 0.73]) describe problems related to children or youth. As we discuss below, these discussions emphasize both crimes/misdemeanors committed by youth and victimization of youth in approximately equal proportion.
    - \*  $n = 3$  (0.06 95% CI: [-0.01, 0.13]) describe calls for police action against a child or youth. Following Gonzalez and Mayka (2022), who code requests for police to “do something” as calls for repression, we code these three meetings as including calls for repression. None of the three of them, however, involves calls for police violence like those that Gonzalez and Mayka observe in São Paolo (most egregiously, when one attendee reported that police

<sup>3</sup>Note that “*joven*” is also an adjective meaning “young,” which risks false-positive classifications. We remedy this in the hand coding below.

killed her neighbor's son because he was stealing, "Other attendees responded 'amen' and that the officer 'performed a service,'" Gonzalez and Mayka 2022, 13).

Among the 30 meeting notes that describe problems related to children or youth, 77% describe crimes—most frequently marijuana use—and 70% describe various forms of victimization of children or youth. Obviously, given these proportions, these discussions are not mutually exclusive. We provide illustrative quotes below. Note that specific names and location information (neighborhoods or addresses) are omitted and replaced with *XX* to preserve anonymity.

For examples of discussion of victimization of children (in different regards):

**Meeting # 116:** "Doña *XX* cuenta que hay como cuatro niños que han cogido bacterias de cuenta del mal manejo que las personas tienen sobre sus mascotas, pues no recogen el excremento."

*Translation:* "Mrs. *XX* says that there are about four children who have bacterial infections due to the negligence of pet owners who do not pick up the excrement."

**Meeting #403:** "por otro lado nos preocupa la cantidad de niños que llegaron de Venezuela enfermos y sin vacunas, están propagando enfermedades, porque ellos aquí no tienen EPS"

*Translation:* "On the other hand, we are worried about the number of children who arrive from Venezuela sick and unvaccinated who are spreading illness because they do not have access to EPS [health insurance]."

Some examples include both victimization of and crimes committed by youth. For example:

**Meeting # 21:** "La mujer joven interviene y dice que a los niños les están robando las bicicletas por *XX*, que lo peor es que los que roban son muchachitos más grandes o adolescentes, que les sacan navaja a los pequeños y se les llevan las bicicletas."

*Translation:* "The young women intervenes and says that they are stealing children's bicycles in *XX*. Worse than that, the robbers are bigger boys or adolescents and that they rob the children at knifepoint."

We report the three instances of calls for police action ("repression" per Gonzalez and Mayka 2022):

**Meeting #108:** "en el barrio queremos que haya más sentido de pertenencia cuando la policía patrulla, porque pasan de largo y no hay complemento de civismo, queremos que paren en ciertos lugares, que no pasen de largo, especialmente a la salida del colegio, pues se presentan jóvenes haciendo piques en la vía del colegio sin cascos representando un peligro para los transeúntes que en su mayoría son menores."

*Translation:* "in the neighborhood, we want the police to patrol with more of a sense of belonging, because right now they just pass by, without any real public spirit. We want them to stop in certain places, not to pass through, especially when [children are] leaving school, because young people show up doing wheelies on motorcycles on the school road without helmets, representing a danger to passers-by who are mostly minors."

**Meeting #340:** "Nos cuenta que ha llamado a la policía para denunciar a unos jóvenes que se acumulan cerca a su casa hacen escándalo y consumen vicio. Dice que la policía le dice que

van en camino pero nunca llegan.”

*Translation:* “They tell us that he has called the police to report some young people who congregate near their house, make disturbances and do drugs. They say the police tell them they’re on their way but they never arrive.”

**Meeting #492:** “Luego el señor de camisa a cuadros toma la palabra y expone que en el taller de motos que hay en la XX hay constantemente consumo de drogas y esto lo hacen sobre toda la vía pública, se hacen pequeños grupos de 3 ó 4 jóvenes a consumir ya ‘planear sus cosas,’ dando a entender que también delinquen, menciona que el dueño del taller también consume drogas junto a los demás jóvenes que van al taller, por eso su petición es que la patrulla pase más, que haga mayor patrullaje en el sector.”

*Translation:* “Then the man in the checkered shirt speaks up and explains that in the motorcycle repair shop on XX[<sup>th</sup> street] there is constant drug use and they do this all over the public street. They form small groups of 3 or 4 young people to consume and ‘plan their things,’ implying that they also commit crimes. He mentions that the owner of the shop also uses drugs along with the other young people who go to the workshop. For this reason his request is that the patrol go more, that he patrol the sector more.”

In contrast, some meeting-goers actually called out actions by the police against young people:

**Meeting #127:** “Una mujer de más o menos 60 años con un problema en su boca que la obligaba a hablar más despacio y con alguna dificultad comentó sobre un caso en el cual a un joven que pocas veces hace ruido, le retuvieron su equipo de sonido y, la persona que siempre ha hecho ruido en su cuadra permanece con el suyo. El patrullero XX le pregunta en qué cuadra vive y dice que verificará el caso; la líder XX sostiene que ese joven sí realizaba ruido y la mujer le responde que era muy poco.”

*Translation:* “A woman around 60 years old with a speech impediment that forced her to speak more slowly and with some difficulty brought up a case in which a young man who rarely makes noise had his sound system withheld [by police] and, a person who has always made noise on her block still has his sound system. Patrolman XX asks her which block she lives on and says that he will verify the case; the community leader XX maintains that this young man did make noise and the [first] woman replies that it was very little.”

Finally, we note that many of the proposed solutions to youth misbehavior involve the intervention of parents, teachers, or the community, *not* police action. These solutions substitute for calls for police action:

**Meeting #3:** “XX les cuenta el caso de un niño en una unidad que salían a trabajar los padres desde las 7 am hasta las 8 pm, llegaban y la empleada decía que todo bien, se despedían del hijo si lo veían despierto, y les pregunta, De quién es la culpa si el niño toma malos hábitos?”

*Translation:* “XX tells us about the case of a child from a household in which the parents work from 7 am until 8 pm, arrive home and the domestic employee says everything was fine, they said goodbye to the child if they even saw him awake, and they wonder: whose fault is it if the child develops bad habits?”

**Meeting #240:** “[El] líder . . . dice que antes se podía trabajar con la rectora, hacer planes conjuntos para atraer a los jóvenes y niños al buen camino, ahora el nuevo rector no colabora

en nada, dice que en el colegio *XX* la problemática de consumo es muy evidente y el rector en vez de tomar medidas que ayuden, lo que hace es alejar más a los jóvenes. Por ejemplo, cuando llegan tarde los devuelve para la casa, pero la mayoría se queda en la calle y ahí aprovechan los “muchachos” para ponerlos como carritos y, de esa manera, los van introduciendo al camino de lo ilegal.”

*Translation:* “The [community] leader . . . says that before it was possible to work with the rector, make common plans to attract young people and children to the right path. Now the new rector does not collaborate at all. The leader says that in the *XX* school, the problem of drug consumption is very evident and instead of taking measures that help, the rector’s actions alienate young people further. For example, when they [students] arrive late, the rector sends them home, but most of them stay on the street and the ‘boys’ take advantage of it to use them as dealers and, in this way, introduce them to the illegal path.”

## A10 Anonymized Pre-Analysis Plan

On pages A-22–A-55, we reproduce an anonymized version of our pre-analysis plan for this experiment. We note that the analysis of the LAPOP survey data was added in response to our findings of positive selection into community-police meetings.

PRE-ANALYSIS PLAN  
THE CONSTRUCTION OF TRUST IN THE STATE:  
EVIDENCE FROM POLICE-COMMUNITY  
RELATIONS IN COLOMBIA



**Abstract**

We outline and pre-specify our analysis for research products from a field experiment on community policing in Medellín, Colombia. The experiment consists of two interventions, crossed in a  $2 \times 2$  factorial design: the creation of beat-level community policing meetings and the dissemination of security-related information (on norms, procedures, and outcomes). We first propose a paper on the evolution of *trust in the police*. Second, we propose a paper on *gender and demand for better policing*. Finally, we will contribute to the Metaketa-IV analyses and joint product with a discussion of how the efficacy of these interventions can be compared across contexts, given the evidence. All products draw on data collected from: baseline and endline surveys of citizens; endline surveys of community police; intervention implementation records; administrative data on crime and population; ethnographic observation of community policing meetings; and qualitative focus groups and interviews.



**Contents**

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Preliminary Statement of Research Outputs</b>	<b>5</b>
<b>3</b>	<b>Theory and Arguments</b>	<b>5</b>
3.1	Theoretical Framework: On Building Trust . . . . .	5
3.2	Theoretical Framework: Gender and Demand for Policing . . . . .	7
3.3	Note: Metaketa Arguments . . . . .	8
<b>4</b>	<b>Experimental Design</b>	<b>8</b>
4.1	Treatments . . . . .	8
4.2	Unit of assignment to treatment . . . . .	10
<b>5</b>	<b>Data and measurement</b>	<b>13</b>
5.1	Overview . . . . .	13
5.1.1	Administrative Data . . . . .	13
5.1.2	Survey Data . . . . .	14
5.1.3	Qualitative Data . . . . .	16
5.2	Outcome measurement: Trust paper . . . . .	17
5.2.1	Manipulation checks . . . . .	17
5.2.2	Citizen Beliefs . . . . .	18
5.2.3	Officer Beliefs . . . . .	20
5.2.4	Citizen Behavior . . . . .	20
5.3	Outcome Measurement: Gender and Demand for Policing . . . . .	21
5.4	Outcome Measurement: Metaketa Contribution . . . . .	22
<b>6</b>	<b>Estimation</b>	<b>22</b>
6.1	ITT Estimators . . . . .	22
6.2	Estimation: Trust Paper . . . . .	23
6.2.1	Baseline Trust Levels: Prior Beliefs . . . . .	23
6.2.2	Does the intervention lead people to update their beliefs? . . . . .	24
6.2.3	Does the intervention lead to different police beliefs about citizens? . . . . .	25
6.2.4	Does the intervention change crime reporting behavior? . . . . .	26
6.3	Estimation for Gender and Demand for Policing Paper . . . . .	27
6.4	Estimation for Metaketa Contribution . . . . .	28
<b>7</b>	<b>Power</b>	<b>28</b>
7.1	Attrition rate and serial correlation . . . . .	29
7.2	Effect Size . . . . .	30
<b>8</b>	<b>Ethics</b>	<b>31</b>
8.1	Risk / Benefit Assessment . . . . .	31

<b>A Appendix</b>	<b>A1</b>
A.1 Complete list of Metaketa hypotheses . . . . .	A1
A.2 Metaketa Estimators . . . . .	A2
A.3 Meta-Analysis Outcomes and Coding . . . . .	A3
A.4 Additional tables and figures . . . . .	A4
A.5 Meeting Protocol . . . . .	A5

## 1 Introduction

Trust in government is the subject of “an industry of research” ([Levi and Stoker, 2000](#)) in political science, sociology, and economics. Since at least the 1960s, researchers have worked to define trust in government, to develop survey instruments that measure it, to pin down its determinants, and to understand its consequences, which include political and economic behaviors as fundamental as voting, obeying the law, or even buying goods.

This large literature has produced robust findings. But most studies focus on trust in an abstract *government* or in national institutions like Congress; less is known about the causes or consequences of trust in specific, local government agents. [Levi and Stoker \(2000\)](#) argue that this type of trust—trust in particular authorities—is essential to understanding “everyday political behavior,” and they therefore urge scholars to expand their inquiries beyond the traditional focus on citizens’ trust in ‘government’ in general” (p. 495–496).

This study investigates the determinants and consequences of citizen trust in a specific, local government institution that performs an elemental function of the state: the police. Criminologists have used survey data to enumerate the correlates of trust in various police forces (e.g., [Tyler, 2005](#); [Sunshine and Tyler, 2003](#)), but empirical political scientists, sociologists, and economists have largely ignored the subject—despite the fact that theorists in all three fields emphasize its importance to successful law enforcement.<sup>1</sup>

Policymakers have not waited for social science. Governments around the world have implemented policies aimed at improving trust in the police. A 2008 police reform in England, for example, eliminated all nationally mandated performance targets for local police forces *except* targets for improving public confidence, as measured by the British Crime Survey ([Home Office, 2008](#), p. 84).

We use a large-scale field experiment in Medellín, Colombia, to study the effects of two initiatives on citizen trust in the police (as reported in surveys), cooperation with law enforcement (as measured in administrative records), and officer perceptions of citizens (as reported in surveys of officers).

The first initiative facilitates *interaction* among police officers and citizens by holding town-hall-style police-community meetings. Other governments have implemented similar programs; in Chicago, for example, police hold regular beat meetings, which “provide an opportunity for police and residents to get acquainted, and to build mutual respect and trust” ([Skogan, 2006](#)).

The second initiative is to provide *information* about police responsibilities, security outcomes, and social norms about reporting. Many governments have invested in information campaigns aimed at improving opinions of the police, and a handful of experimental studies suggest that these campaigns can work (e.g., [Wunsch and Hohl, 2009](#); [Ardanaz, Corbacho, and Ruiz-Vega, 2014](#)).

This pre-analysis plan outlines our planned research outputs (Section 2), describes our theoretical framework and hypotheses (Section 3), experimental design (Section 4), data (Section 5), and hypothesis tests (Section 6).

---

<sup>1</sup>For example, [Akerlof and Yellen \(1994\)](#) emphasize that “the major deterrent to crime is not an active police presence but rather presence of knowledgeable civilians, prepared to report crimes and cooperate in police investigations” (p. 2). See also [Acemoglu and Jackson \(2017\)](#).

## 2 Preliminary Statement of Research Outputs

We propose three central quantitative research outputs based on the data collected here: two papers and a contribution to the Metaketa IV joint output, provisionally a chapter in an edited volume. Our provisional plan for the outputs is as follows:

1. *On Building Trust*: We posit a theory about trust in the state, conceiving of trust as a two-way relationship that requires understanding both citizen beliefs about the state (or agents thereof) in addition to state agents' beliefs about citizens. We use two programs, community policing and an information campaign, to measure how officers and citizens update their beliefs about each other. We measure outcomes using survey data, participation logs, and administrative data on the reporting of crime.
2. *On Demand for Better Policing*: Stronger police-community relationships rely, in part, on citizens' propensity to engage with the police. The structure of our community policing program ensures that a random sample of residents of the *manzana* of each *cuadrante* were invited to meetings. We study who attends these meetings, benchmarking attendance to new (2018) census data. In particular, we examine the empirical basis for a common observation: most attendees are women. We first establish this pattern empirically. Then we use a multi-method approach to understand differences in demand for engagement with police, drawing upon administrative census data, pre-treatment crime reports, survey data, focus group data, and ethnographic notes from each of 522 community policing meetings.
3. *On the Comparison of Policies to Improve Security* (meta-article and Metaketa book chapter): This chapter contains the analyses of the meta-analysis outcomes from the Colombia project. We then consider the merits of comparing the estimated effects of the two treatments: community policing and flier campaign. Evidence-based policy-making requires comparison of different policy options. Yet, comparisons are less than straightforward even with a design as primed for comparison as our  $2 \times 2$  factorial. We examine empirically the pitfalls of "horse-racing" policies options, even when high-quality evidence is available.

## 3 Theory and Arguments

### 3.1 Theoretical Framework: On Building Trust

We seek to understand the dynamics of trust between citizens and police (agents of the state) within the context of community policing. Summarizing a wide literature, [Levi and Stoker \(2000\)](#) argue that trust is "relational [between an] individual ... [and] another individual, group, or institution that has the capacity to do her harm or to betray her" and a "judgment ... expected to inspire courses of action" (476).

We conceptualize trust as a two-way relation. While the study of *citizen trust in government* has generated a large research agenda over several decades, we rarely study *government trust in citizens*. There are obvious reasons for this omission. The literature's focus on "trust in government" in general or abstract terms renders the two-way nature of trust difficult to define: what would it mean for "the government" to trust a citizen? Our focus on agents of a specific institution

(here, the police) clarifies how specific agents of the state (here, street-level officers) can trust or distrust citizens in the same way that citizens trust or distrust police.

In this project, we seek to measure citizens' trust in police officers and officers' trust in citizens, and to understand how that trust evolves. Our concept of trust between citizens and police builds on the [Bhattacharya, Devinney, and Pillutla \(1998\)](#) definition of trust as "an expectancy of positive (or nonnegative) outcomes that one can receive based on the expected action of another party in an interaction characterized by uncertainty" (p. 462). This definition subsumes and generalizes the [Levi and Stoker \(2000\)](#) summary. This definition points to two features of our theoretical framework. Trust is defined in the context of an interaction between at least two actors. Here we consider a citizen and a police officer (or groups thereof). The citizen and police officer face uncertainty about the other actor's type. The police officer holds prior beliefs about citizens' propensity to collaborate; the citizen holds prior beliefs about the police officer's work ethic.

These beliefs condition each player's actions. Specifically, citizens choose how much information to relay to police and police decide how much effort to invest in policing the beat. The level of security in the neighborhood is a then function of the level of information provision and effort expended in policing. With this simple framework, we seek to describe baseline beliefs (priors) and equilibrium actions.

Given the emphasis on uncertainty as a definitional component of trust, we focus on understanding the origins of beliefs and how they change. In principle, changes in either player's beliefs may change equilibrium actions and thereby security. We thus evaluate two interventions aimed at inducing players to update their beliefs and change their behavior accordingly. By observing this updating process, we seek to understand (a) how beliefs change; and (b) how updating affects behavior.

We begin by describing the prior beliefs held by citizens and by police. As described in more detail below, we consider two measures of prior beliefs. First, we will examine self-reported attitudinal measures of citizen beliefs about the police. Second, we treat attendance in the first community meeting by both citizens and police as a behavioral measure of prior beliefs of both actors. In particular, we hypothesize that citizen priors may come from (a) personal experiences with crime or policing or (b) neighborhood-level crime risk indicators. Police priors may emerge from (a) interactions with local citizens or (b) patrolling neighborhood-level risk.

Two different interventions allow us to examine updating of beliefs of both police and citizens. A series of three **community policing meetings** ( $T_m$ ) allow for exchange between citizens and police. They allow citizens to gauge police responsiveness to their concerns and to learn about the efforts and responsibilities of the local (beat-level) police officers. A second **security information treatment** ( $T_f$ ) is a leaflet campaign that provides information about the responsibilities of the police and security outcomes. We anticipate that this will enable citizens to update their beliefs but should not influence the beliefs of police. We denote the structure of the theoretical framework and the interventions in a schematic in Figure 1 for clarity.

We assume that both actors update in a Bayesian manner. This structures the hypotheses that we take to the data. Preliminary descriptive analysis of baseline citizen survey data indicates that citizen priors over police quality vary substantially. Under the assumption that the signal of police quality revealed in the intervention falls somewhere between the most extreme prior beliefs about police quality, we do not expect that all citizens update in the same direction. Unconditional estimates of the ITT of either treatment on beliefs, thus, can only provide information about the signal(s) relative to citizens' priors. While this is useful, it does not provide a direct test of our

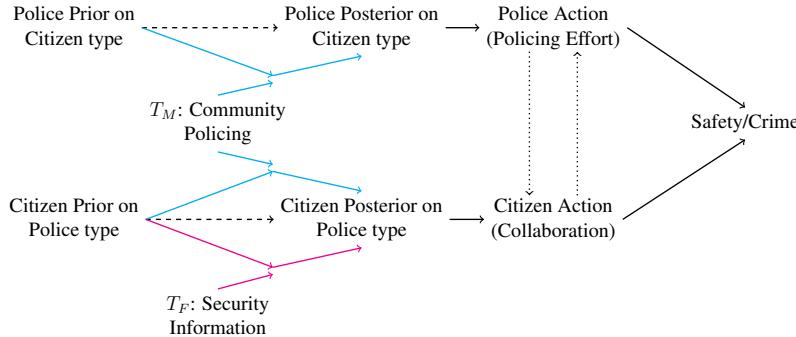


Figure 1: Schematic of theoretical framework. The dashed lines correspond to updating that may occur in the absence of either intervention. The dotted lines denote best response correspondences. The cyan and magenta lines depict the two interventions through which we study belief updating.

argument. We thus also estimate parameters that are more consistent with Bayesian updating. Specifically, we consider treatment effects on the second moment, i.e. the standard deviation, at the quadrante (beat) level, to examine whether citizens' posteriors have lower variance than citizens' priors. We also examine heterogeneity in updating by prior beliefs.

After establishing the effect of the treatments on citizens' beliefs, we then examine one behavioral outcome: 123 calls by citizens (the equivalent of 911). This is a measure of crime reporting, or collaboration with the police. If citizens update positively on police type (propensity to act) because of the meetings or flyers, we expect that, *ceteris paribus*, they will report crime at higher rates. Of course, if the intervention were to affect crime rates themselves, a change in call volume would not necessarily reflect a change in reporting rates—but the intervention was not designed to lower crime rates and we do not expect it to do so.

### 3.2 Theoretical Framework: Gender and Demand for Policing

One immediate observation from initial police-community meetings is that participants are disproportionately (1) female and (2) middle-aged or older. These observations echo research on community health and local infrastructure in Venezuela (Fernandes, 2010; Hanson, 2017, 2018; Zaremburg, 2010) and in Ecuador and Bolivia (Lind, 2002). While previous work provides evidence on differential *response* to police by gender (e.g., Blair, Karim, and Morse, 2019; Cooper, 2018), we know less about differential *demand* for police services or engagement with police.

In this paper, we will evaluate competing explanations for why women (and especially older women) are more likely to participate in police-community meetings. Leveraging 2018 census data and the fact that invitations were distributed to a random sample of citizens, we cleanly measure participation rates. We then evaluate the role of the following factors (informed by the literature and by qualitative observation of initial meetings):

1. Gendered **preferences** on security: Women and men may have different preferences over crime and security. While researchers typically view crime as a valence issue, crime may be more salient for women than men. If this is the case, we expect that baseline assessments of security will correlate with differences in participation in community meetings (and may help explain differential participation by gender).
2. Gendered repertoire of **strategies** to promote security: Just as exclusion from traditional forms of politics may catalyze women's participation in community initiatives (Martin Barbero, 1993), women's exclusion from mechanisms of informal dispute resolution may catalyze their engagement with the police. Reliance on the police should drive desire to engage with officers in police-community meetings. If men have more access to informal means of justice, differences in participation may be highest where rates of insecurity are highest. Similarly, participation in voluntary meetings may substitute for participation in more formal forms of community governance, such as *juntas de acción communal* (JAC). In this case, we would observe higher female attendance at police-community meetings where there were large (baseline) gender differences in JAC participation.
3. Differential **costs** of participation: Meetings require time, typically during afternoon hours. This may have been a barrier to the participation of people working outside the home or outside the neighborhood. To evaluate this hypothesis, we describe how the day and hour of police-community meetings correlates with the composition of participants.
4. Police as a **gendered institution**: Beat-level police officers—the police participants in community meetings—are overwhelmingly young men (< 30 years). Men and women may have different perceptions of officers' attributes. To evaluate this hypothesis, we will describe how officers' characteristics (documented in ethnographic notes) correlate with the composition of participants.

### 3.3 Note: Metaketa Arguments

This study is part of a set of coordinated experiments in six countries, which together comprise the *Metaketa IV* initiative from Evidence in Governance and Politics (EGAP). The meta-analysis of data from the six experiments will test a long list of hypotheses; this list appears in Appendix A.1. Our study tests the shorter list of hypotheses described in Sections 3.1 and 3.2 above, both because (a) these are the hypotheses that follow directly from the questions and arguments we posit as relevant to the Medellín context and because (b) our study is not powered to test additional hypotheses.

## 4 Experimental Design

### 4.1 Treatments

Our intervention comprises two treatments: a *police-community meetings* treatment and an *information* treatment.

**In the police-community meetings treatment,** police officers will hold town-hall-style meetings with residents of their respective police-beats (*cuadrantes*). The meetings will take place approximately bi-monthly over a period of six to eight months, so that residents of each treated *cuadrante* (beat) are invited to three meetings.

The overall intervention will take place between July 2018 and May/June 2019, with some police stations starting and finishing meetings at different points in time.

The research team will invite residents to the meetings via printed flyers distributed door-to-door at residents' homes. The flyers will be distributed within sub-beat areas called a *micro-neighborhood* (see Section 4.2). A pair of enumerators begin in one randomly selected extreme of the micro-neighborhood and work their way around it. If there are fewer than 350 households within the micro-neighborhood, the enumerator team expands delivery to adjacent blocks, as long as those blocks fall within the police beat (*cuadrante*).

The invitations, pictured in Figure 2, introduce our implementing partner, *Estrategia & Territorio*.



Figure 2: Tri-fold flyer for inviting neighborhood to police-community meetings

These meetings differ from police-community dialogues previously held in Medellín. Past police-community meetings generally involved high-ranking officers (such as station chiefs) and were organized at the level of large geographic units called *comunas*. In contrast, the meetings associated with the intervention involve beat officers and are organized at the level of small neighborhoods; this local focus is designed to facilitate relationships among citizens and the agents with

whom they would otherwise interact (that is, the police officers whom they would be likely to encounter outside of the intervention).

Each meeting will begin with a presentation from the police officers about who they are, how the *cuadrante* (beat) operates, and recent activities and plans. The presentation will be followed by an open discussion of local problems. Officers and residents will then develop a *Cooperation Agreement*, in which both officers and residents agree to take specific, concrete actions toward addressing problems in the neighborhood. Subsequent meetings will then begin with a review of the previous meeting's Cooperation Agreement, with the goal of providing a sense of mutual accountability. A team of about 25 local research assistants will assist beat officers in moderating the meetings. And to facilitate ongoing communication, participants will be invited to participate in WhatsApp groups with their neighbors. The full meeting protocol, devised in partnership with the City Government and the police, appears in Appendix A.5.

**In the information treatment,** the research team will distribute information about the police and security, broadly defined. We will deliver three different information flyers twice to each treated *cuadrante* (all occurring in the later half of the implementation, i.e., 2019). The flyers present information about (a) resources for victims of domestic violence as well as resources for reporting; (b) Colombia's new Police Code, which empowers officers to issue fines for behaviors such loud noises, as well as resources for reporting; and (c) comuna-level crime trends.

The information treatment is motivated both by literature documenting (often) inaccurate perceptions of crime rates and of state security institutions (e.g. Ardanaz, Corbacho, and Ruiz-Vega, 2014), and by our own qualitative fieldwork in Medellín, which highlighted minimal familiarity with police responsibilities and achievements. In this regard, the Police Code flyer (Figure 3) emphasizes some of these little-known responsibilities in the context of one of the most recurrent issues: noise complaints. In the same spirit, the flyer in Figure 4 compares of homicide rates in the 2009–2013 period to homicide rates in the 2014–2018 period. We choose these periods because they are informative about the trend, but yet shorter than in previous experiments (c.f. with 10 years in Ardanaz, Corbacho, and Ruiz-Vega, 2014) so to increase salience. We decided to use within-comuna comparisons rather than cross-comuna comparison because existing evidence points to null additional effects (Arias et al., 2019).

Finally, we use a social-norms marketing approach for presenting information about the importance of crime reporting, emphasizing that domestic violence is not a private matter (Figure 5).

## 4.2 Unit of assignment to treatment

The unit of randomization is the *cuadrante* (police beat). The total number of *cuadrantes* in Medellín (at the time of the design) is 413. We exclude from our sample 66 quadrantes that are either (a) located in remote areas of the city or (b) non-residential (e.g., the local airport). Figure 6 illustrates the police beats included in and excluded from our sample.

Within each of the remaining 347 police beats, we define a *micro-neighborhood* (or 'prioritized blocks') as the set of inhabited, contiguous city blocks closest to the centroid of the police beat. Each micro-neighborhood comprises about four blocks, depending on the residential density, so as to ensure similar populations across micro-neighborhoods. When the centroid of the



Figure 3: Tri-fold flyer for with security related information about the Código de Policía and reporting

police beat falls in (for example) a park, we begin the micro-neighborhood at the inhabited block closest to the centroid. Cuadrantes have an average of 5,348 residents (in the 2005 census), our micro-neighborhoods contain approximately 1,200 residents, or about 400 households. Figure 7 illustrates the location of micro-neighborhoods located within a sample of police beats. It is within these micro-neighborhoods that we deliver invitations and informational flyers to 350 households.

**There is thus one micro-neighborhood per police beat; we assign each beat to one of four conditions:** control, meetings only ( $T_m$ ), information only ( $T_f$ ), or meetings and information ( $(T_m)$  and ( $T_f$ )).

**To assign police beats (and thereby micro-neighborhoods) to treatment conditions, we block-randomize.** Each block contains four police beats that (a) belong to the same police station (of which we coded 14)<sup>2</sup> and (b) have the same treatment status (treated or control) in a simultaneous intervention conducted by other researchers. (In other words, we cross-randomize with another

<sup>2</sup>Aranjuez, Belén, Buenos Aires, Candelaria, Castilla, Doce de Octubre, Laureles, Manrique, Poblado, Popular, San Antonio de Prado, San Javier, Santa Cruz, and Villa Hermosa. However, some of the stations are relatively small, and thus we grouped 6 of these into 3, for a total of 11 police station groups. We grouped Aranjuez with Manrique, Buenos Aires with Vila Hermosa, and Popular with Santa Cruz.



Figure 4: Tri-fold flyer for with security related information about homicide trends

trial taking place in Medellín). Within each block, we randomly assign one police beat to each of the four treatment conditions (see Table 1).<sup>3</sup> This simple blocking strategy is sufficient to produce balance on all observable demographic, socioeconomic, and crime characteristics, as shown in Appendix Table A1.

	¬ Leaflets ¬ $T_f$	Leaflets $T_f$
¬ Community Meetings	¬ $T_m$	(Control, $Z_0$ ) $N = 87$
Community Meetings	$T_m$	$Z_m$ $N = 87$

Table 1: Treatment conditions

<sup>3</sup>There is one block of three police beats, each of which we assign to one of the four conditions with equal probability.

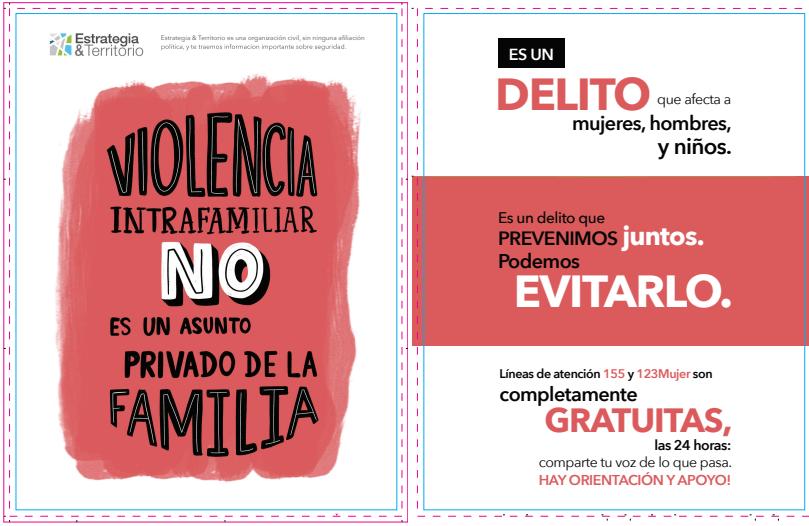


Figure 5: Tri-fold flyer for security related information and social marketing message about domestic violence

## 5 Data and measurement

### 5.1 Overview

Our research design leverages three types of data to test our hypothesis and guide the interpretation of our results. We rely on (1) administrative data on crime reporting, (2) survey data of both citizens and police officers, and (3) qualitative notes, interviews, and observations.

#### 5.1.1 Administrative Data

Our first source of administrative data is census data on population by *cuadrante*. Here, we measure variation in citizen characteristics by *cuadrante* in order to benchmark meeting participation to the eligible population. We will use the 2018 Census data once it becomes available.

We will analyze administrative micro-data on citizen crime reporting. Broadly speaking, we have four different outcomes: reported theft, reported domestic violence, reported public misconduct, and calls to emergency line NUSE 123, akin to 911. We use pre-treatment measures of these variables, in addition to homicide rates, to estimate crime rates in neighborhoods prior to the treatment.

These crime reports are time-stamped and geo-located. All the data is provided by the Infor-

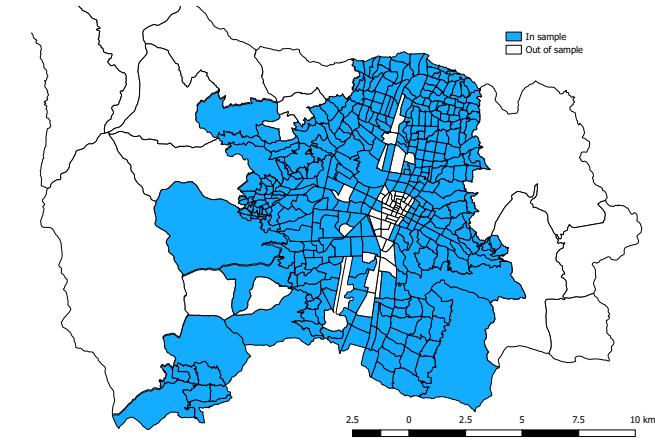


Figure 6: The 347 police beats in our sample

mation Office (SISC) of the Security Secretariat of the City of Medellín. Table 2 describes the crime data.

#### 5.1.2 Survey Data

**Citizens** We conduct baseline and endline surveys of citizens. The baseline will include 15 respondents per *cuadrante*, for a total of 5,205 respondents (one per household). Households are surveyed randomly within micro-neighborhood through a random walk method, with a random starting point.

The expected attrition rate was about 0.3, implying an endline sample of approximately 3,643 (approximately 10 to 11 endline respondents per micro-neighborhood). All surveys are conducted in-person, using tablets, with printed visual aids for scales. Enumerators will introduce themselves as part of INVAMER, a well-known survey firm in Medellín, asking opinions about community life and security. Broadly speaking, the survey covers:

1. Security perception
2. Trust and perception of institutions (focus on police)
3. Crime victimization
4. Community relations and behavior
5. Political attitudes
6. Demographics

Behavior	Information	Source
Theft reported	Type of property stolen (e.g. motor vehicle, cellphone), weapon (e.g. firearm), victim characteristics (e.g. age, gender).	Fiscalía General de la Nación and SIJIN (Seccionales de Investigación Criminal) via SISC
Domestic violence reported	Type of violence (e.g. physical, threat), motive (e.g. jealousy), institution where report is made (e.g. police station, permanencia), victim characteristics (e.g. age, gender).	Subsecretaría de gobierno local y convivencia via SISC
'Public misconduct' ( <i>Conductas contrarias a la convivencia</i> ) reported	<i>Comparendos</i> (subpoena) on fights, brawls, gunfire, public consumption of alcohol, etc. as well as perpetrator characteristics such as age and gender.	NUSE 123 via SISC
NUSE 123 calls	All calls to the NUSE 123 line, with reason for calling (e.g. domestic violence, tips, etc.).	NUSE 123 via SISC
Homicide (covariate only)	Type of weapon used, if any (e.g. firearm), notes on likely motive (e.g. linked to theft, linked to gangs, etc), as well as background information on the victim.	INML (Instituto Nacional de Medicina Legal y Ciencias Forenses), SIJIN (Seccionales de Investigación Criminal) and CTI (Cuerpo Técnico de Investigación, under La Fiscalía General de la Nación) via SISC, Secretaría de Seguridad y Convivencia as technical observer

Table 2: Summary of administrative crime data: outcomes measuring crime and crime reports.

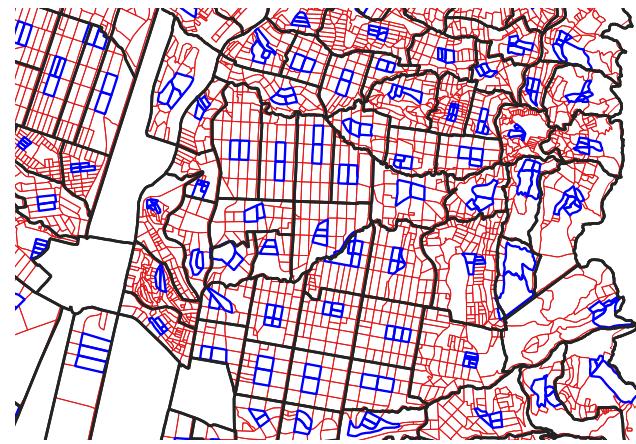


Figure 7: Location of micro-neighborhoods within each police beat

Note that the recontact rate for the endline survey was lower than anticipated ( $\approx 50\%$  across the full sample). As a result, we augmented the endline survey with new subjects such that the endline sample in each *cuadrante* represents 70% of the baseline sample size. For the new respondents in the endline, we do not have baseline measures, so we impute baseline *cuadrante* means and include an indicator for new respondents in any covariate-adjusted analyses.

**Patrol officers** We conduct an endline survey of patrol officers. The target for the endline police survey was two patrol officers per *cuadrante*, for a total sample size of 694. Due to implementation challenges, we were unable to reach that many police officers. The surveys were conducted in *station*-level meetings of patrol officers. There exists some variation in attendance and, as such, the number of police officers per *cuadrante*. Broadly speaking, the survey covers:

1. Information acquisition for policing
2. Police perceptions of citizens
3. Police perceptions of citizens' perceptions about police/behavior toward police
4. Demographics (including experience in the Police)

### 5.1.3 Qualitative Data

Finally, we will collect three types of qualitative data: (1) semi-structured interviews conducted before and after the intervention, (2) observation (i.e., detailed ethnographic notes) of 100% of the police-community meetings, and (3) post-intervention focus groups conducted.

- Semi-structured interviews.** These include citizens, local leaders, and police officers.
- Observation reports of meetings.** Our local enumerators have a strict protocol for delivering reports on every single meeting. Thus, we will have a minimum of 558 reports (some meetings produce two reports, one from each enumerator). The reports include detailed accounts of the meeting interactions, questions, reactions, attitudes and corporal language.
- Focus groups.** We will conduct one focus group in each of 51 quadrantes (51 focus groups total). The quadrantes are randomly sampled within each of three blocks: (i) quadrante assigned to ‘no-meeting’ condition, (ii) quadrante assigned to meeting, with attendance higher than *comuna*-level median attendance, and (iii) quadrante assigned to meeting, with attendance lower than *comuna*-level median attendance.

Across the board, a constant comparison analysis of qualitative data will be used to illuminate the mechanisms underlying any observed treatment effects on community attitudes and/or police perceptions.

## 5.2 Outcome measurement: Trust paper

### 5.2.1 Manipulation checks

Before estimating the effect of the intervention on beliefs or behavior, we compare awareness of the intervention itself across citizens and officers in treatment and control neighborhoods. To do so, we use responses to questions detailed in Table 3.

Outcome	Data Type	Details	Range
Citizens: meetings	Survey	In the past 12 months, do you remember receiving invitations to meetings with police officers from your neighborhood? → If no: In the past 12 months, do you remember hearing about meetings between citizens and police officers from someone in your neighborhood? In the past 12 months have you attended meetings between citizens and police from your neighborhood?	{0,1}
Citizens: leaflets	Survey	In the past 12 months, do you remember receiving flyers with security-related information? → If no: In the past 12 months, do you remember hearing about flyers with security-related information that were distributed in your neighborhood? Comprehension check – questions on whether citizens remember the content of the flyers conditional on having received them?	{0,1}
Police Officers	Survey	In the past 12 months, have you attended meetings with the residents of your beat? → [If yes] How many of these meetings have you attended?	{0,1}

Table 3: Manipulation checks. All outcomes come from the endline citizen and police survey.

### 5.2.2 Citizen Beliefs

The theoretical framework discussed in Section 3 predicts that the interventions will affect citizen beliefs about police officers, officer beliefs about citizens, and the behavior of both citizens and police officers. This section specifies how we measure these beliefs and behaviors.

We use two principal survey-based measures of citizen beliefs: trust and views about police quality, as outlined in Table 4. Naturally, responses in the baseline survey correspond to prior beliefs, and responses in the endline survey correspond to posterior beliefs.

Outcome	Data Type	Details	Range
Trust	Survey	<i>Responses to the question:</i> How much do you trust the following institutions or groups? → The police  → The police officers in your neighborhood.	{1, 2, 3, 4, 5}
Police Quality	Survey	<i>Index of responses to:</i> To what extent do you agree or disagree with each of the following statements? → The police act upon citizen comments and complaints about security in my community. → The police take cases seriously and investigate them. → The police are corrupt. → The police provide the same quality of service to all citizens. → The police have the capacity to respond to incidents of crime in a timely manner. → The police have the capacity to investigate crimes and gather evidence effectively.	Continuous index, $\mu = 0, \sigma = 1$
Other Institutions Quality	Survey	Same as above but with respect to → The District Attorney's office.	Continuous index, $\mu = 0, \sigma = 1$
Police relative quality	Survey	Difference between Police Quality Index and District Attorney Quality Index	Continuous index $\in [-1, 1]$
Security Perception	Survey	<i>Index of responses to</i> the following questions: → Overall, how safe do you feel in your neighborhood? (Level). → With respect to six months ago, how safe do you feel in your neighborhood? (Change). Out of fear, during the last 12 months did you ever ... (Yes/No) → avoid going out alone at night? → avoid certain streets or roads at night? → avoid using public transportation? → avoid new purchases as they could be stolen? → avoid letting children play on the street? → purchase any type of firearm?	Continuous index, $\mu = 0, \sigma = 1$

Table 4: Outcomes measuring citizen beliefs. All outcomes come from the endline citizen survey. The indices are constructed using a *z*-score index.

These are scaled using a *z*-score index, with mean 0, standard deviation 1 (with respect to the control group).

We ask the same trust and quality questions about other government agencies, in addition to

the police. Specifically, we also ask about the district attorney's office; this is important because citizens could conflate the responsibilities of the police and the D.A.'s office. For this same reason, the endline survey will also pose questions that attempt to disentangle citizens' trust *in the specific officers assigned to their neighborhood police beat* from trust *in the police institution as a whole*.

We also study the effect of the intervention on perceptions of security, as described in the bottom of Table 4.

### 5.2.3 Officer Beliefs

Table 5 describes our survey measures in the (endline) police survey.

Outcome	Data Type	Details	Range
Citizens cooperation	Survey	<i>Index of responses to:</i> To what extent do you agree or disagree with each of the following statements? → Citizens are concerned for the well-being of police officers on my beat. → Most of the citizens on my beat are cooperative and respectful. → The information that citizens report is useful.	Continuous index, $\mu = 0, \sigma = 1$

Table 5: Outcomes measuring police beliefs. All outcomes come from the endline police survey. The index is constructed using a *z*-score index.

### 5.2.4 Citizen Behavior

We use both (1) self-reported crime reporting (from our survey data) and (2) administrative data to measure the rate at which citizens report crimes.

Outcome	Data Type	Details	Range
Crime Reporting	Survey	<i>For each of the crimes listed below, we ask victims about their reporting behavior and non-victims about their hypothetical reporting behavior, using the following sequence of questions:</i> Thinking about the last six months, → were you or a member of your household a victim of [crime]? → was a friend or acquaintance in this neighborhood a victim of [crime]? → [If yes] Did the victim report the crime? → [If yes] To whom did the victim report the crime? → [If no] If you or someone you know were a victim of [crime], would you report it, or not?	Continuous index, $\mu = 0, \sigma = 1$
Crime Reporting	Admin. data	<i>Geocoded, time-stamped administrative data to construct two measures of crime reporting at the micro-neighborhood level</i> → Calls to the emergency line NUSE 123 → Reports of theft, domestic violence, and “public misconduct”	Continuous index, $\mu = 0, \sigma = 1$

Table 6: Outcomes measuring citizen reporting. All outcomes come from the endline citizen survey. The index is constructed using a  $z$ -score index.

For self-reported crime reporting based on our survey data, we ask victims about their reporting behavior and non-victims about their hypothetical reporting behavior. We do so with questions outlined in Table 6. Moreover, we do so for (i) crimes not committed by the police and (ii) crimes committed by the police:

- a. Robbery, auto theft, fist fights; *asked only about victimization of neighbors, not about own or own-household victimization*: domestic violence, sexual abuse, homicide
- b. Police physical abuse, police verbal abuse

We then use responses to these questions to construct two indices of self-reported crime reporting behavior:

$$\text{Report}_t^a \equiv \text{Count of actual and hypothetical reports of crimes in group (a)}$$

$$\text{Report}_t^b \equiv \text{Count of actual and hypothetical reports of crimes in group (b)}$$

where the subscript  $t \in \{1, 2\}$  distinguishes baseline from endline measures.

### 5.3 Outcome Measurement: Gender and Demand for Policing

Our primary measures for the first portion of this project come from census data and the sign-up sheets and notes collected by our research assistants during the police–community meetings. We evaluate rates of participation at both the *cuadrante* level (the cluster) as well as the individual (population) level, as enumerated in Table 7.

Outcome	Data Type	Details	Range
Individual participation	Census + meeting logs	Create a long-form dataset including rows for each adult resident of a <i>cuadrante</i> according to gender and age designations, i.e. group $g$ . Make this a panel for meetings $t \in \{1, 2, 3\}$ . Impute a 1 for attendance matching each type of individual in attendance in meeting $t$ .	[0,1]
Community-level participation, by group, $g$	Census + meeting logs	The ratio of: Participants in meeting $t$ belonging to group $g$ / <i>Cuadrante</i> population belonging to group $g$	[0,1]

Table 7: Outcomes measuring rates of participation across subgroups in the population

### 5.4 Outcome Measurement: Metaketa Contribution

See Appendix A.1 for an enumeration of outcome measures used in the meta-analysis and our contribution to the joint product.

In our article, we will also focus on comparability across the two interventions. As a reminder, we care about implementation, uptake, and the effects of each intervention. Conveniently, there were three meetings and three sets of fliers delivered by the field team.

	Community policing	Informational fliers
Flier delivery	<ul style="list-style-type: none"> <li>- Dates of delivery of invitations for meetings 1, 2, and 3</li> <li>- Indicator for whether fliers were delivered</li> </ul>	<ul style="list-style-type: none"> <li>- Dates of delivery of fliers for deliveries 1, 2, 3</li> <li>- Indicator for whether fliers were delivered</li> </ul>
Recall of fliers	<ul style="list-style-type: none"> <li>- Recall receiving at least one invitation to a meeting in survey (binary)</li> </ul>	<ul style="list-style-type: none"> <li>- Recall receiving at least one flier in survey (binary)</li> </ul>
Information processing of fliers		<ul style="list-style-type: none"> <li>- Recall content of fliers: chooses 3 correct topics from list of 4 options (binary)</li> </ul>
Participation in meetings	<ul style="list-style-type: none"> <li>- Attendance at each meeting by citizens from meeting notes (count) [non-experimental]</li> <li>- Attendance at each meeting by police from meeting notes (count) [non-experimental]</li> <li>- Reported attendance by citizens (binary) [experimental]</li> </ul>	<ul style="list-style-type: none"> <li>- Attendance at each meeting by citizens from meeting notes (count) [non-experimental]</li> <li>- Attendance at each meeting by police from meeting notes (count) [non-experimental]</li> <li>- Reported attendance by citizens (binary) [experimental]</li> </ul>

Table 8: Outcomes on implementation and compliance for study in the comparison of two policies.

## 6 Estimation

### 6.1 ITT Estimators

Below, we enumerate our estimators and analysis plans, by project. The factorial design implies two (broad) sets of estimators, both with possible tradeoffs. The main estimator specified in the Meta-PAP estimates:

$$Y_{ijb} = \beta_1 T_{jb}^m + \beta_2 T_{jb}^f + \gamma_b + \epsilon_{ijb} \quad (1)$$

where  $Y_{ijb}$  is an outcome measured at the level of an individual survey respondent in quadrante  $j$  in block  $b$ ;  $T_{jb}^m$  is an indicator for assignment to the police-community meetings treatment;  $T_{jb}^f$  is an indicator for assignment to the flyer treatment; and  $\lambda_b$  is a vector of block fixed effects. In this specification,  $\beta_1$  is the estimator of the ITT of community policing (alone);  $\beta_2$  is the estimator of the ITT of the flyer campaign (alone); and  $\beta_1 + \beta_2 + \beta_3$  is the estimator of the ITT of the combined treatment condition (versus control).

The alternative estimator implied by the factorial design interacts the indicators for each of the treatment arms,  $T_{jb}^m$  and  $T_{jb}^f$ :

$$Y_{ijb} = \beta_1 T_{jb}^m + \beta_2 T_{jb}^f + \beta_3 T_{jb}^m T_{jb}^f + \gamma_b + \epsilon_{ijb} \quad (2)$$

A comparison of the two estimators reveals a bias/power tradeoff. Equation 1 is better powered, but may be “biased” in the presence of strong complementarities between the treatment arms. We do not theorize such complementarities and opt for the better powered estimator for the primary meta-analysis specifications. Please see Appendix A.1 for additional details on the meta-analysis estimators.

## 6.2 Estimation: Trust Paper

Section 3 motivated our two principal hypotheses: that police-community interaction and/or information about police performance affect (a) trust and perceptions surrounding the police and (b) crime reporting. This section describes how we evaluate these hypotheses.

### 6.2.1 Baseline Trust Levels: Prior Beliefs

Because our theory focuses on updating of beliefs, characterizing prior beliefs is important. We use two sources of data to characterize prior beliefs. First, we leverage pre-treatment qualitative data. This includes 69 semi-structured baseline interviews, including 12 police officers, 50 community leaders, and 7 local outreach coordinators from the mayor’s office. We also convened a focus group of citizens. Because we do not have baseline survey data for the police, we use qualitative data to aid in understanding the quantitative findings from the endline survey of police officers.

Second, we use the baseline survey of citizens. We characterize how citizens’ priors covary with three sets of predictors:

- Cuadrante-level measures of crime and policing, including measures of crime and gang presence from administrative data
- Individual demographic characteristics: age, gender, social class (*estrato*), and education
- Individual experiences with crime and the police as recorded in the baseline survey, including crime victimization and reporting.

Finally, we estimate the first-stage ATE on citizen attendance using the survey data. Specifically, we estimate:

$$\text{Attended a meeting}_{ijb} = \beta_0 + \beta_1 T_{jb}^m + \gamma_b + \epsilon_{ijb} \quad (3)$$

$$\text{Attended a meeting}_{ijb} = \beta_0 + \beta_1 T_{jb}^m + \delta \mathbf{X}_{ijb} + \kappa T_{jb}^m \mathbf{X}_{ijb} + \gamma_b + \epsilon_{ijb} \quad (4)$$

where  $\text{Attended a meeting}_{ijb}$  indicates whether citizen  $i$  in quadrante  $j$  in block  $b$  attended a police-community meeting during the intervention. We expect  $\beta_1 > 0$  in Equation 3.

In Equation 4, we explore heterogeneity by citizen and *cuadrante* characteristics  $\mathbf{X}_{ijb}$ . We will select the covariates  $\mathbf{X}_{ijb}$  for inclusion by Lasso with cross-validation, as implemented in the *glmnet* R package. The loss function in Lasso requires specification of a tuning parameter,  $\lambda$ , that sets the penalty for the inclusion of additional covariates. We cannot pre-specify  $\lambda$ , but we will use 10-fold cross validation on all pre-treatment covariates collected (at the individual and *cuadrante* level) to select the  $\lambda$  that minimizes cross-validation error for each outcome.

We plan to study how attendance evolves over the three meetings. In particular, we will estimate how citizen and police attendance at the first meeting ( $a_c \in \{0, 1, \dots, 61\}$ ,  $a_p \in \{0, 1, 2\}$ ) predicts attendance at later meetings. Because flyers were distributed to 350 households in all quadrantes (beats), we can meaningfully compare these participation counts regardless of population size.

### 6.2.2 Does the intervention lead people to update their beliefs?

We aim to estimate the effects of assignment to treatment along each margin, namely:

$$ITT_m = E[\text{Citizen beliefs}|T_m = 1] - E[\text{Citizen beliefs}|T_m = 0]$$

$$ITT_f = E[\text{Citizen beliefs}|T_f = 1] - E[\text{Citizen beliefs}|T_f = 0]$$

These estimands ( $ITT_m$  and  $ITT_f$ ) are ITTs; they capture the average effect of meetings and the average effect of the information treatment (averaging over the rows or columns in the  $2 \times 2$  factorial design). Our baseline specification is:

$$\text{Beliefs}_{ijb}^{t=1} = \delta \text{Beliefs}_{ijb}^{t=0} + \beta_1 T_{jb}^m + \beta_2 T_{jb}^f + \gamma_b + \epsilon_{ijb} \quad (5)$$

where, as above,  $i$  indexes individual survey respondents,  $j$  indexes *cuadrantes*, and  $t$  indexes the survey round (1 indicates endline, 0 indicates baseline).  $\gamma_b$  is a vector of block fixed effects.

We will also estimate a version of Equation 7 that adjusts for covariates. When using covariate adjustment, we will demean our covariates  $\mathbf{X}_{ijb}$  and fully interact them with our treatment indicators to produce unbiased and consistent average effects even in the presence of heterogeneity along these covariates (Lin, 2013; Imbens and Rubin, 2015). We will select covariates using Lasso, as described above in Section 6.2.1.

We also allow for the possibility that the two treatments function as complements or as substitutes, though our power calculations (Section 7) indicate that we will only be able to test for complementarities (or substitutabilities) if there is a high degree of correlation between baseline and endline outcomes. In the event of high correlation between baseline and endline outcomes (which we do not expect), we estimate:

$$\text{Beliefs}_{ijb}^{t=1} = \beta_1 T_{jb}^m + \beta_2 T_{jb}^f + \beta_3 T_{jb}^m T_{jb}^f + \gamma_b + \delta \text{Beliefs}_{ijb}^{t=0} + \epsilon_{ijb}^{t=1} \quad (6)$$

Rejection of the null hypothesis that  $\beta_3 = 0$  would indicate that the two treatments are either substitutes or complements among compliers (depending on the direction of the difference).

While we estimate Equations 6 and 7 for the sake of comparison with estimates for other studies, we do not consider them our primary specifications of interest. Because our theoretical framework focuses on updating, we expect the direction of updating to depend on an individual's prior and the signal (Arias et al., 2018). For this reason, we first interact prior beliefs with the treatment to look at conditional updating as a function of the priors:

$$\text{Beliefs}_{ijb}^{t=1} = \delta \text{Beliefs}_{ijb}^{t=0} + \beta_1 T_{jb}^m + \psi_1 T_{jb}^m \text{Beliefs}_{ijb}^{t=0} + \beta_2 T_{jb}^f + \psi_2 T_{jb}^f \text{Beliefs}_{ijb}^{t=0} + \gamma_b + \epsilon_{ijb}^{t=1} \quad (7)$$

For robustness, we also (1) discretize  $\text{Beliefs}_{ijb}^{t=0}$  into terciles and estimate separate interactions for each tercile, and (2) create a graph of the nonparametric relationship between posteriors ( $\text{Beliefs}_{ijb}^{t=1}$ ) and priors, separately for treatment and control groups.

We also estimate treatment effects on the second moment of the distribution of beliefs. To do so, we estimate Std. Dev.  $\text{Beliefs}_{jb}^{t=1}$ , the standard deviation of posteriors within quadrante  $j$  in block  $b$ , and estimate:

$$\text{Std. Dev. Beliefs}_{jb}^{t=1} = \delta \text{Std. Dev. Beliefs}_{jb}^{t=0} + \beta_1 T_{jb}^m + \beta_2 T_{jb}^f + \gamma_b + \epsilon_{jb} \quad (8)$$

Alongside this quantitative analysis, our post-intervention focus groups and detailed meeting notes will qualitatively illustrate the extent to which citizens acted and reacted positively or negatively during the meetings; similarly, the meeting notes together and post-intervention interviews with police officers will qualitatively describe police behavior and reactions.

### 6.2.3 Does the intervention lead to different police beliefs about citizens?

Because we do not have a baseline survey of police, our quantitative analysis of officer beliefs relies entirely on the endline survey. We use this survey to estimate (average) differences in endline beliefs of officers who were assigned to attend police-community meetings, compared to officers not assigned to meetings. (We do not anticipate an effect of the flyers on police beliefs.) We will estimate:

$$\text{OfficerBeliefs}_{ijb} = \beta_0 + \beta_1 T_{jb}^m + \beta_2 T_{jb}^f + \gamma_b + \epsilon_{ijb} \quad (9)$$

with and without covariate adjustment, using the covariate selection procedure described above. We cluster standard errors at the *cuadrante* level. Exploratory tests may allow us to assess differential updating among different types of officers (e.g. experience/rank), quadrante (e.g. level of crime/insecurity), or priors, as proxied by attendance at the first meeting.

Note that there were several complications in the administration and design of the police survey. Specifically:

- We were unable to sample the prespecified 694 officers (2 per *cuadrante*). The surveys were administered at station meetings of beat-level patrol officers with permission from each station commander. The main constraint was attendance at these meetings.
- Some police officers were not willing to provide their *cuadrante* number such that we cannot determine their assignment to treatment or control.

The *experimental* comparisons using police survey data will use the following procedure:

1. Condition the sample on those officers that reported their *cuadrante* number. Analyze the data for differential survey response by regressing a count of completed surveys with *cuadrante* identification from each *cuadrante* on treatment assignment using the ITT estimator above. Rejection of the null hypothesis on meetings constitutes grounds for police survey outcome exclusion from the meta-analysis.
2. Manipulation check. Given the issues with the design, we will establish a first stage in the police survey between assignment to meetings and participation in meetings. To do so, we regress the answers to the compliance questions #10 and an indicator for an affirmative answer to #10 and/or #11 on Equation 9.
3. Analyses. Further experimental analyses will be estimated on equation 9 with standard errors clustered at the *cuadrante* level.

We will also examine observationally the associations between self-reported participation in community policing meetings (measured in questions #10 and #11 on the police survey) and beliefs about citizens so that we can use the full survey sample (even when *cuadrante* is not reported). We will delineate such comparisons from the experimental analysis.

### 6.2.4 Does the intervention change crime reporting behavior?

To understand the extent to which our interventions affected crime reporting, we estimate:

$$\begin{aligned} ITT_m &= E(\text{Crime reporting}|T_m) - E(\text{Crime reporting}|\neg T_m) \\ ITT_f &= E(\text{Crime reporting}|T_f) - E(\text{Crime reporting}|\neg T_f) \end{aligned}$$

To estimate these parameters, we first use survey data to create measures of  $\text{Reporting}_{ijb}^t$  at baseline and endline. We then re-estimate Equation 3, replacing  $\text{Beliefs}_{ijb}^t$  with  $\text{Reporting}_{ijb}^t$ .

Second, we use administrative data, aggregated to the *cuadrante*-day level. Specifically, we estimate

$$\text{Reporting}_{jbt} = \beta_1 T_j^m + \beta_2 T_j^f + \gamma_b + \kappa_t + \epsilon_{jbt} \quad (10)$$

where  $\text{Reporting}_{jbt}$  is a measure of crime reports from quadrante  $j$  in block  $b$  on day  $t$  (aggregated from administrative microdata; see Section 5),  $\kappa_t$  is a vector of day fixed effects, and all other variables are as defined above. We cluster standard errors by quadrante. When estimating Equation 10, we weight the observations by the share of households in each quadrante that received leaflets. (For control quadrantes, we use the share of households receiving leaflets in the average treated quadrante in the same block). This appropriately downweights large quadrantes in which only a small fraction of households received leaflets.

While Equation 10 provides a useful first cut, it does not leverage the fact that the timing of flyer delivery and police-community meetings was staggered (though not randomized) across neighborhoods. Because we would expect behavior to change only *after* the beginning of treatment, we define  $\text{Post}_{jbt}^m$ , an indicator taking a value of one after the first community meeting in a

cadrante assigned to the meetings treatment; and  $Post_{jbt}^f$ , an indicator taking a value of one after the first flyer distribution in a cadrante assigned to the flyers treatment; in order to estimate:

$$\text{Reporting}_{jbt} = \beta_1 T_{jb}^m + \theta_1 Post_{jbt}^m + \beta_2 T_{jb}^f + \theta_2 Post_{jbt}^f + \gamma_b + \kappa_t + \epsilon_{jbt} \quad (11)$$

We will test the null hypotheses that  $\theta_1 = 0$  and that  $\theta_2 = 0$ . We can examine anticipation by testing the hypotheses  $\beta_1 = 0$  and  $\beta_2 = 0$ .

Because the distribution of flyers and invitations was targeted to households in the center (*manzana*) of each *cadrante*, we will also conduct the analysis for reporting originating in each *manzana*. The logic for this is straightforward. If exposure to treatment is higher in the *manzana* than in the area outside of it, we may attenuate estimates by “averaging” over occurrences in the whole *cadrante*.

It might also be the case that effects of the intervention on reporting behavior are short-lived. To consider this possibility, we re-estimate Equation 11 thirty times, each time re-defining  $Post_{jbt}$  as an indicator for the first  $\{1, 2, \dots, 30\}$  days after flyer distribution began or after a meeting was held. We then plot the thirty coefficients on  $Post_{jbt}$ , with the objective of observing whether effects decay over time. This will allow us to more flexibly examine empirical patterns of reporting in response to the treatments.

### 6.3 Estimation for Gender and Demand for Policing Paper

To estimate participation rates by gender, we estimate:

$$\text{Attendance}_{ijt} = \beta_0 + \beta_1 \text{Female}_i + \beta_2 \text{Age group}_i + \epsilon_{ijt} \quad (12)$$

$$\text{Prop. Attendance}_{ij} = \beta_0 + \beta_1 \text{Female}_j + \beta_2 \text{Age group}_j + \epsilon_{ij} \quad (13)$$

These specifications will estimate differences in participation by age and gender. In particular  $\beta_1$  tests our hypothesis on gender;  $\beta_2$  tests our hypothesis on age. Equation 12 estimates differences in rates of attendance at the individual level; Equation 13 estimates differences in rates of attendance at the community level (effectively downweighting people in large *cadrantes*). To describe which community-level factors affect differential participation by gender or age, we interact  $\text{Female}_i$  and  $\text{Age group}_i$  with covariates such as the pre-intervention crime rate. We select covariates using Lasso, as described above.

We will also use variants of Equation 12 to evaluate the hypotheses laid out in Section 3.2. In particular, in Equation 12 we will interact  $\text{Female}_i$  and  $\text{Age group}_i$  with:

1. Baseline assessments of beliefs about safety, as measured in the *security perception* index of Table 4.
2. The proportion of the local *Junta de Acción Comunal* (JAC) that was female prior to the start of the intervention.
3. Respondents’ employment status.
4. The *age* of officers assigned to that *cadrante* and the length of their tenure there.

### 6.4 Estimation for Metaketa Contribution

See Appendix A.1 for details on our estimators for use in the meta-analysis. We defer to the meta-analysis pre-analysis plan (Blair, Christia, and Weinstein, 2020) in the case of any conflicts between these documents or lack of clarity.

For the comparison of policies, we will use the same estimators as in the meta-analysis whenever experimental comparisons are feasible. When we examine patterns of treatment implementation from the qualitative data, we will assess rates within the treatment group. For example, we cannot directly measure meeting attendance in control (where meetings were not held). Any reporting on this form of uptake of the treatment is necessarily non-experimental but informs our understanding of the relative efficacy of each treatment.

## 7 Power

We use simulation together with our baseline data to estimate our statistical power for testing hypotheses about citizens’ trust in the police. The outcome, described in Section 5.2, is an *z*-score index of responses to questions about several outcomes. We focus on citizen beliefs for the purpose of this simulation. To do so, we begin the simulation by calculating this index in our baseline data.<sup>4</sup> Denote the baseline value of this index as  $\text{Beliefs}_{ijb}^{t=0}$ . We merge this data with the treatment assignment.

We simulate outcomes of the index at endline, denoted  $\text{Beliefs}_{ij}^{t=1}$ , under various assumptions about the data generating process. Note that the intraclass correlation (ICC) of the baseline index is low (0.027). The following simulated data generating process uses the actual treatment indicators and randomly draws individual-level treatment effects. This DGP roughly preserves the low ICC at small effect sizes, including the magnitudes of those simulated here.

$$\text{Beliefs}_{ij}^{t=1} = \begin{cases} \text{Beliefs}_{ij}^{t=0} + \phi_i + \tau_i^m T_{ij}^m + \tau_i^f T_{ij}^f + \tau_i^{mf} T_{ij}^f T_{ij}^m & \text{if } D_i = 0 \\ \text{missing} & \text{if } D_i = 1 \end{cases}$$

$$D_i \sim \text{Bernoulli}(p)$$

$$\phi_i \sim \mathcal{N}(0, \sigma)$$

$$\tau_i^m \sim \mathcal{N}(\mu_m, \sigma_m)$$

$$\tau_i^f \sim \mathcal{N}(\mu_f, \sigma_f)$$

$$\tau_i^{mf} \sim \mathcal{N}(\mu_{mf}, \sigma_{mf})$$

Here  $p$  is the probability that an individual is not located at endline (the attrition rate). We assume that attrition is independent of treatment assignment and of geographic location (*cadrante*).  $D_i$  is thus an indicator for endline missingness. Where we relocate a respondent, we measure  $\text{Beliefs}_{ij}^{t=1}$ . This is a function of the baseline trust index  $\text{Beliefs}_{ij}^{t=0}$ , noise ( $\phi_i$ ), treatment assignment  $\tau_i$ , and individual treatment effects ( $\tau_i^m$ ,  $\tau_i^f$ , and  $\tau_i^{mf}$ ). We vary eight underlying parameters to assess the implications for statistical power:  $p$ ,  $\sigma$ ,  $\mu_m$ ,  $\sigma_m$ ,  $\mu_f$ ,  $\sigma_f$ ,  $\mu_{mf}$ , and  $\sigma_{mf}$ .

To illustrate, we examine the power of the two proposed estimators in Equations 4 (denoted “margins”) and 6 (denoted “with interaction”). The covariate adjustment set includes only decile

<sup>4</sup>Where data is missing, we impute the *cadrante* median.

bins of the lagged (baseline) dependent variable. We estimate using OLS with standard errors clustered at the level of the *cuadrante* (police beat).

Given the number of variants of the data generating process, we proceed in three steps. First, we examine the implications of the attrition rate ( $p$ ) and serial correlation of the untreated potential outcomes ( $\sigma$ ), holding treatment effects fixed. Second, we look at the implications of varying the magnitude of the treatment effects, holding attrition and serial correlation fixed. Finally, we discuss the implications of varying the other parameters.

### 7.1 Attrition rate and serial correlation

Serial correlation in this context is a function of both measurement error and how individuals' trust varies over time. We examine the power of the experiment to detect constant standardized treatment effects of .1 of both factors with no complementarities ( $\mu_m = .1, \mu_f = .1, \mu_{mf} = 0, \sigma_m = 0, \sigma_f = 0, \sigma_{mf} = 0$ ) across a grid of  $p \in \{0.2, .0.3, 0.4\}$  and  $\sigma \in \{1, 1.5, 2.25\}$ .<sup>5</sup>

The results in Table 9 suggest that for small effect sizes, the power of the design depends crucially on the correlation between the baseline and endline outcomes. If this correlation is high (e.g.  $> 0.7$ ), we are able to detect very small standardized treatment effects. The power losses from attrition (within the simulated  $p$ 's) are much more modest. Our survey firm has advised us to expect a survey attrition rate of around 30%.

Serial correlation	Attrition rate	Power				
		Marginal Effect Estimator		Interaction effect Estimator		
		$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_3$
$\sigma$	$\approx$ Cor.	$p$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$
1	0.7	0.2	0.94	0.96	0.69	0.74
1	0.7	0.3	0.76	0.79	0.52	0.53
1	0.7	0.4	0.57	0.58	0.33	0.35
1.35	0.6	0.2	0.92	0.92	0.66	0.68
1.35	0.6	0.3	0.74	0.79	0.45	0.47
1.35	0.6	0.4	0.58	0.56	0.30	0.33
1.75	0.5	0.2	0.88	0.90	0.62	0.64
1.75	0.5	0.3	0.66	0.71	0.42	0.45
1.75	0.5	0.4	0.50	0.56	0.27	0.32
Implied standardized effect size		0.15	0.15	0.1	0.1	0

Table 9: This table estimates the power of the design under different levels of serial correlation in the control group outcomes (denoted in columns 1-2) and levels of attrition in the survey (denoted in column 3). We assume  $\mu_m = .1, \mu_f = .1, \mu_{mf} = 0, \sigma_m = 0, \sigma_f = 0, \sigma_{mf} = 0$ . Each cell represents 700 simulations. Imbalances in the power in columns 4-5 and 6-7 come from both the assignment and from simulation error. Hypothesis tests are one-tailed

<sup>5</sup>Under these  $\sigma$  values  $Cor(Y_{ij}^B, Y_{ij}^E | T_{ij}^m = 0, T_{ij}^f = 0) \in \{0.7, 0.6, 0.5\}$ , respectively.

### 7.2 Effect Size

Here we fix  $\sigma = 1.35$  (serial correlation  $\approx 0.6$ ) and  $p = 0.3$  at expected levels. We further set  $\mu_{mf} = 0$  (no complementarities);  $\sigma_m = 0, \sigma_f = 0$ , and  $\sigma_{mf} = 0$  (constant treatment effects); while varying  $\mu_f$  and  $\mu_m$ , the principal treatment effects of interest. Table 10 suggests that at expected levels of attrition and serial correlation, the design is well-powered to detect effects of 0.2 baseline standard deviations. As anticipated, the marginal effect estimator is better powered than the three-arm estimator.

		Power				
		Marginal Effect Estimator		Interaction Effect Estimator		
$\mu_m$	$\mu_f$	$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_3$
0.1	0.1	0.71	0.78	0.44	0.48	0.07
0.1	0.15	0.69	0.98	0.43	0.80	0.06
0.1	0.2	0.73	0.99	0.43	0.93	0.05
0.15	0.1	0.96	0.77	0.76	0.49	0.03
0.15	0.15	0.97	0.97	0.76	0.79	0.05
0.15	0.2	0.96	1	0.76	0.92	0.07
0.2	0.1	1.00	0.79	0.93	0.47	0.06
0.2	0.15	1.00	0.97	0.91	0.79	0.05
0.2	0.2	1.00	1.00	0.94	0.94	0.05
Implied standardized effect size		$\frac{2\mu_m + \mu_f}{2}$	$\frac{2\mu_f + \mu_m}{2}$	$\mu_m$	$\mu_f$	0

Table 10: We assume  $p = 0.3, \sigma = 1.35, \mu_{mf} = 0, \sigma_m = 0, \sigma_f = 0, \sigma_{mf} = 0$ . Each cell estimated from 700 simulation runs. Hypothesis tests are one-tailed.

Finally, we simulate a non-zero interaction effect in Table 11, demonstrating that we are likely severely underpowered to detect reasonable interaction effects between the two arms.

		Power				
		Marginal Effect Estimator		Interaction Effect Estimator		
$\mu_{mf}$		$\beta_1$	$\beta_2$	$\beta_1$	$\beta_2$	$\beta_3$
0.1		0.29	0.38	0.04	0.06	0.34
0.2		0.72	0.76	0.06	0.07	0.71
0.3		0.96	0.94	0.05	0.06	0.96
0.4		1.00	1.00	0.05	0.06	1.00
Implied standardized effect size		$\frac{\mu_{mf}}{2}$	$\frac{\mu_{mf}}{2}$	0	0	$\mu_{mf}$

Table 11: We assume  $p = 0.3, \sigma = 1.35, \mu_m = 0, \mu_m = 0, \sigma_m = 0, \sigma_f = 0, \sigma_{mf} = 0$ . Each cell estimated from 700 simulation runs. Hypothesis tests are one-tailed.

While we omit further simulations, increasing the noise on estimated treatment effects (e.g.  $\sigma_m$  and  $\sigma_f$ ) implies very small reductions in the power of the design. Introducing complementarities

between the treatments ( $\tau_{mf} > 0$ ) increases power predictably for both estimators. The data generating process simulated here may be helpful in ex-post power calculations.

## 8 Ethics

### 8.1 Risk / Benefit Assessment

**Potential risks to participants.** Like all studies that use surveys and interviews methods, ours entails the risk that certain questions will make respondents feel uncomfortable. For example, asking them whether they trust the police, or whether they would cooperate with a police investigation if they were able to provide relevant information, may make certain respondents feel uneasy—especially if the answers to these questions are “no.” Likewise, questions about past crime victimization (“Have you been the victim of a robbery?”) may bring up unpleasant memories. Similarly, asking police officers about their perceptions of citizens may make them feel uncomfortable.

We seek to minimize these risks by, first, emphasizing to all interviewees that their participation is voluntary and that they may end the interview at any time. These risks are commensurate with those already faced by the target population: academic and other researchers frequently field surveys in Medellín, including about crime and security topics. For example, EAFIT University fields a Victimization and Perceptions Survey (Encuesta de Percepción y Victimización) that includes questions very similar to those of our survey instrument.

In addition to the risk of discomfort experienced by interviewees, there are risks associated with the intervention itself. While the objective of bringing officers, citizens, and city officials together in a town-hall format is to facilitate open and constructive conversation, there is also the possibility that these conversations become contentious and conflictive. Like the survey-related risks, however, these risks are commensurate with those already faced by the target population. The City of Medellín together with the police have held town-hall-style meetings with citizens in the past, with largely positive results. These meetings have disproportionately been attended by women and the elderly.

**Potential Study Benefits.** We expect that the interventions—both the police-community meetings and the provision of information about crime trends—will have direct benefits for subjects assigned to the treatment groups. In particular, we expect that the meetings will improve trust between citizens and the police and thereby police-community relations. Moreover, in increasing citizen awareness of mechanisms for communicating with the police and reporting crime (such as an existing cell-phone app called Seguridad en Línea), the intervention may increase access to the police for those residents assigned to treatment. For the officers, participation in the intervention may improve their awareness of and understanding of the communities they serve. The information treatment also entails benefits for participants: a more accurate view of crime trends may well increase citizens’ feelings of safety in their own communities.

If our randomized controlled trial were to find evidence of these benefits, the City of Medellín might well extend the program to control areas, thereby bringing the benefits to more citizens.

For the research community and for policymakers outside Medellin, the study will improve our understanding of the effects of police-community meetings and of information provision on citizen and police attitudes, and on citizen reporting behavior. Given that dozens of cities across

the Americas have spent scarce resources on such interventions without any experimental evidence as to their effectiveness, this study would provide highly relevant inputs to active policymaking processes.

## References

- Acemoglu, Daron, and Matthew Jackson. 2017. “Social Norms and the Enforcement of Laws.” *Journal of the European Economic Association*.
- Akerlof, George, and Janet Yellen. 1994. Gang Behavior, Law Enforcement, and Community Values. In *Values and Public Policy*, ed. Henry Aaron, Thomas Mann, and Timothy Taylor.
- Ardanaz, Martin, Ana Corbacho, and Mauricio Ruiz-Vega. 2014. “Mind the Gap: Bridging the Perception and Reality of Crime Rates with Information.” IDB Working Paper Series No. IDB-WP-530.
- Arias, Eric, Horacio Larreguy, John Marshall, and Pablo Querubin. 2018. “Priors rule: When do malfeasance revelations help and hurt incumbent parties?” NBER Working Paper 24888.
- Arias, Eric, Horacio Larreguy, John Marshall, and Pablo Querubin. 2019. “Does the Content and Mode of Delivery of Information Matter for Electoral Accountability? Evidence from a Field Experiment in Mexico.” Working Paper.
- Bhattacharya, Rajeev, Timothy M. Devinney, and Madan M. Pillutla. 1998. “A Formal Model of Trust Based on Outcomes.” *The Academy of Management Review* 23 (3): 459–472.
- Blair, Graeme, Fotini Christia, and Jeremy Weinstein. 2020. “Meta-Analysis Pre-Analysis Plan: Community Policing Metaketa Project, version 3.” Available at <http://egap.org/registration/5154>.
- Blair, Robert A., Sabrina M. Karim, and Benjamin S. Morse. 2019. “Establishing the Rule of Law in Weak and War-torn States: Evidence from a Field Experiment with the Liberian National Police.” *American Political Science Review* Forthcoming: 1–17.
- Cooper, Jasper. 2018. “State Capacity and Gender Inequality: Experimental Evidence from Papua New Guinea.” Working Paper.
- Fernandes, Sujatha. 2010. *Gender and Populism in Latin America: Passionate Politics*. University Park, PA: Penn State University Press chapter Gender, Popular Participation, and the State in Chávez’s Venezuela.
- Hanson, Rebecca. 2017. Civilian Policing, Socialist Revolution, and Violent Pluralism in Venezuela PhD thesis University of Georgia.
- Hanson, Rebecca. 2018. “Deepening Distrust: Why Participatory Experiments Aren’t Always Good for Democracy.” *The Sociological Quarterly* 59 (1): 145–147.

- Home Office. 2008. "From the neighbourhood to the national: policing our communities together." *Green Paper*.
- Imbens, Guido, and Donald Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.
- Levi, Margaret, and Laura Stoker. 2000. "Political trust and trustworthiness." *Annual Review of Political Science*.
- Lin, Winston. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *Annals of Applied Statistics* 7: 295–318.
- Lind, Amy. 2002. "Making Feminist Sense of Neoliberalism: The Institutionalization of Women's Struggles for Survival in Ecuador and Bolivia." *Journal of Developing Societies* 18: 228–258.
- Martin Barbero, Jesus. 1993. "Latin America: Cultures in the Communication Media." *Journal of Communication* 43 (2): 18–30.
- Skogan, Wesley. 2006. *Justice and Safety in America's Immigrant Communities*. Princeton University chapter Community Policing and the New Immigrants: Latinos in Chicago.
- Skogan, Wesley, and Susan Hartnett. 1999. *Community Policing, Chicago Style*. Oxford University Press.
- Sunshine, Jason, and Tom Tyler. 2003. "The role of procedural justice and legitimacy in shaping public support for policing." *Law & Society Review*.
- Tyler, Tom. 2005. "Policing in Black and White: Ethnic Group Differences in Trust and Confidence in the Police." *Police Quarterly*.
- Wunsch, Daniela, and Katrin Hohl. 2009. "Evidencing a 'Good Practice Model' of Police Communication: The Impact of Local Policing Newsletters on Public Confidence." *Policing: A Journal of Policy and Practice*.
- Zaremba, Giseal. 2010. "Género verus pueblo?: Movilización, cooptación y participación en Venezuela, Brasil, y Nicaragua." *Latin American Research Review* 51 (1): 84–108.

## A Appendix

### A.1 Complete list of Metaketa hypotheses

As noted in Section 3, our study is part of a set of coordinated experiments in six countries. These experiments together comprise the *Metaketa IV* initiative from Experiments in Governance and Politics. Note that due to the wording and construction of indices, we do not necessarily expect effects on all outcomes in the Colombian context. The meta-analysis of data from these experiments will test a larger set of hypotheses, reproduced from [Blair, Christia, and Weinstein \(2020\)](#) below.

#### Primary Outcome Family 1: Security of Life and Property

- 1a. Negative effect on incidence of crime<sup>†</sup>
  - Note that in our case, the survey measures of crime incidence are composed of binary indicators, not counts of crime occurrence as in the other sites. This is clarified in the table of outcome measures.
- 1b. Positive effect on perceptions of safety (personal, land, and possessions)

#### Primary Outcome Family 2: Citizen Perceptions of the Police

2. Positive effect on citizen perceptions of police<sup>6</sup>

#### Primary Outcome Family 3: Police Perceptions of and Behaviors Toward Citizens

- 3a. Positive effect on perceptions of police empathy, accountability, and abuse and taking corruption seriously.
- 3b. Negative effect reporting of police abuse and bribery<sup>†</sup>

#### Primary Outcome Family 4: Behavioral Cooperation of Citizens with the Police

- 4a. Positive effect on reporting of crime victimization
- 4b. Positive effect on reporting of crime prevention tips
- 4c. Positive effect on reporting of victimization by the police

#### Mechanism Family 1: Perceived Costs to Citizens Cooperating with the Police

- M1a. Positive effect on beliefs about police intentions
- M1b. Positive effect on knowledge of criminal justice system

---

<sup>6</sup>The Meta-PAP does not posit that citizens learn in a Bayesian manner. If Bayesian learning does occur, we may expect people to update in different directions, based on the relation between the signal and prior, which may undermine ability to detect changes.

M1c. Positive effect on norms of citizens cooperation with police

**Mechanism Family 2:** Perceived Returns to Citizens Cooperating with the Police

M2a. Positive effect on beliefs about police capacity

M2b. Positive effect on perceptions of responsiveness to citizen feedback

**Secondary Outcome Family 1:** Trust in the State

S1. Positive effect on trust in the state

**Secondary Outcome Family 2:** Communal Trust

S2. Positive effect on communal trust

**Compliance with Treatment:** Citizen Interactions with Police

C. Positive effect on rate of citizen interactions with police

## A.2 Metaketa Estimators

The estimators that will be employed in the meta analysis are as follows. For survey-based outcomes of police or citizens, we use an estimator analogous to Equation 3, where  $Y_{ijb}^{t=1}$  is the survey outcome of interest for individual  $i$  in *cuadrante*  $j$  in block  $b$  at endline ( $t = 1$ ). For the meta analysis, will weight by the inverse of the sampling probability of citizens in a *cuadrante*. The relevant estimator for the meta-analysis is  $\nu_m$ :

$$Y_{ijb}^{t=1} = \nu_m T_{ij}^m + \nu_f T_{ij}^f + \gamma_b + \delta Y_{ijb}^{t=0} + u_{ijb}^t \quad (\text{A1})$$

where  $T_{ij}^m$  is the community meeting treatment indicator,  $\nu_f T_{ij}^f$  is the alternative treatment arm treatment indicator, and  $\gamma_b$  is a vector of block fixed effects.  $Y_{ijb}^{t=0}$  is the baseline measurement of the outcome  $Y_{ijb}^t$ , where present. We do not have baseline measurements for any officer survey outcomes (Hypothesis 3 outcomes). Further, we do not have baseline measurements for the outcomes: *satis\_general*, *bribe\_freq*, *bribe\_amt*, *know\_law\_suspect*, *know\_law\_lawyer*, *know\_law\_fees*, *know\_law\_vaw*, *know\_report\_followup*, *know\_report\_station*, *obeynorm*, *know\_law\_idx*, and *know\_report\_idx*, all from the citizen survey.

For the administrative data, we follow the estimator in Equation 10 using outcome data from the six months following the first meeting in each block. We define the start of the post-treatment period as the month in which the first meeting was held for any *cuadrante* in a block. Note that we collapse over the pre-treatment period to calculate  $Y_{ijb}^{t=0}$  (for the six months preceding the first meeting in each block) and over the post-treatment period to calculate  $Y_{ijb}^{t=1}$  (for the six months succeeding the first meeting in each block). The relevant estimator for the meta-analysis is  $\xi_m$ .

$$Y_{jb}^{t=1} = \alpha + \xi^m T_{jb}^m + \xi^f T_{jb}^f + \gamma_b + \delta Y_{jb}^{t=0} + e_{jb} \quad (\text{A2})$$

A2

For all estimators, we cluster standard errors at the level of treatment assignment: the *cuadrante* wherever the unit  $i$  is not equivalent to the unit assignment  $j$ .

We follow the meta-analysis PAP for guidance on measuring compliance, attrition, and index construction. Note two threats to inference realized in the course of data collection:

- The recontact rate for the endline survey was lower than anticipated ( $\approx 50\%$  across the full sample). As a result, we augmented the endline survey with new subjects such that the endline sample in each *cuadrante* represents 70% of the baseline sample size. For the new respondents in the endline, we do not have baseline measures, so we impute baseline *cuadrante* means and include an indicator for new respondents in any covariate-adjusted analyses.
- Reiterating the discussion on page 25, the police survey was beset by a number of unanticipated issues, as follows:
  - We were unable to sample the prespecified 694 officers (2 per *cuadrante*). The surveys were administered at station meetings of beat-level patrol officers with permission from each station commander. The main constraint was attendance at these meetings.
  - Some police officers were not willing to provide their *cuadrante* number such that we cannot determine their assignment to treatment or control.

The experimental comparisons using police survey data will use the following procedure:

1. Condition the sample on those officers that reported their *cuadrante* number. Analyze the data for differential survey response by regressing a count of completed surveys with *cuadrante* identification from each *cuadrante* on treatment assignment using the ITT estimator above. Rejection of the null hypothesis on meetings constitutes grounds for police survey outcome exclusion from the meta-analysis.
2. Manipulation check. Given the issues with the design, we will establish a first stage in the police survey between assignment to meetings and participation in meetings. To do so, we regress the answers to the compliance questions #10 and an indicator for an affirmative answer to #10 and/or #11 on Equation 9.
3. Analyses. Further experimental analyses will be estimated on equation 9 and standard errors clustered at the *cuadrante* level.

## A.3 Meta-Analysis Outcomes and Coding

We replicate the outcomes from [Blair, Christia, and Weinstein \(2020\)](#) in the following table. Where our questions diverge in content (not simply translation) we note these departures in . We will aggregate the variables as prespecified in [Blair, Christia, and Weinstein \(2020\)](#).

A3

**Variable Coding and Survey Questionnaire.** A check in column *B* indicates that variable was measured at baseline and a check in column *E* indicates that the variable was collected at endline. The absence of checkmarks indicates the variable is absent from the dataset. Comments in red indicate site-specific adaptions of questions or measurement.

<i>B</i>	<i>E</i>	Variable name	Question text	Response options	Variable construction	Data Source
<b>PRIMARY OUTCOME FAMILY 1: SECURITY OF LIFE AND PROPERTY</b>						
1a. Negative effect on incidence of crime						
✓	✓	armedrob_num <sup>1</sup>	In the past 6 months, were you or any member of your household the victim of any ROBBERY? [IF YES:] In the last incident, was any type of arm or other object used to threaten or attack you? [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>2</sup>	Numeric	Coded as 1 if subject responds "yes" to both questions, else 0. Binary variable.	Citizen survey
✓	✓	armedrob_bin		Numeric	Equivalent to armedrob_num: coded 1 if armedrob_num = 1; 0 if armedrob_num = 0	Citizen survey
✓	✓	burglary_num <sup>3</sup>	(a) In the past 6 months, were you or any member of your household the victim of any ROBBERY? [IF YES:] In the last incident, was any type of arm or other object used to threaten or attack you? *OR* In the past 6 months, have you experienced theft of a vehicle or its parts? [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>4</sup>	Numeric	Coded as 1 if subject responds "yes" to the first question but "no" to the second of (a) OR "yes" to (b). Binary variable.	Citizen survey
✓	✓	burglary_bin		Freeform	Equivalent to burglary_num: coded 1 if burglary_num = 0; 0 if burglary_num = 0	Citizen survey
✓	✓	simpleassault_num <sup>5</sup>	In the past 6 months, has anyone attacked you or any member of your household? [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>6</sup>	Numeric	Binary measure.	Citizen survey.

<sup>1</sup>Adapted from Blair et al. (2017).

<sup>2</sup>Adapted from Blair et al. (2017).

<sup>3</sup>Adapted from Blair et al. (2017).

<sup>4</sup>Adapted from Blair et al. (2017).

<sup>5</sup>Blair et al. (2017).

<sup>6</sup>Blair et al. (2017).

<i>B</i>	<i>E</i>	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	simpleassault_bin		Numeric	Equivalent to burglary_num: coded 1 if simpleassault_num = 1; 0 if simpleassault_num = 0	Citizen survey
		aggrav assault_num	Besides any armed robbery, in the past 6 months, has anyone attacked you or any member of your household WITH A WEAPON? (INCLUDING GUNS, CUTLASSES, STICKS, ETC.) [IF YES:] How many times did this happen in the past 6 months?	Numeric	Not differentiated from simpleassault_num on survey instrument.	Citizen survey
		sexual_num	In the past 6 months, have you or any member of your household been a victim of SEXUAL ABUSE OR RAPE? (INCLUDING RAPE) [IF YES:] How many times did this happen in the past 6 months?	Numeric	Not asked about self/own household.	Citizen survey
		domestic_phys_num	Besides any sexual abuse, in the past 6 months, has anyone in your household ever PHYSICALLY ABUSED you? (INCLUDING PUSHING, SLAPPING, PUNCHING, KICKING, CHOKING, ETC.) [IF YES:] How many times did this happen in the past 6 months?	Numeric	Not asked about self/own household.	Citizen survey
		domestic_verbal_num	Besides any physical abuse, in the past 6 months, has anyone in your household ever VERBALLY ABUSED you? [INCLUDING SHOUTING, CUSSING, THREATS OF ABUSE, ETC.]	Numeric	Not asked because it is not a crime.	Citizen survey
		land_any	In the past 6 months, did you or a member of your household have a LAND DISPUTE over your house land or farm land? This include disputes that ended in the past 6 months or disputes that are still ongoing up to now. [IF YES:] Was there any violence or property destruction due to this dispute?	Numeric	Not asked because of limited contextual relevance.	Citizen survey
		other_any	In the past 6 months, were you or any member of your household a victim of any OTHER CRIME that we haven't mentioned already?		Not asked at baseline or endline due to low response rate during piloting.	Citizen survey
		other_any_violent	Coded as other_any if other_any is a violent crime	Freeform	Not asked at baseline or endline due to low response rate during piloting.	Citizen survey

B	E	Variable name	Question text	Response options	Variable construction	Data Source
		other_any_nonviolent	Coded as other_any if other_any is a non-violent crime	Freeform	Not asked at baseline or endline due to low response rate during piloting. Sum of armedrob_num, simpleassault_num	Citizen survey
✓	✓	violentcrime_num			Equivalent to burglary_num Equivalent to violentcrime_num	Citizen survey
✓	✓	nonviolentcrime_num			Equivalent to violentcrime_num	Citizen survey
✓	✓	violentcrime_num_exp			Equivalent to nonviolentcrime_num	Citizen survey
✓	✓	nonviolentcrime_num_exp			Because all crime reports are binary, equivalent to violentcrime_num	Citizen survey
✓	✓	violentcrime_bin			Because all crime reports are binary, equivalent to nonviolentcrime_num	Citizen survey
		nonviolentcrime_bin				
			Now I want to ask you some questions about different types of crimes that may have happened to SOMEONE ELSE IN THIS COMMUNITY. This can include your neighbors, friends, relatives, or any other person you know that's living IN THIS COMMUNITY?			
✓	✓	carmedrob_num <sup>8</sup>	In the past 6 months, were you or any member of your household the victim of any ROBBERY? [IF YES:] In the last incident, was any type of arm or other object used to threaten or attack you? [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>9</sup>	Numeric	Coded as 1 if subject responds "yes" to both questions, else 0. Binary variable.	Citizen survey
✓	✓	carmedrob_bin		Numeric	Equivalent to carmedrob_num: coded 1 if carmedrob_num = 1; 0 if carmedrob_num = 0	Citizen survey

<sup>7</sup>Adapted from Blair et al. (2017).

<sup>8</sup>Adapted from Blair et al. (2017).

<sup>9</sup>Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	cburglary_num <sup>10</sup>	(a) In the past 6 months, was anyone you know in this community a victim of BURGLARY or THEFT? (ROBBERY WITHOUT WEAPON) [IF YES:] In the last incident, was any type of arm or other object used to threaten or attack you? *OR * (b) In the past 6 months, have you experienced theft of a vehicle or its parts?	Numeric	Coded as 1 if subject responds "yes" to the first question but "no" to the second of (a), or "yes" to (b), else 0. Binary variable.	Citizen survey
✓	✓	cburglary_bin		Numeric	Equivalent to cburglary_num: coded 1 if armedrob_num = 1; 0 if cburglary_num = 0	Citizen survey
		caggravassault_num <sup>11</sup>	Besides any armed robbery, in the past 6 months, was anyone you know in this community attacked WITH A WEAPON? (INCLUDING GUNS, CUTLASSES, STICKS, ETC.) [IF YES:] How many times did this happen in the past 6 months?	Numeric	Not differentiated from csimpleassault_num on survey instrument.	Citizen survey
		caggravassault_bin		Numeric	Not differentiated from csimpleassault_bin on survey instrument.	Citizen survey
✓	✓	csimpleassault_num <sup>12</sup>	In the past 6 months, was anyone you know in this community attacked? [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>13</sup>	Numeric	Binary measure.	Citizen survey
✓	✓	csimpleassault_bin		Numeric	Equivalent to csimpleassault_num. Binary.	Citizen survey
		csexual_num	In the past 6 months, was anyone you know in this community SEXUALLY ABUSED? (INCLUDING RAPE) [IF MORE THAN 1:] I want to ask about the MOST RECENT incident.	Numeric	Citizen survey	
✓	✓	csexual_bin		Numeric	Equivalent to csexual_num. Binary.	Citizen survey
✓	✓	cdomestic_phys_num	Besides any sexual abuse, in the past 6 months, was anyone you know in this community PHYSICALLY ABUSED by someone in their own household? (INCLUDING PUSHING, SLAPPING, PUNCHING, KICKING, CHOKING, ETC.)	Numeric	Citizen survey	

<sup>10</sup>Adapted from Blair et al. (2017).

<sup>11</sup>Adapted from Blair et al. (2017). Only collected at endline in the Colombia study.

<sup>12</sup>Adapted from Blair et al. (2017).

<sup>13</sup>Adapted from Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
		[IF MORE THAN 1:] I want to ask about the MOST RECENT incident.				
✓	✓	cdomestic_phys_bin	In the past 6 months, was anyone you know in this community MURDERED?	Numeric	Equivalent to <code>cdomestic_phys_num</code>	Citizen survey
✓	✓	cmurder_num	[IF MORE THAN 1:] I want to ask about the MOST RECENT incident.	Numeric	Equivalent to <code>cdomestic_phys_num</code>	Citizen survey
✓	✓	cmurder_bin	In the past 6 months, did anyone you know in this community have a LAND DISPUTE over their house land or farm land? This includes disputes that ended in the past 6 months or disputes that are still ongoing up to now.	Numeric	Equivalent to <code>cmurder_num</code> Not asked because it is not a crime.	Citizen survey
		cldand_any	Besides any physical abuse, in the past 6 months, was anyone you know in this community been VERBALLY ABUSED by someone in their own household? [INCLUDING SHOUTING, CUSSING, THREATS OF ABUSE, ETC.]		Not asked because of limited contextual relevance.	Citizen survey
		cmob_num	In the past 6 months, were there any incidents of MOB JUSTICE in this community (i.e. beating or flogging of someone suspected of committing a crime)? [IF YES:] How many times did this happen in the past 6 months?		Not asked because of limited contextual relevance.	Citizen survey
		cother_any	In the past 6 months, was anyone you know in this community a victim of any OTHER CRIME that we haven't mentioned already?		Not asked at baseline or endline due to low response rate during piloting.	Citizen survey
		cother_any_violent	Coded as <code>cother_any</code> if <code>cother_any</code> is a violent crime (see general coding rule for violent crimes)		Not asked at baseline or endline due to low response rate during piloting.	Citizen survey
		cother_any_nonviolent	Coded as <code>cother_any</code> if <code>cother_any</code> is a non-violent crime (see general coding rule for non-violent crimes)		Not asked at baseline or endline due to low response rate during piloting.	Citizen survey

9

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	cviolentcrime_num				Sum of <code>carmedrob_num</code> , <code>csimpleassault_num</code> , <code>csexual_num</code> , <code>cdomestic_phys_num</code> , <code>cmurder_num</code>
✓	✓	cnonviolentcrime_num				Equivalent to <code>cburglary_num</code>
✓	✓	cviolentcrime_num_exp				Sum of <code>carmedrob_num</code> , <code>caggressa_num</code> , <code>csimpleassault_num</code> , <code>csexual_num</code> , <code>cdomestic_phys_num</code> , <code>cmurder_num</code> . Equivalent to <code>cviolentcrime_num</code> .
✓	✓	cnonviolentcrime_num_exp				Equivalent to <code>cburglary_num</code> .
✓	✓	cviolentcrime_bin				Sum of <code>carmedrob_bin</code> , <code>csimpleassault_bin</code> , <code>csexual_bin</code> , <code>cdomestic_phys_bin</code> , <code>cmurder_bin</code> . Equivalent to <code>cviolentcrime_num</code> since component variables are binary.
✓	✓	cnonviolentcrime_bin				Equivalent to <code>cnonviolentcrime_num</code> since component variable is binary.
✓	✓	crime_victim_idx				Index of <code>violentcrime_num</code> , <code>nonviolentcrime_num</code> , <code>cviolentcrime_num</code> , <code>cnonviolentcrime_num</code> Index of <code>violentcrime_num_exp</code> , <code>nonviolentcrime_num_exp</code> , <code>cviolentcrime_num_exp</code> , <code>cnonviolentcrime_num_exp</code>
✓	✓	crime_victim_idx_exp				

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	crime_victim_idx_bin			Index of violentcrime_bin, nonviolentcrime_bin, civiolentcrime_bin, cnonviolentcrime_bin	
		aarmedrob_num	Number of reports of armed robbery in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
✓	✓	aburglary_num	Number of reports of burglary or theft in community in past 6 months		Count of events in crime categories: "hurto a entidad financiera," "hurto a establecimiento comercial," "hurto a personas," "hurto a residencia," "hurto de carro," "hurto de moto," "hurto de semoviente," and "hurto por piratería terrestre."	Administrative
		aaggassault_num	Number of reports of aggravated assault in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
		asimpleassault_num	Number of reports of simple assault in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
		asexual_num	Number of reports of sexual abuse in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
✓	✓	adomestic_phys_num	Number of reports of domestic violence (physical) in community in past 6 months		Count of events in crime categories: "Lev 1098" (=child abuse) and "reincidencia de violencia intrafamiliar." Not a crime.	Administrative
		adomestic_verbal_num	Number of reports of domestic violence (verbal) in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
		aland_num	Number of reports of land disputes in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
		aland_violent_num	Number of reports of violent land disputes in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
		amob_num	Number of reports of mob justice in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
		ariot_num	Number of reports of riots in community in past 6 months		Not identifiable in the administrative crime data.	Administrative
✓	✓	amurder_num	Number of reports of murder in community in past 6 months		Count of homicides.	Administrative

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	another_num	Number of reports of other crimes in community in past 6 months		Count of kidnappings.	Administrative
✓	✓	another_num_violent	Coded as another_num if another_num is a violent crime (see general coding rule for violent crimes)		Count of kidnappings (violent) and misdemeanor (nonviolent) offences.	Administrative
✓	✓	another_num_nonviolent	Coded as another_num if another_num is a non-violent crime (see general coding rule for violent crimes)		Equivalent to another_num. Violations of <i>convictencia</i> . ≈	Administrative
✓	✓	aviolentcrime_num			Sum of adomestic_phys_num, asunder_num, another_num_violent	
		anonviolentcrime_num			Sum of aburglary_num and another_num_nonviolent.	
		ib. Positive effect on perceptions of safety (personal, land, and possessions)				
✓	✓	fear_violent <sup>14</sup>	How worried are you that you or a member of your household will be the victim of a CRIME in the coming year?	o-Not at all worried; 1-Somewhat worried; 2-Worried; 3-Very worried		Citizen survey
		fear_nonviolent <sup>15</sup>	How worried are you that you or a member of your household will be the victim of a NON-VIOLENT CRIME in the coming year? [INCLUDING BURGLARY, THEFT, ETC.]	o-Not at all worried; 1-Somewhat worried; 2-Worried; 3-Very worried	Collapsed into the previous question	Citizen survey
✓	✓	feared_walk <sup>16</sup>	In the past 6 months, have you or anyone in your family felt unsafe going out in your neighborhood? future_insecurity_idx	o-No; 1-Yes	Preserved question (in binary form) from baseline. Index of fear_violent, feared_walk	Citizen survey
		<b>PRIMARY OUTCOME FAMILY 2: CITIZEN PERCEPTIONS OF THE POLICE</b>				
		2. Positive effect on citizen perceptions of police				
		✓	I generally trust the police. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Agree; 3-Strongly agree; NA-Do not know/Refuse to answer	Outcome scale preserved from baseline	Citizen survey

<sup>14</sup>Adapted from Cheema et al. (2017).

<sup>15</sup>Adapted from Cheema et al. (2017).

<sup>16</sup>Adapted from Afrobarometer (2016).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓		satis_general	I am satisfied with the service that the police provide. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; NA-Do not know/Refuse to answer	Was not asked in baseline.	Citizen survey
✓	✓	satis_idx			Index of satis_trust and satis_general	Citizen survey
<b>PRIMARY OUTCOME FAMILY 3: POLICE PERCEPTIONS OF AND BEHAVIORS TOWARD CITIZENS</b>						
3a.		Positive effect on perceptions of police empathy, accountability, and abuse and corruption concerns				
✓		empathy_complaints	When people complain about the police, they usually have a good reason. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Agree; 3-Strongly agree; NA-Do not know/Refuse to answer		Officer survey
✓		empathy_reports	Most things that people report to the police are worth taking seriously. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Agree; 3-Strongly agree; NA-Do not know/Refuse to answer		Officer survey
✓		empathy_idx			Index of empathy_complaints, empathy_reports	
✓		account_pol_matter	The police leadership takes citizen complaints about officers seriously. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Agree; 3-Strongly agree; NA-Do not know/Refuse to answer		Officer survey
✓		hypothetical2	If an officer in your agency engaged in this behavior and was discovered doing so, what if any discipline do YOU think WILL follow?	o-None; 1-Verbal reprimand; 2-Written reprimand; 3-Period of suspension without pay; 4-Demotion in rank; 5-Dismissal; NA-Don't know/Refuse to answer	This was asked as an open response question. We code responses qualitatively to accord with this scale. The category "2" encompasses many types of punishment.	Officer survey
✓		hypothetical2_reportself	Do you think YOU would report a fellow police officer who engaged in this behavior?	o-Definitely not; 1-Probably not; 2-Probably yes; 3-Definitely yes; NA-Do not know/Refuse to answer; 99-other		Officer survey

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓		hypothetical2_reportthe	Do you think MOST POLICE OFFICERS would report a fellow police officer who engaged in this behavior?	o-Definitely not; 1-Probably not; 2-Probably yes; 3-Definitely yes; NA-Do not know/Refuse to answer		Officer survey
✓		hypothetical3_punishment	If an officer in your agency engaged in this behavior and was discovered doing so, what if any discipline do YOU think WILL follow?	o-None; 1-Verbal reprimand; 2-Written reprimand; 3-Period of suspension without pay; 4-Demotion in rank; 5-Dismissal; NA-Don't know/Refuse to answer	This was asked as an open response question. We code responses qualitatively to accord with this scale. The category "2" encompasses many types of punishment.	Officer survey
✓		hypothetical3_reportself	Do you think YOU would report a fellow police officer who engaged in this behavior?	o-Definitely not; 1-Probably not; 2-Probably yes; 3-Definitely yes; NA-Do not know/Refuse to answer		Officer survey
✓		hypothetical3_reportthe	Do you think MOST POLICE OFFICERS would report a fellow police officer who engaged in this behavior?	o-Definitely not; 1-Probably not; 2-Probably yes; 3-Definitely yes; NA-Do not know/Refuse to answer		Officer survey
✓		hypothetical5_punishment	If an officer in your agency engaged in this behavior and was discovered doing so, what if any discipline do YOU think WILL follow?	o-None; 1-Verbal reprimand; 2-Written reprimand; 3-Period of suspension without pay; 4-Demotion in rank; 5-Dismissal; NA-Do not know/Refuse to answer	This was asked as an open response question. We code responses qualitatively to accord with this scale. The category "2" encompasses many types of punishment.	Officer survey
✓		hypothetical5_reportself	Do you think YOU would report a fellow police officer who engaged in this behavior?	o-Definitely not; 1-Probably not; 2-Probably yes; 3-Definitely yes; NA-Do not know/Refuse to answer		Officer survey
✓		hypothetical5_reportthe	Do you think MOST POLICE OFFICERS would report a fellow police officer who engaged in this behavior?	o-Definitely not; 1-Probably not; 2-Probably yes; 3-Definitely yes; NA-Do not know/Refuse to answer		Officer survey

B	E	Variable name	Question text	Response options	Variable construction	Data Source
12		✓ accountability_idx		Index of account_pol_matter, hypothetical2_punishment, hypothetical2_reportself, hypotheti- cal2_reportothers, hypothetical3_punishment, hypothetical3_reportself, hypotheti- cal3_reportothers, hypothetical5_punishment, hypothetical5_reportself, hypothetical5_reportothers		
		✓ hypothetical5_abusese	Two police officers on foot patrol surprise a man who is attempting to break into an automobile. The man flees. They chase him for about two blocks before apprehending him by tackling him and wrestling him to the ground. After he is under control, both officers punch him a couple of times in the stomach as punishment for fleeing and resisting. Do YOU CONSIDER THIS BEHAVIOR TO BE SERIOUS MISCONDUCT?	o-Not at all serious; 1-Somewhat serious; 2-Serious; 3-Very serious; 97-Do not know; 98-Refuse to answer	Officer survey	
		✓ hypothetical5_abuseoth	Do MOST POLICE OFFICERS consider this behavior to be serious misconduct?	o-Not at all serious; 1-Somewhat serious; 2-Serious; 3-Very serious; 97-Do not know; 98-Refuse to answer	Officer survey	
		✓ abuse_idx		Index of hypothetical5_abusese, hypothetical5_abuseoth		
		✓ hypothetical2_corruptsel	A police officer routinely accepts money or gifts in exchange for ignoring the activities of an armed group/gang in his beat. Do you consider this behavior to be serious misconduct?	o-Not at all serious; 1-Somewhat serious; 2-Serious; 3-Very serious; NA-Do not know/Refuse to answer	Officer survey	
		✓ hypothetical2_corruptoth	A police officer routinely accepts money or gifts in exchange for ignoring the activities of an armed group/gang in his beat. Do MOST POLICE OFFICERS consider this behavior to be serious misconduct?	o-Not at all serious; 1-Somewhat serious; 2-Serious; 3-Very serious; NA-Do not know/Refuse to answer	Officer survey	
13		✓ hypothetical3_corruptsel	A police officer stops a motorcyclist for a traffic violation. The officer agrees to accept a personal gift of half of the amount of the fine in exchange for not issuing a citation. Do you consider this behavior to be serious misconduct?	o-Not at all serious; 1-Somewhat serious; 2-Serious; 3-Very serious; NA-Do not know/Refuse to answer	Officer survey	
		✓ hypothetical3_corruptoth	A police officer stops a motorcyclist for a traffic violation. The officer agrees to accept a personal gift of half of the amount of the fine in exchange for not issuing a citation. Do MOST POLICE OFFICERS consider this behavior to be serious misconduct?	o-Not at all serious; 1-Somewhat serious; 2-Serious; 3-Very serious; NA-Do not know/Refuse to answer	Officer survey	
		✓ corrupt_idx		Index of hypothetical2_corruptsel; hypothetical2_corruptoth;		
		✓ officer_attitude_idx		hypothetical3_corruptoth; Index of corrupt_idx, abuse_idx, accountability_idx, empathy_idx		
		3a. Negative effect reporting of police abuse and bribery	In the past 6 months, have you ever witnessed or heard about police officers PHYSICALLY ABUSING people from your community? [INCLUDING PUSHING, SLAPPING, PUNCHING, KICKING, CHOKING, ETC.]	o-No; 1-Yes; NA- Do not know/Refuse to answer	Citizen survey	
		✓ ✓ policeabuse_phys_any <sup>17</sup>				
		✓ ✓ policeabuse_verbal_any <sup>18</sup>	Besides any incidents of physical abuse, in the past 6 months, have you ever witnessed or heard about police officers VERBALLY ABUSING people from your community? [INCLUDING SHOUTING, CÜSSING, ETC.] This includes verbal abuse against you or someone in your family.	o-No; 1-Yes; NA-Do not know/Refuse to answer	Citizen survey	

<sup>17</sup>Adapted from Blair et al. (2017).

<sup>18</sup>Adapted from Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	policeabuse_any			Recoded 0 if policeabuse_verbal_any = 0 and policeabuse_phys_any = 0; 1 if policeabuse_verbal_any = 1 or policeabuse_phys_any = 1	Citizen survey
✓	✓	policeabuse_phys_num <sup>19</sup>	In the past 6 months, have you ever witnessed or heard about police officers PHYSICALLY ABUSING people from your community? (INCLUDING PUSHING, SLAPPING, PUNCHING, KICKING, CHOKING, ETC.) [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>20</sup>	Numeric	Binary. Equivalent to policeabuse_phys_any	Citizen survey
✓	✓	policeabuse_verbal_num <sup>21</sup>	Besides any incidents of physical abuse, in the past 6 months, have you ever witnessed or heard about police officers VERBALLY ABUSING people from your community? [INCLUDING SHOUTING, CUSSING, ETC.] This includes verbal abuse against you or someone in your family. [IF YES:] How many times did this happen in the past 6 months? [IF MORE THAN 1:] I want to ask about the MOST RECENT incident. <sup>22</sup>	Numeric	Binary. Equivalent to policeabuse_verbal_any	Citizen survey
✓	✓	policeabuse_num			Sum of number of incidents of verbal (policeabuse_verbal_num) or physical abuse (policeabuse_phys_num) by police officers in the past 6 months	

<sup>19</sup> Adapted from Blair et al. (2017).

<sup>20</sup> Blair et al. (2017).

<sup>21</sup> Adapted from Blair et al. (2017).

<sup>22</sup> Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	policeabuse_verbal_repor	To the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	o-No; 1-Community leaders; 2-Police station or police commander; 3-Any other government agency; 4-NGO; 5-Journalist; 6-OTHER COUNTRY-SPECIFIC VENUES]; 97-Do not know; 98-Refuse to answer	Recoded 0 if policeabuse_verbal_num = 0 and policeabuse_verbal_repor = 0; 1 if policeabuse_verbal_num > 0 and policeabuse_verbal_num = 2	Citizen survey
✓	✓	policeabuse_phys_report	To the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	o-No; 1-Community leaders; 2-Police station or police commander; 3-Any other government agency; 4-NGO; 5-Journalist; 6-OTHER COUNTRY-SPECIFIC VENUES]; 97-Do not know; 98-Refuse to answer	Due to an error in survey programming, this question was not asked correctly.	Citizen survey
✓	✓	policeabuse_report			Recoded 0 if policeabuse_verbal_repor = 0 and policeabuse_phys_report = 0; 1 if policeabuse_verbal_repor > 0 or policeabuse_phys_report > 0	
✓		bribe_freq	How many times in the past 6 months have you made an unofficial payment to the police?	1-None; 2-Once; 3-Between 2 and 5 times; 4-More than 5 times; NA-Do not know/Refuse to answer	Categorical variable for frequency of unofficial payments to the police in the past 6 months	Citizen survey
✓		bribe_amt	[IF ANY:] The last time you made an unofficial payment to the police, how much was it?	Numeric	Note: this was asked as an open-ended question. Some citizens reported what they were asked for; others reported what they paid. We cannot fully distinguish between the two amounts.	Citizen survey

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	police_abuse_idx			Index of policeabuse_any, policeabuse_num, bribe_freq, bribe_amt	
<b>PRIMARY OUTCOME FAMILY 4: BEHAVIORAL COOPERATION OF CITIZENS WITH THE POLICE</b>						
<b>4a. Positive effect on reporting of crime victimization</b>						
		acrime_hline	Total number of reports of crimes to hotline	Total number of hotline calls is atips_hline. We do not have any indicator of victimization specifically.	Administrative	
		aviolent_hline	Number of reports of violent crimes to hotline	Not disaggregated by crime/accusation.	Administrative	
		anonviolent_hline	Number of reports of non-violent crimes to hotline	Not disaggregated by crime/accusation.	Administrative	
		facrime_station	Total number of reports of crimes to nearest police station	No records of station reports.	Administrative	
		aviolent_station	Number of reports of violent crimes to nearest police station	No records of station reports.	Administrative	
		anonviolent_station	Number of reports of non-violent crimes to nearest police station	No records of station reports.	Administrative	
		aburglary_hline	Number of reports of burglary to hotline	Not disaggregated by crime/accusation.	Administrative	
		aarmedrob_hline	Number of reports of armed robbery to hotline	Not disaggregated by crime/accusation.	Administrative	
		arape_hline	Number of reports of rape to hotline	Not disaggregated by crime/accusation.	Administrative	
		amurder_hline	Number of reports of murder to hotline	Not disaggregated by crime/accusation.	Administrative	
		asimpleassault_hline	Number of reports of simple assault to hotline	Not disaggregated by crime/accusation.	Administrative	
		aaggassault_hline	Number of reports of aggravated assault to hotline	Not disaggregated by crime/accusation.	Administrative	
		atheft_hline	Number of reports of theft to hotline	Not disaggregated by crime/accusation.	Administrative	
		aburglary_station	Number of reports of burglary to nearest police station	No records of station reports.	Administrative	
		aarmedrob_station	Number of reports of armed robbery to nearest police station	No records of station reports.	Administrative	

B	E	Variable name	Question text	Response options	Variable construction	Data Source
		arape_station	Number of reports of rape to nearest police station		No records of station reports.	Administrative
		amurder_station	Number of reports of murder to nearest police station		No records of station reports.	Administrative
		asimpleassault_station	Number of reports of simple assault to nearest police station		No records of station reports.	Administrative
		aaggassault_station	Number of reports of aggravated assault to nearest police station		No records of station reports.	Administrative
		atheft_station	Number of reports of theft to nearest police station		No records of station reports.	Administrative
<i>Actual crime (survey)</i>						
✓	✓	aarmedrob_report <sup>23</sup>	In the past 6 months, were you or any member of your household the victim of any ARMED ROBBERY? (ROBBERY WITH ANY KIND OF WEAPON, INCLUDING GUNS, CUTLASSES, STICKS, ETC.) Where did you report this case? [SELECT ALL THAT APPLY]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum; 8-Other country specific forum; 88-Other; 98-Refuse to Answer	Recoded 0 if aarmedrob_num = 0 or aarmedrob_report = 0; 1 if aarmedrob_num > 0 and aarmedrob_report = 1	Citizen survey
✓	✓	burglary_report <sup>24</sup>	Besides any armed robbery, in the past 6 months, were you or any member of your household the victim of BURGLARY or THEFT? [ROBBERY WITHOUT WEAPON]. Where did you report this case? [SELECT ALL THAT APPLY]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum; 8-Other country specific forum; 88-Other	Recoded 0 if burglary_num = 0 or burglary_report = 0; 1 if burglary_num > 0 and burglary_report = 1	Citizen survey

<sup>23</sup>Adapted from Blair et al. (2017).<sup>24</sup>Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	simpleassault_report <sup>25</sup>	Besides any armed robbery, in the past 6 months, has anyone attacked you or any member of your household WITH A WEAPON? [INCLUDING GUNS, CUTLASSES, STICKS, ETC.] Where did you report this case? [SELECT ALL THAT APPLY]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Recoded 0 if simpleassault_num = 0 or simpleassault_report = 0; 1 if simpleassault_num > 0 and simpleassault_report = 1	Citizen survey
		other_report <sup>26</sup>	In the past 6 months, were you or any member of your household a victim of any OTHER CRIME that we haven't mentioned already? Where did you report this case? [SELECT ALL THAT APPLY]		Questions about other crimes not asked	Citizen survey
		other_report_violent			textcoloredQuestions about other crimes not asked	
		other_report_nonviolent			textcoloredQuestions about other crimes not asked	
✓	✓	violentcrime_report_num			Sum of armedrob_report, simpleassault_report	
✓	✓	nonviolentcrime_report_n			Equivalent to burglary_report	
✓	✓	carmedrob_report <sup>27</sup>	In the past 6 months, was anyone you know in this community a victim of ARMED ROBBERY and [ROBBERY WITH ANY KIND OF WEAPON, INCLUDING GUNS, CUTLASSES, STICKS, ETC.] to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Recoded 0 if carmedrob_num = 0 or carmedrob_report = 0; 1 if carmedrob_num > 0 and carmedrob_report = 1	Citizen survey

<sup>25</sup>Blair et al. (2017).<sup>26</sup>Blair et al. (2017). Only collected at endline in the Colombia study.<sup>27</sup>Adapted from Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	cburglary_report <sup>28</sup>	Besides any armed robbery, in the past 6 months, was anyone you know in this community a victim of BURGLARY or THEFT and [ROBBERY WITHOUT WEAPON] to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Recoded 0 if cburglary_num = 0 or cburglary_report = 0; 1 if cburglary_num > 0 and cburglary_report = 1	Citizen survey
		caggaassault_report <sup>29</sup>	Besides any armed robbery, in the past 6 months, was anyone you know in this community attacked WITH A WEAPON and [INCLUDING GUNS, CUTLASSES, STICKS, ETC.] to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]			
✓	✓	csimpleassault_report <sup>30</sup>	In the past 6 months, was anyone you know in this community attacked WITHOUT a weapon and to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Recoded 0 if csimpleassault_num = 0 or csimpleassault_report = 0; 1 if csimpleassault_num > 0 and csimpleassault_report = 1	Citizen survey
		csexual_report	In the past 6 months, was anyone you know in this community SEXUALLY ABUSED? [INCLUDING RAPE] and to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]			

<sup>28</sup>Adapted from Blair et al. (2017).<sup>29</sup>Adapted from Blair et al. (2017). Only collected at endline in the Colombia study.<sup>30</sup>Adapted from Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓	✓	cdomestic_phys_report	Besides any sexual abuse, in the past 6 months, was anyone you know in this community PHYSICALLY ABUSED by someone in their own household and [INCLUDING PUSHING, SLAPPING, PUNCHING, KICKING, CHOKING, ETC.] to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Recoded o if cdomestic_phys_num = 0 or cdomestic_phys_report = o: 1 if cdomestic_phys_num > 0 and cdomestic_phys_report = 1	Citizen survey
✓	✓	cmurder_report <sup>31</sup>	In the past 6 months, was anyone you know in this community MURDERED and to the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Recoded o if cmurder_num = 0 or cmurder_report = o; 1 if cmurder_num > 0 and cmurder_report = 1	Citizen survey
		cother_report <sup>32</sup>	In the past 6 months, was anyone you know in this community a victim of any OTHER CRIME that we haven't mentioned already? To the best of your knowledge, was this incident reported to anyone? [SELECT ALL THAT APPLY]	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 88-Other	Questions about other crimes not asked.	Citizen survey
		cother_report_violent	Coded as cother_report if cother_any is a violent crime (see general coding rule for violent crimes)	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 87-Don't know; 98-Refuse to answer	Questions about other violent crimes not asked.	Citizen survey
		cother_report_nonviolent	Coded as cother_report if cother_any is a non-violent crime (see general coding rule for non-violent crimes)	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 87-Don't know; 98-Refuse to answer	Questions about other non-violent crimes not asked.	Citizen survey
✓	✓	cviolentcrime_report_num		Sum of cassrob_report, cimplesassault_report, csexual_report, cdomestic_phys_report, cmurder_report	Equivalent to burglary_report	
✓	✓	cnonviolentcrime_report_				

<sup>31</sup>Only collected at endline in the Colombia study.<sup>32</sup>Only collected at endline in the Colombia study.

B	E	Variable name	Question text	Response options	Variable construction	Data Source
		Hypothetical crime (survey)				
			There are many places you can go to solve your crime here in [COUNTRY]. We got the POLICE, community leaders, [COUNTRY SPECIFIC FORUM 1], and [COUNTRY SPECIFIC FORUM 2]. Now I want to ask about what you think should happen for different types of crime that might happen in your community. <sup>33</sup>			
✓		burglaryres <sup>34</sup>	If there's a BURGLARY in your community, who you would most like to resolve the situation? [DO NOT READ OPTIONS]	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 97-Don't know; 98-Refuse to answer	Recoded such that burglaryres = 1 if respondent prefers the police or courts to resolve the situation; burglaryres = 0 if otherwise.	Citizen survey
✓		dviolres <sup>35</sup>	If a MAN BEAT HIS WOMAN in your community, who you would most like to resolve the situation? [DO NOT READ OPTIONS]	o-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum2; 8-Other country specific forum3; 97-Don't know; 98-Refuse to answer	Recoded such that dviolres = 1 if respondent prefers the police or courts to resolve the situation; dviolres = 0 if otherwise.	Citizen survey

<sup>33</sup>Blair et al. (2017).<sup>34</sup>Blair et al. (2017). Only collected at endline in the Colombia study.<sup>35</sup>Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source
✓		armedrobres <sup>36</sup>	If there's an ARMED ROBBERY in your community, who you would most like to resolve the situation? [DO NOT READ OPTIONS]	0-Nowhere; 1-Police; 2-Court; 3-Town chief or elders; 4-Community watch group; 5-Settled directly with the perpetrator; 6-Other country specific forum; 7-Other country-specific forum; 8-Other country specific forum; 97-Don't know; 98-Refuse to answer	Recoded such that armedrobres = 1 if respondent prefers the police or courts to resolve the situation; armedrobres = 0 if otherwise.	Citizen survey
✓		crimeres_idx			Index of burglaryres, dviores, and armedrobres	Citizen survey
✓		crime_reporting_idx			Index of violentcrime_report_num, nonviolentcrime_report_n, violentcrime_report_num, nonviolentcrime_report_n, crimeres_idx.	Citizen survey
<b>4b. Positive effect on reporting of crime prevention tips</b>						
✓	✓	atips_hline <sup>37</sup>	Number of crime prevention tips reported via hotline (if available in both T and C locations)		Number of crime prevention tips reported	Administrative
		atips_box	ADMIN: Number of crime prevention tips reported via comment boxes (if available in both T and C locations)	We do not have data on any comment boxes (if they exist)	We do not have data on any comment boxes (if they exist)	Administrative
✓	✓	contact_pol_susp_activit	In the past 6 months, have you ever contacted the police to alert them to suspicious or criminal activity in your community?	0-No; 1-Yes; NA-Do not know; 98-Refuse to answer		Citizen survey
✓	✓	give_info_pol_investigat	In the past 6 months, have you ever given information to the police to assist with an investigation?	0-No; 1-Yes; 97-Do not know; 98-Refuse to answer		Citizen survey
✓	✓	crime_tips_idx			Index of contact_pol_susp_activit and give_info_pol_investigat	
✓	✓	tips_idx			Index of atips_hline, atips_box, crime_tips_idx	

<sup>36</sup>Plair et al. (2017).<sup>37</sup>Not collected in Liberia.

B	E	Variable name	Question text	Response options	Variable construction	Data Source
<b>4c. Positive effect on reporting of victimization by the police</b>						
		apolvtm_hline	Number of incidents of victimization by the police reported via hotline (if available in both T and C locations)	1-Very unlikely; 2-Unlikely; 3-Likely; 4-Very likely; 97-Don't know; 98-Refuse to answer	We do not have access to this data if it exists.	Administrative
		apolvtm_cmtbox	Number of incidents of victimization by the police reported via comment boxes (if available in both T and C locations)	1-Very unlikely; 2-Unlikely; 3-Likely; 4-Very likely; 97-Don't know; 98-Refuse to answer	We do not have access to this data if it exists.	Administrative
		apolvtm_station	Number of incidents of victimization by the police reported to nearest station	1-Very unlikely; 2-Unlikely; 3-Likely; 4-Very likely; 97-Don't know; 98-Refuse to answer	We do not have access to this data if it exists.	Administrative
<i>Reporting of victimization by the police (hypothetical)</i>						
		dutydrink_report	Suppose you see a uniformed police officer drinking alcohol in your community. How likely would you be to report that situation?	1-Very unlikely; 2-Unlikely; 3-Likely; 4-Very likely; 97-Don't know; 98-Refuse to answer	Not collected because officers often wear uniforms off-duty so distinguishing on-duty drinking is difficult	Citizen survey
✓	✓	policebeating_report	Suppose you see a group of officers unjustifiably beating someone in your community. How likely would you be to report that situation?	1-Very unlikely; 2-Unlikely; 3-Likely; 4-Very likely; 97-Don't know; 98-Refuse to answer		Citizen survey
✓	✓	police_abuse_report_idx				Index of dutydrink_report or policeabuse_report_idx.
<b>MECHANISM FAMILY 1: PERCEIVED COST TO CITIZENS COOPERATING WITH THE POLICE</b>						
<b>M1a. Positive effect on beliefs about police intentions</b>						
<i>Perceptions of police intentions (case management)</i>						
		polcaseserious	Imagine someone is a VICTIM of an armed robbery in your community and they take the case to the POLICE. I want to ask you what you think will happen. The police will take the case seriously and investigate. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer		Citizen survey
		polcasefair	The police will be fair to both complainant and defendant in the investigation. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer	This outcome is modified to refer to handling the case fairly, rather than the investigation, which are not conducted by the police but rather the public prosecutor.	Citizen survey

B	E	Variable name	Question text	Response options	Variable construction	Data Source
<b>Perceptions of police intentions (general)</b>						
			Ok, now I want to ask you about what you think about the police in general.			
✓	✓	polint_corrupt <sup>38</sup>	The police are corrupt. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer	In our construction of this variable we reverse the order of this variable to ensure that a higher value indicates a positive effect on citizen belief about police intentions.	Citizen survey
✓	✓	polint_quality <sup>39</sup>	The police provide the same quality of service to all citizens. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer		Citizen survey
✓	✓	polint_idx			Index of polint_corrupt and polint_quality	
✓	✓	intentions_idx			Index of polcaseserious, polcasefair, polint_idx	
<b>M1b. Positive effect on knowledge of criminal justice system</b>						
✓	✓	know_law_suspect	If you see a dead body lying in the street and you report it to the police, [COUNTRY] law says the police must hold you as a suspect. True or false?	o-False; 1-True; 97-Do not know; 98-Refuse to answer		Citizen survey
✓	✓	know_law_lawyer <sup>40</sup>	If you take your case to court and you don't have money to pay a lawyer, [COUNTRY] law says the government must provide a lawyer for you. True or false?	o-False; 1-True; 97-Do not know; 98-Refuse to answer		Citizen survey
✓	✓	know_law_fees	If you take a case to the police, [COUNTRY] law says the police can charge a fee to register the case. True or false?	o-False; 1-True; 97-Do not know; 98-Refuse to answer		Citizen survey
✓	✓	know_law_vaw	According to [COUNTRY] law, it is a crime to beat on one's wife. True or false?	o-False; 1-True; 97-Do not know; 98-Refuse to answer		Citizen survey

<sup>38</sup>Adapted from Sunshine and Tyler (2005).<sup>39</sup>Sunshine and Tyler (2005).<sup>40</sup>Blair et al. (2017). Only collected at endline in the Colombia study.

B	E	Variable name	Question text	Response options	Variable construction	Data Source
<b>Mr. Positive effect on norms of citizens cooperation with police</b>						
✓	✓	know_law_idx			Index of know_law_suspect, know_law_lawyer, know_law_fees, and know_law_vaw	
✓	✓	know_report_followup	If a crime is reported to the police using the hotline, an officer must follow up with the complainant in person in order for the crime to be recorded by the police. True or False? [ENUMERATOR: IS RESPONDENT CORRECT?]	o-No; 1-Yes		Citizen survey
✓	✓	know_report_station	Do you know where the nearest police station is? [ENUMERATOR: IS RESPONDENT CORRECT?]	o-No; 1-Yes		Collected as a self-reported binary variable due to proliferation of police facilities.
✓	✓	know_report_idx			Index of know_report_followup, know_report_station	
✓	✓	know_idx			Index of know_law_idx, know_report_idx	
<b>Mrc. Positive effect on norms of citizens cooperation with police</b>						
✓	✓	reportnorm_theft <sup>41</sup>	If there is a BURGLARY in your community, people can get angry if you take it to the police. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer		In our construction of this variable we reverse the order for the responses to ensure that a higher value indicates a positive effect on norms of citizen cooperation with police.
✓	✓	reportnorm_abuse <sup>42</sup>	If a MAN BEATS HIS WIFE in your community, people can get angry if you take it to the police. Agree or disagree?	o-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer		In our construction of this variable we reverse the order for the responses to ensure that a higher value indicates a positive effect on norms of citizen cooperation with police.

<sup>41</sup>Blair et al. (2017).<sup>42</sup>Blair et al. (2017).

B	E	Variable name	Question text	Response options	Variable construction	Data Source	
✓	✓	obeynorm <sup>43</sup>	You should do what the police tell you to do even when you do not understand the reasons for their decisions. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer	In our construction of this variable we reverse the order for the responses to ensure that a higher value indicates a positive effect on norms of citizen cooperation with police.	Citizen survey	
✓	✓	norm_idx			Index of reportnorm_theft, reportnorm_abuse, obeynorm		
<b>MECHANISM FAMILY 2: PERCEIVED RETURNS TO CITIZENS COOPERATING WITH THE POLICE</b>							
✓	✓	M2a. Positive effect on beliefs about police capacity	M2a. Positive effect on beliefs about police capacity	The police have the capacity to respond to incidents of crime in a timely manner. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer	Citizen survey	
✓	✓	polcap_timely					
✓	✓	polcap_investigate	The police have the capacity to investigate crimes and gather evidence effectively. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer	Citizen survey		
✓	✓	police_capacity_idx			Index of polcap_timely, polcap_investigate		
✓	✓	M2b. Positive effect on perceptions of responsiveness to citizen feedback	M2b. Positive effect on perceptions of responsiveness to citizen feedback	The police act upon citizen comments and complaints about security in my community. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Neither agree nor disagree; 3-Agree; 4-Strongly agree; 97-Do not know; 98-Refuse to answer	Citizen survey	
<b>SECONDARY OUTCOME FAMILY 1: INCREASE IN TRUST IN THE STATE</b>							
✓	✓	S1. Positive effect on trust in the state	S1. Positive effect on trust in the state	How much do you trust the government of [COUNTRY]?	1-Not at all; 2-just a little; 3-Somewhat; 4-A lot; 97-Don't know; 98-Refuse to answer		
<b>SECONDARY OUTCOME FAMILY 2: INCREASE IN COMMUNAL TRUST</b>							
✓	✓	S2. Positive effect on communal trust	S2. Positive effect on communal trust				
26							
B	E	Variable name	Question text	Response options	Variable construction	Data Source	
✓	✓	trust_community	Most people in my community can be trusted. Agree or disagree?	0-Strongly disagree; 1-Disagree; 2-Agree; 3-Strongly agree; 97-Do not know; 98-Refuse to answer		Citizen survey	
<b>COMPLIANCE WITH TREATMENT: CITIZEN INTERACTIONS WITH POLICE</b>							
✓	✓	C. Positive effect on rate of citizen interactions with police	C. Positive effect on rate of citizen interactions with police	ameeting_count	Attendance sheets at community meetings	Percentage Attendance	
✓	✓	compliance_patrol	About how often do you see police officers patrolling your area on FOOT?	1-Daily; 2-Weekly; 3-Monthly; 4-Seasonally; 5-Less than seasonally; 97-Do not know; 98-Refuse to answer		Sum of citizen attendance across three meetings as a proportion of invitations distributed. No meetings during baseline, so measured only at endline.	Administrative
✓	✓	compliance_freq	About how often do you see police officers patrolling your area while in a vehicle or on a motorbike?	1-Daily; 2-Weekly; 3-Monthly; 4-Seasonally; 5-Less than seasonally; 97-Do not know; 98-Refuse to answer		In our construction of this variable we reverse the order of this variable such that a higher value indicates a positive effect on citizen interactions with the police.	Citizen survey
✓	✓	compliance_meeting	In the past 6 months, have you HEARD ABOUT, SEEN, OR ATTENDED community meetings with police officers taking place in your area?	o-No; 1-Yes; 97-Do not know; 98-Refuse to answer		In our construction of this variable we reverse the order of this variable such that a higher value indicates a positive effect on citizen interactions with the police.	Citizen survey
✓	✓	compliance_idx				Index of compliance_patrol, compliance_freq, compliance_meeting	
27							

<sup>43</sup>Sunshine and Tyler (2003). Only collected at endline in the Colombia study.

#### A.4 Additional tables and figures

Figure A1: Flyers with information about the 123 Line



Figure A2: Flyers with information about the Seguridad en Línea platform



#### A.5 Meeting Protocol

Facilitator:

1. Introduce self
2. General introduction to the meeting
3. Clarify that the meeting is a space of respect
4. Attendance sheet + explanation of the Whatsapp group

Police officers:

1. Introduce the officers (names, ranks, cell phone numbers) (1 minute)
2. Name of the station commander (who serves as supervisor of the officers for the cuadrante/quadrant) and cell phone number (1 minute)
3. Services the police offer the community (3 minutes)
  - (a) Name the services
  - (b) What the officers do
4. Characterization of the *cuadrante* (police beat): territory, number of officers assigned to the cuadrante, shifts, etc. (2 minutes)
5. Identify local police stations (1 minute)
6. Review how to report crimes (1 minute)
  - (a) Line 123: emergency reports
  - (b) Segurida en Línea: anonymous reports
  - (c) To report: commercial theft, personal theft, residential theft, computer crime, child pornography, extortion
7. Review activities and interventions implemented in the beat during the past weeks
8. Distribute pamphlets, if applicable
9. Dialogue with participants
  - (a) Questions for patrolmen
  - (b) Patrolmen respond and explain what problems they can address directly and which ones are the responsibility of another entity
10. Agreement between participants and patrolmen
  - (a) Come to an agreement over objectives for the next meeting
  - (b) Identify concrete measures they can take to reach those objectives
11. Participants and patrolmen sign the agreement
12. Information about the next meeting (expected date, time, and place)

Table A1: Balance tests

	Meetings	Leaflet	Community & Leaflet	Control Mean	Control Std. dev.	p-value for dif.
Population	69.701	-(582.57)	-352.08	-(586.49)	-221.56	-(569.71)
Households	13.529	-(155.04)	-82.425	-(156.01)	-9.167	-(156.63)
People per household	0.06	-(0.055)	-0.034	-(0.050)	-0.032	-(0.050)
Share of rented homes	0.002	-(0.004)	-0.002	-(0.003)	0.002	-(0.004)
Avg. bedrooms per household	-0.027	-(0.052)	0.071	-(0.054)	-0.009	-(0.052)
Share of overcrowding households	0.009	-(0.009)	0	-(0.009)	0.008	-(0.010)
Household shares cooking with electricity	0.001	-(0.004)	-0.003	-(0.003)	0.001	-(0.005)
Household shares cooking with piped gas	0.004	-(0.012)	-0.01	-(0.010)	0.004	-(0.014)
Household shares with fridge or TV	0.004	-(0.005)	-0.001	-(0.004)	0.004	-(0.006)
Household shares with computer	0.004	-(0.008)	-0.006	-(0.006)	0.003	-(0.010)
Household shares with motorcycle	0.005	-(0.008)	-0.006	-(0.007)	0.005	-(0.010)
Household shares with landline	0.004	-(0.014)	-0.002	-(0.013)	0.016	-(0.012)
Household shares with pipeline gas	-0.007	-(0.017)	-0.003	-(0.019)	-0.018	-(0.017)
Employed per household	0.004	-(0.020)	0.005	-(0.018)	-0.013	-(0.020)
Unemployed per household	0.001	-(0.006)	0	-(0.005)	-0.005	-(0.005)
Retirees per household	0.001	-(0.001)	-0.001	-(0.001)	0.002	-(0.001)
Household shares with family living abroad	0	-(0.004)	0.003	-(0.004)	0.005	-(0.004)
Share of males per household	-0.009*	-(0.005)	-0.001	-(0.004)	-0.008	-(0.006)
Share of females per household	0.002	-(0.006)	0.006	-(0.005)	0	-(0.007)
Share of under-aged per household	0.002	-(0.007)	0.004	-(0.007)	-0.006	-(0.007)
Share of seniors per household	0.008	-(0.009)	-0.005	-(0.007)	0.014	-(0.010)
Household shares with male head	-0.007	-(0.007)	0.003	-(0.006)	-0.008	-(0.008)
Household shares with single-male parent	-0.008	-(0.005)	-0.002	-(0.005)	-0.010**	-(0.005)
Household shares with single-female parent	-0.003	-(0.006)	-0.002	-(0.006)	-0.007	-(0.006)
Household shares w/o children at home	-0.01	-(0.009)	0.001	-(0.008)	-0.007	-(0.009)
Household shares with university students	0.033	-(0.041)	-0.056	-(0.039)	-0.004	-(0.040)
Household head born in Colombia	-0.01	-(0.009)	0.003	-(0.007)	-0.017	-(0.010)
Household head born in Medellín	0.004	-(0.012)	-0.004	-(0.011)	-0.002	-(0.011)
Log of monthly rent	0.1	-(0.075)	0.049	-(0.072)	0.027	-(0.073)
Homicide rate	-5.493	-(12.530)	-16.499	-(11.523)	15.001	-(19.392)
Theft rate	-4.739	-(5.615)	-6.461	-(5.515)	-4.83	-(5.089)
Lagged homicide rate	3.053	-(6.980)	-1.413	-(6.095)	9.098	-(7.425)
Lagged theft rate	-43.301	-(40.045)	-46.651	-(40.743)	-36.574	-(39.085)
Second study treatment condition	0.034	-(0.024)	0.023	-(0.022)	0.039*	-(0.022)
Second study control condition	-0.023	-(0.023)	-0.034	-(0.023)	-0.019	-(0.020)

## **Supplementary Appendix: References**

Gonzalez, Yanilda, and Lindsay Mayka. 2022. “Policing, Democratic Participation, and the Reproduction of Asymmetric Citizenship.” *American Political Science Review* pp. 1–17.