



USAID Regional Governance Activity (RGA) in Colombia, 2015-2020: Endline Annexes

Authors

Michael G. Findley
Professor, University of Texas at Austin
Member, AidData Research Consortium
p: 512.749.3815 e: mfindley@gmail.com

Oliver Kaplan
Associate Professor, University of Denver
p: 303.871.2377 e: oliver.kaplan@du.edu

Ana Carolina Marrugo
Ph.D. Student, University of Pittsburgh
p: 202.770.7941; e: acm189@pitt.edu

Alejandro Ponce de Leon
Ph.D. Candidate, University of California, Davis
p: 530.761.3221 e: poncedeleon@ucdavis.edu

Daniel Walker
Independent Data Scientist
p: 661.607.8606; e: dannywalker101@gmail.com

Joseph K. Young
Professor, American University
Member, AidData Research Consortium
p: 850.284.7192 e: jyoung@american.edu

30 April 2021

Table of Contents

Authors	1
Annex 1: Recommendations Actions Matrix	3
Annex 2: Supplemental Analysis and Data	4
Annex 2a: Contextual Data and Results	4
Annex 2b: Component 2 Data and Results	2
Annex 2c: Component 4 Data and Results	6
Public Accountings	6
Citizen oversight.....	10
Annex 2d: All Regression Results	14
Annex 2e: Arauca Specific Dynamics: Endline Relative to Midline	15
Annex 2f: Other Region-Specific Dynamics	17
Annex 2g: Rescaled Financial Indicators	18
As a Proportion of Taxes	18
Other Proportion Graphs	19
In 2019 Pesos.....	21
Annex 3: Findings, Conclusions, and Recommendations Table	23
Annex 4: Detailed Methodology	26
Annex 4a: Municipality Sampling	26
Annex 4b: Analysis	31
Annex 5: Instruments	32
Annex 6: Sources of Information	32
Annex 7: Evaluation Statement of Work	34
Annex 8: Summary Information about Evaluation Team Members	34
Annex 9: Statement of Differences (when applicable)	35

Annex I: Recommendations Actions Matrix

Table AI.1 below details the key recommendations from the endline evaluation along with the targeted audience for the recommendation. We summarize the core recommendation in column 2 and then check boxes for the each of the relevant audiences for the evaluation.

Table AI.1: Recommendations Actions Matrix

No	Recommendation	Audience					
		USAID Mission	USAID HQ	GOC	Other Donors	NGOs/CBOs	Municipalities
	Overall						
1	Encourage local elites to raise awareness of programming among the public	X					X
2	Target future programming to address region-specific gaps in outcomes	X		X			
3	Communicate the links between programming, communication with local officials, and the peace process	X		X	X		
4	Conduct a follow-up evaluation of RGA to assess slower indicators	X					
	Context						
5	Consider the periodicity of mayoral terms when interpreting RGA program effects	X	X				
6	Continue tracking security indicators to see whether they improve as public works activities increase	X		X			
7	Consider pausing programming unless GoC addresses security risks to social leaders	X			X		
	Component 2						
8	Emphasize early program results (revenue/ royalty gains) to gain trust	X	X	X	X	X	
9	Communication strategies should encourage realistic expectations about slower pace of service delivery	X	X	X			X
10	Devote resources to best-managed localities for proper use of resources	X		X			
11	Identify why trainings of municipal staffs may not lead to improved municipal performance	X					X
	Component 4						
12	Prioritize citizen oversight and participation in regions with higher royalties for accountable investments	X		X			
13	Consider burdens that programs may impose on local leaders, such as exposure to security risks	X					
14	Address security risks for local leaders and citizens for safe participation	X		X			
15	Educate officials about the utility of citizen oversight	X					X
16	Encourage / train new oversight committees to monitor COVID-19 assistance funds	X		X			
17	Continue to include ethnic minorities, women, and vulnerable populations in participatory oversight processes	X					X

Annex 2: Supplemental Analysis and Data

Annex 2a: Contextual Data and Results

From the survey, reported victimization increased by midline, but decreased by endline. See Figure A2.1 for the overall results, and Figure A2.2 for results separated out by gender. These figures show that victimization decreased by endline in both control and treatment, though more in control than in treatment. It is important to recall that, at endline, rates of non-response increased especially for some of the more sensitive questions, such as victimization. As such, the decreased rates of victimization may or may not be meaningful. Ultimately, it is difficult to say for sure, but we suspect that the lower rates of victimization were in part due to self-censoring in reporting. In the elite interviews, potential respondents were, indeed, reluctant to participate due to security concerns. To the extent the results are meaningful. Note that at endline women were victimized at higher rates in treatment relative to control, and the opposite was the case for men. This suggests that, contextually, the situation was worse for women than men at endline, and that the RGA municipalities faced particularly difficult circumstances for successful implementation.

Figure A2.1: Victim of Crime Across All Waves of the Evaluation

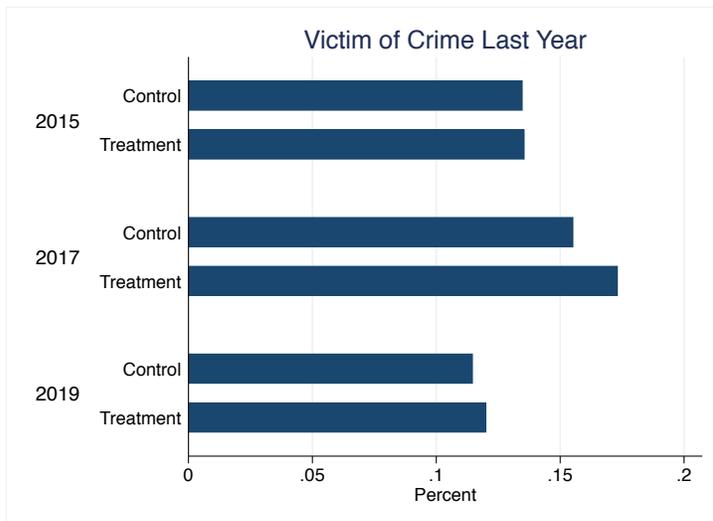
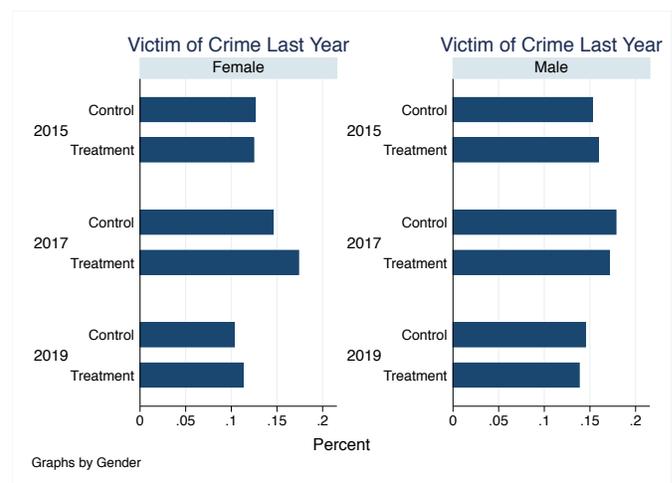


Figure A2.2: Victim of Crime by Self-Reported Gender



Figures A2.3 and A2.4 show homicide rates by region across the waves of the evaluation. Homicide rates are much higher in treatment municipalities relative to controls across all three waves, and homicides appear to be increasing in treatment municipalities while decreasing in controls. Regionally, Valle de Cauca, Antioquia, and Putumayo experienced substantial decreases in homicides from baseline to endline, suggesting important improvements, whereas Nariño and Cauca saw large increases. This is consistent evidence from interviews with community leaders. Across treatment and control municipalities, leaders stated feeling insecurity had increased in their municipalities, especially for social leaders who were increasingly being threatened. Municipal administrators were also more reluctant in the endline than in the baseline and midline to talk about security conditions due to security concerns.

Table A2.2: Do armed groups impact mayor's office performance? Municipal administrators

	Control			Treatment		
	2015	2017	2019	2015	2017	2019
No	48.9%	70.8%	30.8%	55.3%	64.5%	16.7%
Yes	44.4%	29.2%	61.5%	38.2%	25.8%	50.0%
NA	6.7%	0.0%	7.7%	6.6%	9.7%	33.3%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Given the difficult security environment and mixed results, we also examined the basic correlations of key outcome indicators and the security environment. Table A2.3 shows those correlations and highlights (shading and *) any in which the correlation is higher than 0.3 (or -0.3). The 0.3 threshold is arbitrary but draws attention to what may be considered possibly important relationships. It appears that the safer people feel, the more likely they are to report greater access to services (0.4106) and the better the security conditions the more likely people are to report involvement with participatory budgeting (0.3685). Predictably, higher feelings of safety are correlated with a lower likelihood of being a victim of crime last year (-0.324), two different measures of a similar dynamic. Finally, the trust in the mayor's office is associated with not being a victim of crime.

As with many of the other results in the endline, the results here are mixed, but do suggest that security dynamics did, in fact, complicate participation in some important ways. Indeed, even those indicators with lower correlations seem to trend in this direction. For example, ease of participation in political life is positively correlated with better security conditions. Being a victim of crime, moreover, tracks closely with feelings of safety, trust, and security conditions. These results contribute to the larger theme in the endline that pronounced positive results of RGA are few, and despite some initial improvement in key areas, by endline effects were not strong and likely compounded by a difficult contextual environment.

Table A2.3: Correlations between Security and Key Program Indicators

	Feelings of safety	Security conditions	Mayor's office
Property tax	0.1631	-0.0347	-0.134
Access to services	0.4106*	0.0438	0.2647
Easy to participate in political life	0.1552	0.2606	0.1712
Discouragement to oversight	-0.1807	0.0025	-0.1114
Participatory budgeting	0.1249	0.3685*	-0.1573
Municipal responsiveness	0.2582	0.2188	0.2619

	Feelings of safety	Security conditions	Mayor's office
Peace process support	-0.0402	0.0275	0.1142
Armed groups move freely	0.031	0.022	-0.0138
Homicide rate	0.0006	-0.232	-0.1762
Victim of crime last year	-0.324*	-0.2411	-0.4082*

Annex 2b: Component 2 Data and Results

Component 2 of RGA aimed to improve municipal fiscal capacity and management. Although there were some improvements, especially with respect to royalties, there were only marginal changes in many of the standard fiscal and management indicators. The main report begins by presenting data on total revenues in control and treatment municipalities across waves as well as by region.

Even if revenues did not increase significantly, municipal administrators feel that the support from RGA has been crucial to understand better the SGR and more efficiently plan and report their financial activity. These skills may bring long-term benefits to the extent that knowledge is effectively passed to future administrations. Qualitative accounts show great enthusiasm for these forms of assistance. As a municipal administrator from Montelíbano, Córdoba described their experience with the program, “RGA program professionals helped us a lot in formulating projects in the SGR. The process brought so many benefits for us, we saved money, we spent it faster.”

Figures A2.5 and A2.6 shows the results for Fiscal Performance both across waves and regions. Across waves, there was almost no change over time. Across regions, changes are mainly observed in Putumayo (declines in performance). Fiscal performance ratings therefore do not seem to have been impacted by RGA activity.

Figure A2.5: Fiscal Performance Across Waves and RGA Activity

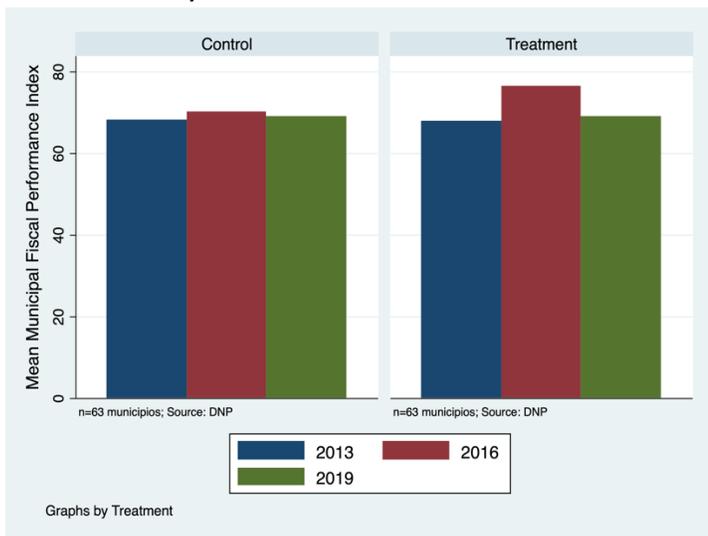
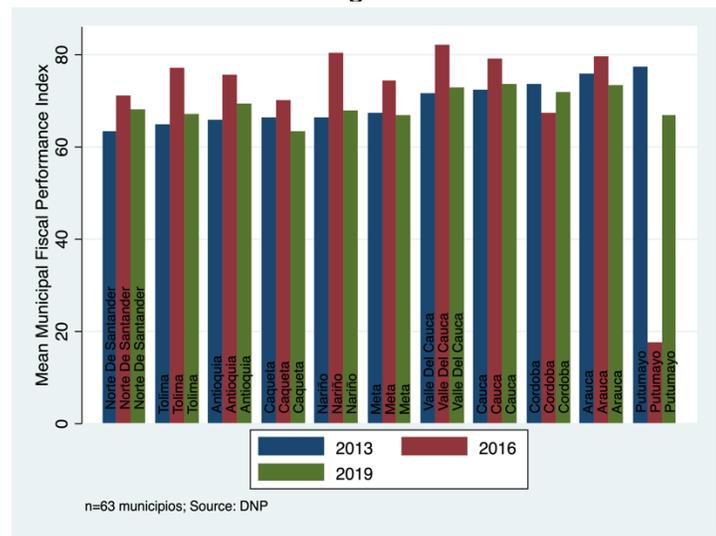


Figure A2.6: Fiscal Performance Across Waves and Regions



Another key indicator is the Municipal Performance Index, reported in Figures A2.7 and A2.8. The results are marginally negatively significant, in contrast to expectations from the theory of change of a positive result, the RGA is associated with no meaningful difference.

Figure A2.7: Municipal Performance Index Across Waves and RGA Activity

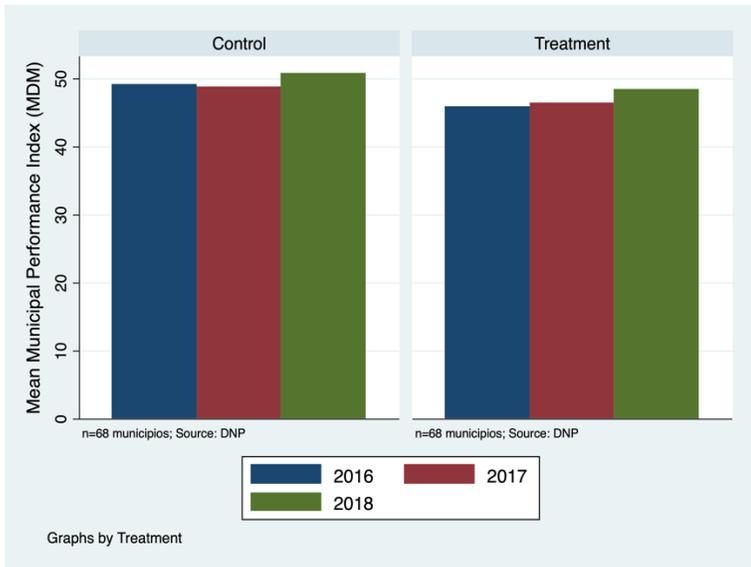
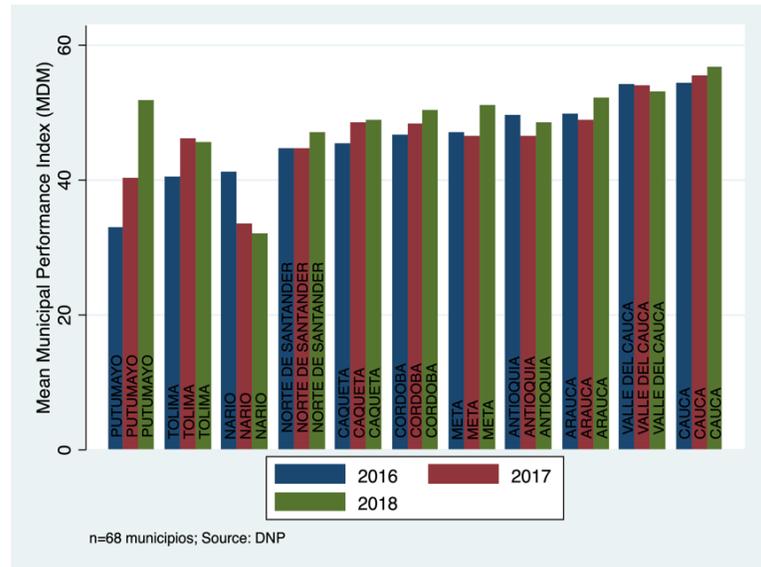
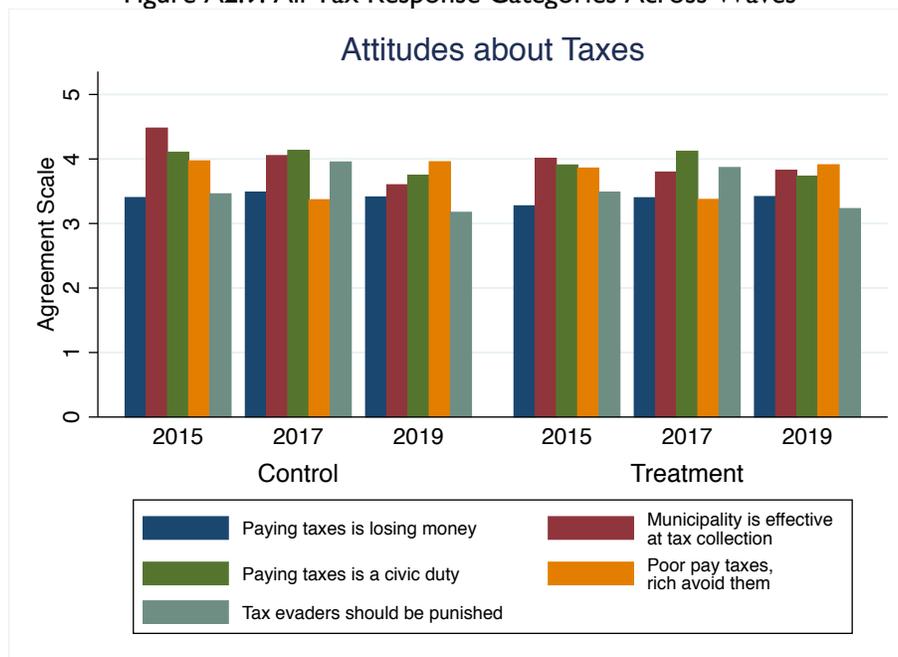


Figure A2.8: Municipal Performance Index Across Waves and Regions



We report some specific attitudes about paying taxes in the main report, but present the full results here. As Figure A2.9 shows, attitudes are relatively constant over time. There are a few differences, but none that are systematically meaningful in treatment relative to control. Thus, the RGA program, while focused on a number of outcomes related to financial management, did not seem to have widespread effects on citizen attitudes about taxes. That includes the five separate indicators of perceptions about taxes reported in Figure A2.9.

Figure A2.9: All Tax Response Categories Across Waves

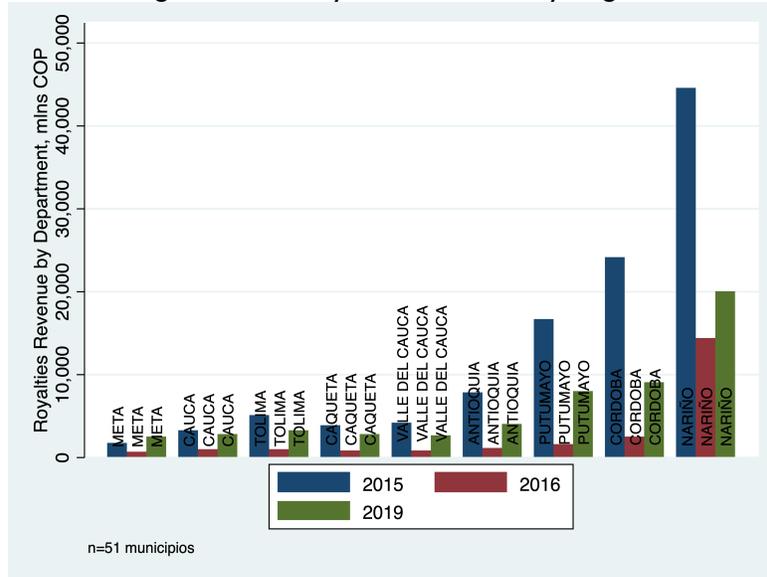


Focus group participants and community leaders agree that the lack of information on property and land use is a central factor inhibiting a positive “tax culture” in municipalities. Community leaders usually point

to the lack of “civic culture” and citizens’ limited comprehension of their duties as impeding tax collection. Lay citizens in focus groups more frequently mentioned the precarious economic conditions in small municipalities as a main reason for low tax payment rates. Finally, municipal administrators emphasized the positive impacts of educating taxpayers for increasing municipal revenue.

By region, royalties revenues declined in most regions across the waves of the study. In some cases, such as Nariño, the declines are substantial. See Figure A2.10. As discussed in the main report, there were relative improvements in treatment versus control in spite of these regional and overall trends.

Figure A2.10: Royalties Revenues by Region



Taking a broader look at tributary revenues suggests no meaningful differences *between treatment and control municipalities*. Similar to property revenues, tributary revenues increased from the baseline to the later stages across the board, but more in control than in treatment. This increase was also uneven across departments. See Figures A2.11 and A2.12.

Figure A2.11: Tributary Revenue by Wave and RGA

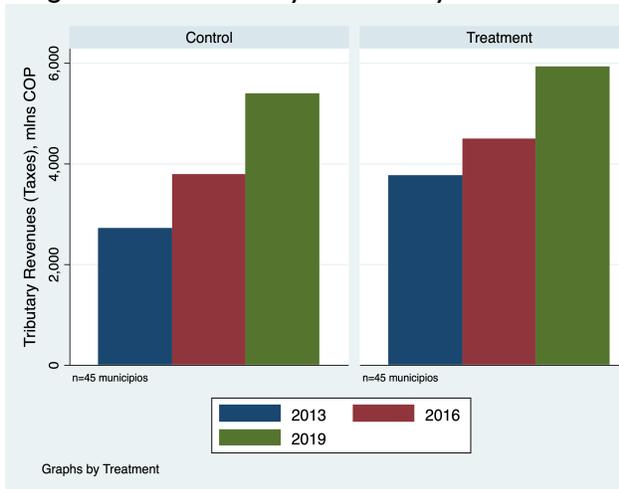
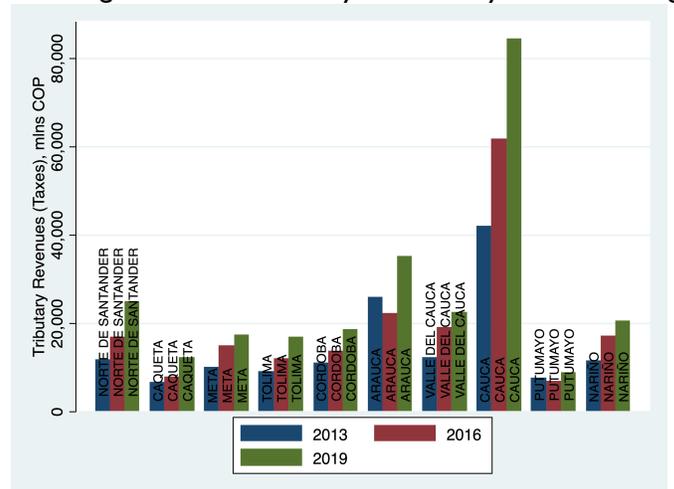


Figure A2.12: Tributary Revenue by Wave and Region



Expenditures showed little change in both control and treatment over time, and the results are not significantly meaningful comparing treatment and control. The regional graphs also do not show major differences across regions, with the notable exception of Nariño. See Figures A2.13 and A2.14.

Figure A2.13: Total Expenditure by Wave and RGA Activity

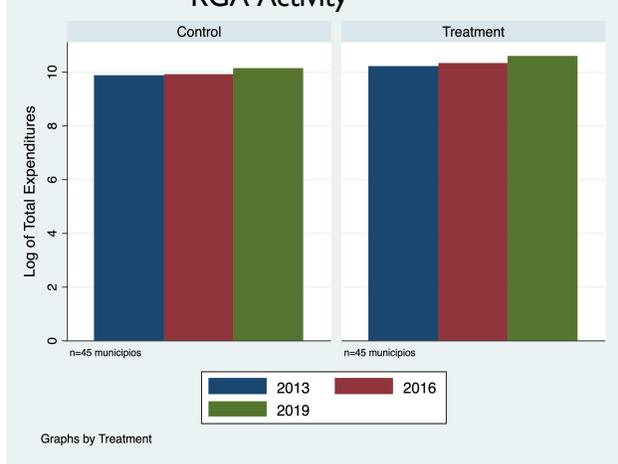
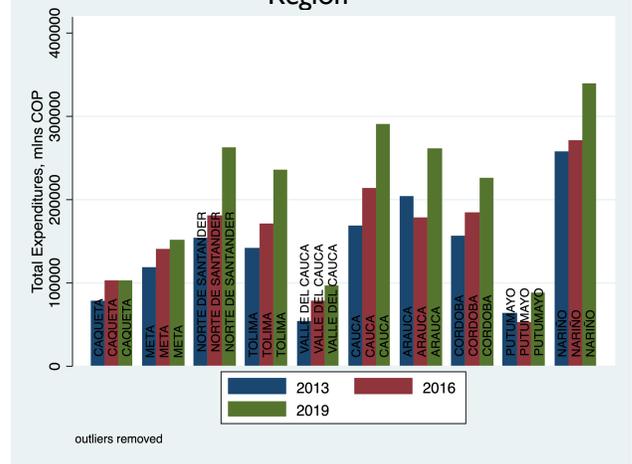


Figure A2.14: Total Expenditure by Wave and Region



Figures A2.15 and A2.16 display deficits and surpluses over time and across municipalities. In each graph pane, both deficits and surpluses are shown for each evaluation wave. The two far-left bars in the left pane of Figure A2.15 show the deficits and surpluses in control and treatment municipalities before the baseline. In the left pane, deficits and surpluses change within control municipalities. Comparing the left pane to the right pane, deficits and surpluses are varying across control and treatment and over time. The graphs show that both deficits and surpluses declined over time within control and within treatment, which is both good and bad—while deficits declined, surpluses also declined, dampening net gains in revenues. By the endline, deficits in treatment municipalities were higher than in controls and surpluses in controls were higher than in treatment. While not statistically meaningful, these results are not what would be expected from the perspective of good financial management.

Trust and Legitimacy

Our interviews, surveys, and focus groups do not show significant changes in levels of trust in government or non-government institutions, either across municipalities or between baseline and endline measurements. See Table A2.4. However, interviews show that, at the endline, participants had additional information with which to assess mayoral offices, and factors such as corruption and unfulfilled promises stand out. Community elites' level of trust in local government decreased from baseline to midline and midline to endline in treatment and control municipalities alike. In the baseline measurement, most mayors had just started their terms so the early positive perceptions may have reflected signs of hope. At the endline, as mayors were finishing their mandates, communities saw little or no improvement in conditions, which likely undermined trust in the local government.

Table A2.4: Do you trust the local government? Community Elite

	Control			Treatment		
	2015	2017	2019	2015	2017	2019
No	16.55%	33.3%	37.2%	16.83%	31.4%	29.5%
Yes	82.07%	66.7%	50.4%	77.23%	66.9%	44.5%
NA	1.38%	0%	12.4%	4.03%	1.7%	26.0%*
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

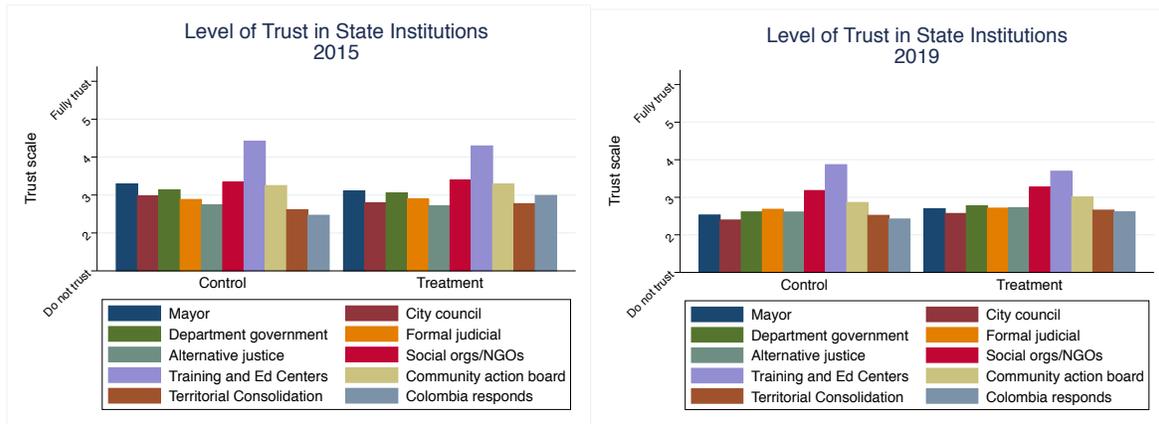
*Note that many interviews did not ask this question at endline prior to the mayoral elections. Some interviews were therefore conducted after mayoral elections, when respondents were less comfortable answering reporting on a new mayor.

Trust and legitimacy are also reflected in indicators about trust in armed actors, trust in institutions involved in conflict resolution, and with respect to the impact of illegal armed actors on municipal performance. Citizens tend to trust more in legal institutions and tend to turn to them (or other citizens) to solve disputes. Surveys and community elite interviews show that, over time, communities became slightly more aware of the impact of illegal armed groups on municipal performance.

We also asked citizens about their trust in a variety of state institutions. The data indicate that people have mixed levels of trust in most state institutions. In the baseline, midline, and endline evaluations, training and education centers, such as SENA, stand out as being the most trusted institutions, as reported in Figures A2.19 and A2.20. Overall, the results in these figures show a slight decline in trust across institutions in Colombia from baseline to endline, rather than improvement. Colombia Responde is among the institutions with the lowest levels of trust, according to the survey data. It is likely that increasing trust in institutions is a process that takes time, especially given the low baseline from which most trust levels began.

Trust declined a little more in some treatment municipalities relative to control municipalities. For example, 19% of respondents had scored trust in the mayor as high (5 or 6) in the control condition at endline, a decline by 4 percentage points since baseline. This is compared to 10% in the treatment group that scored trust in the mayor high at endline, a decline in 9 percentage points. Even for the highly Training and Education centers there were declines in trust that across control and treatment. Finally, in some areas like social organizations/NGOs, the level of trust across time and control/treatment stayed almost the same.

Figures A2.19 and A2.20: Level of Trust at Baseline and Endline



Government and community elites shared their thoughts about trust through the interviews. In focus groups and interviews, the mayors’ offices were at the center of the discussion about trust and legitimacy in state institutions. As in the baseline, corruption was a frequent reason for distrust in local government. Nevertheless, at the endline, participants had new information on which to judge mayors’ trustworthiness, given mayors had been in office for a longer period. Focus group participants, for instance, spoke about changes in mayors’ attitudes towards rural communities, earlier making campaign promises only later to fall by the wayside. As a participant from Tibu said:

“Before being mayor, he went to every house and asked about the needs of the people. One of my daughters lost her house one day, the mayor said he was going to help her to get a piece of land to build a new one. Until now he hasn’t delivered on his promise. While he was campaigning, we saw him everywhere, after he was elected he’s never around, he’s always travelling”

Based on the interview and focus group data, we also found that the prime factors that increase citizen trust in the mayors’ offices remained the same as in the baseline: their close relationships with communities and their human qualities. A community elite in Puerto Libertador, Córdoba describes a mayor who they trust in the following way:

“...[the mayor] was born and raised here, he has lived here his whole live and has suffered like us, had the same needs we do. We are in his hands now, we believed in him and we put the administration of the municipality in his hands... he’s a God fearing person and that will help him do things right.”

Another leader from Segovia, Antioquia said:

“Of course [we trust him], he is one of us and he is very effective. He listens to people and really cares for us. He goes to see works done around the municipality and makes sure the police are vigilant as well.”

Given that, at the time of the endline, mayors had been in office for a more extended period, some of the main reasons to trust them were the visibility of public investment and their direct involvement with

community organizations. For instance, in Corinto, Cauca, trust in the mayor came from training programs for Indigenous and Afro communities:

“We trust the mayor because he has developed various trainings for social organizations, peasants, Afro-Colombians, and other vulnerable groups.”

In Angostura, Antioquia another leader justified trust in mayor in the following statement:

“We trust him because we have seen the infrastructure works and the investment. He also supports and works with the JAC.”

Similarly, municipal officials held more positive perspectives about the trust that citizens have in the administration. As in the baseline, most of them believe that citizens trust their administrations, although they do not provide clear evidence of this. Most positive sentiments indicate that actions that enhanced citizen participation also resulted in increased trust from citizens.

A government official from Nechí, Antioquia told us:

“We always had a relation based on trust. We always had the trust of the community and we trusted the veedurías too.”

Another government official from Montelíbano, Córdoba said:

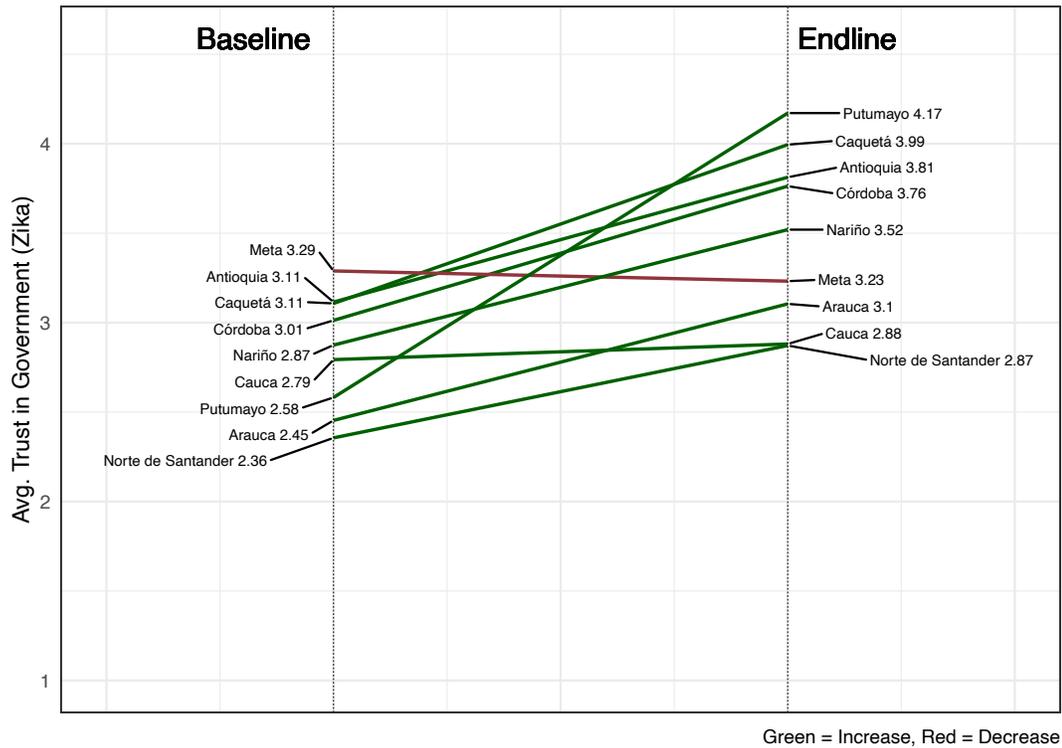
“With the help of the community we gained their trust. The mayor is a very popular person, he meets with the presidents of the JACs, with religious leaders. The community is always willing to participate in debates.”

To understand citizens’ perception of armed actors, we first asked citizens about levels of support for various legal and illegal armed actors (i.e., government army, government police, and illegal armed groups). The results show no statistically higher ratings of government actors, though ratings are not statistically higher for armed actors either. (See full regression results in Table A2.9 below.) Second, we asked people who they would turn to in the case of a dispute. Respondents mostly reported they would turn to neighbors and government institutions rather than illegal armed groups, a result that is largely consistent with the baseline (See full regression results in Table A.2.9 below). Third, we asked how citizens perceive the effect of the presence of illegal armed groups on municipal performance. According to the results in Figure 5 of the main report, presence of illegal armed groups is a core concern that undermines broader trust in the system to work well.

In sum, as stated by RGA, a key hope is that by increasing the capacity of municipalities and departments to deliver services, governance will improve while *trust, credibility and legitimacy will be built*. That said, building trust can be a slow process. The survey data and community interviews suggest RGA municipalities have sustained similar, though not terribly high, levels of citizen trust in government. Citizens trust their local administrations insofar as they feel mayors and their teams are committed to the well-being of the community. Qualitative data show that the engagement of mayors is gauged by the amount of time they spent in their municipalities or public appearances, and by the number of “visible” public works (e.g., road improvement, parks, community centers, etc.). One of the challenges is that citizens may have expectations about government performance that are different from government capabilities. Similarly, corruption and patronage are pervasive forces across many Colombian institutions, and still negatively affect citizen’s trust of mayoral performance in some municipalities.

To make trust more concrete, we considered a specific health issue. While people lacked trust in different levels of institutions, from the municipal to national levels, trust did improve across the regions where the government of Colombia handled Zika well (See Figure A2.21). This is potentially a positive sign as the Colombian government, like other governments around the world, struggle in dealing with COVID19.

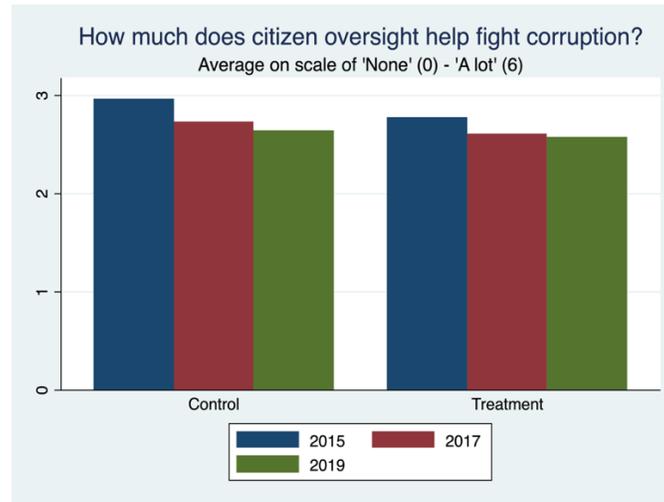
Figure A2.21: Trust in the Government to Handle Zika Across Different Regions



Citizen oversight

In addition to what we consider in the main report (Table 8 and Figure 28), here we report three survey-based measures on citizen oversight processes: whether respondents know about citizen oversight (Figure A2.22), whether they participate in citizen oversight (Figure A2.23), and whether oversight helps with corruption (Figure A2.24). For knowledge about oversight, respondents in treatment areas were slightly more knowledgeable than in controls, but there was not much improvement from baseline to endline in either group. For participation, roughly 20% of respondents in treatment and control participated in citizen oversight, but there were no differences between treatment and control, nor much change over time. On whether oversight bodies help fight corruption, the reported rates at all three waves are high, but they gently decline over time in both treatment and control.

Figure A2.24: How much does oversight help fight corruption



We also asked municipal administrators about their perception of oversight groups (Table A2.6). As discussed in the report, perceptions about these organizations and their contributions to local governance and accountability were positive. However, it was also common to hear disclaimers about how veedurías were “politicized” as well as suspicions that they look for ways to thwart the work of the mayor.

Table A2.6: Percent of coded reference. Perception of oversight groups (veedurías). Municipal administrators

Perception of oversight groups (veedurías). Municipal administrators						
	Control			Treatment		
	2015	2017	2019	2015	2017	2019
Contribute to local governance/improve municipality performance	55.6%	36.6%	69.2%	55.3%	64.4%	50.0%
Provide different point of view	8.9%	2.4%	0.0%	3.9%	0.0%	0.0%
No oversight groups (veedurías) in municipality	4.4%	4.9%	0.0%	3.9%	4.4%	0.0%
Obstacle to municipality's performance	0.0%	4.9%	7.7%	2.6%	2.2%	0.0%
Politicized	4.4%	4.9%	7.7%	3.9%	0.0%	16.7%
They are a citizen's right	15.6%	12.2%	0.0%	10.5%	13.3%	8.3%
Not properly trained	2.2%	9.8%	0.0%	2.6%	2.2%	0.0%
Strengthen participation in municipality	6.7%	4.9%	0.0%	3.9%	4.4%	0.0%
Improve relation between community and local government	0.0%	9.8%	15.4%	0.0%	4.4%	8.3%
Corrupt	0.0%	0.0%	0.0%	0.0%	0.0%	8.3%
NA	2.2%	9.8%	0.0%	13.2%	4.4%	8.3%
Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

With respect to broader participation, we asked community elites about their participation in political life, and how that has increased or decreased. Tables A2.7 and A2.8 demonstrate that fewer think it is easier,

Annex 2d: All Regression Results

Table A2.9: All raw regression results

Description	N	coef	se	t	p	sig_10	sig_5	sig_1
Can Group Move Freely?: Criminal bands	69	0.036	0.062	0.583	0.562			
Can Group Move Freely?: ELN	69	0.05	0.03	1.698	0.094 *			
Can Group Move Freely?: FARC	69	0.199	0.064	3.107	0.003 *	*	*	
Can Group Move Freely?: Paramilitaries	69	0.07	0.034	2.077	0.042 *	*		
Can Group Move Freely?: Guerrillas	69	0.12	0.042	2.84	0.006 *	*	*	
Can Group Move Freely?: Drug traffickers	69	0.052	0.033	1.563	0.123			
Level of trust in: Mayor's office	69	0.055	0.126	0.439	0.662			
Level of trust in: city council	69	0.085	0.089	0.958	0.341			
Level of trust in: departmental government	69	0.127	0.095	1.328	0.189			
Level of trust in: national government	69	0.032	0.092	0.346	0.731			
Level of trust in: formal justice institutions	69	0.032	0.078	0.406	0.686			
Level of trust in: alternative justice institutions	69	0.086	0.1	0.86	0.393			
Level of trust in: agricultural institutions	69	0.055	0.103	0.531	0.597			
Level of trust in: social organizations/NGOs	69	0.106	0.114	0.929	0.356			
Level of trust in: training and educational centers	69	-0.001	0.119	0.007	0.995			
Level of trust in: ombudsman	69	0.062	0.074	0.833	0.408			
Level of trust in: national police	69	0.048	0.068	0.711	0.479			
Level of trust in: national army	69	0.015	0.097	0.157	0.875			
Level of trust in: national navy	69	0.037	0.074	0.507	0.614			
Level of trust in: municipal ombudsman	69	-0.003	0.087	0.03	0.976			
Level of trust in: community action board	69	0.093	0.085	1.093	0.278			
Level of trust in: The Unit for Attention and Integral Reparation of Victims	69	0.033	0.066	0.502	0.618			
Level of trust in: the unit for territorial consolidation	69	0.073	0.088	0.838	0.405			
Level of trust in: the land restitution unit	69	0.028	0.058	0.478	0.634			
Level of trust in: Colombia Responds	69	0.193	0.069	2.795	0.007 *	*	*	
Admin capacity: 2013 vs. 2017	69	3.42	6.98	0.49	0.626			
Fiscal performance: 2013 vs. 2019	67	1.259	5.586	0.225	0.822			
Total revenue: 2013 vs. 2019	63	0.481	0.233	2.067	0.044 *	*		
Property revenues: 2013 vs. 2019	62	0.443	0.193	2.291	0.027 *	*		
Taxes: Paying taxes is losing money	69	0.039	0.095	0.414	0.68			
Taxes: The municipality is effective at tax collection	69	-0.01	0.121	0.087	0.931			
Taxes: Paying taxes is a civic duty	69	-0.101	0.087	1.172	0.245			
Taxes: The poor pay taxes; the rich avoid them	69	-0.007	0.109	0.066	0.947			
Taxes: Tax evaders should be punished	69	0.147	0.144	1.024	0.309			
Royalties: 2015 vs. 2019	51	0.401	0.084	4.751	0 *	*	*	
Access to public services: Getting worse	69	-2.058	2.052	1.003	0.319			
Access to public services: Staying the same	69	-1.627	1.969	0.826	0.412			
Access to public services: Getting better	69	-2.27	2.635	0.862	0.392			
Corruption	69	-0.089	0.156	0.573	0.568			
Participation in event: Cultural, sporting	69	-0.001	0.024	0.022	0.983			
Participation in event: Marches, demonstrations	69	0.005	0.018	0.275	0.784			
Participation in event: Improve security	69	0.015	0.015	1.027	0.308			
Participation in event: Political party events	69	0.002	0.017	0.099	0.921			
Participation in event: Government	69	0.017	0.011	1.576	0.12			
Participation in event: Labor strikes, demonstrations	69	0.012	0.01	1.197	0.235			
Participation in event: Fundraising events	69	0.023	0.021	1.104	0.273			
Participation in event: Indigenous marches (Mingas)	69	0.031	0.01	2.943	0.004 *	*	*	
Participation in event: Participatory budgeting	69	0.016	0.011	1.414	0.162			
Responsiveness	69	-0.018	0.112	0.157	0.876			
Victim last year	69	-0.027	0.02	1.339	0.185			
Who can be a leader in your community: A member of the opposition party?	69	0.015	0.037	0.408	0.685			
Who can be a leader in your community: A woman?	69	0.027	0.012	2.213	0.03 *	*		
Who can be a leader in your community: An Afro-Colombian?	69	0.019	0.043	0.434	0.666			
Who can be a leader in your community: A former combatant?	69	-0.045	0.039	1.155	0.252			
Who can be a leader in your community: A young person?	69	0.014	0.025	0.556	0.58			
Who can be a leader in your community: An LGBT person?	69	0.004	0.051	0.074	0.941			
Homicide rate: 2015 vs. 2019	66	3.406	9.363	0.364	0.717			
Solving disputes	69	0.148	0.159	0.931	0.355			
Armed actors	69	0.044	0.05	0.894	0.374			
Renditions: 2015 vs. 2019	49	-0.323	0.498	0.648	0.523			
Deficit: 2013 vs. 2019	62	-2770	2830.9	0.979	0.333			
Tributary revenues: 2013 vs. 2019	62	0.017	0.206	0.084	0.934			
Municipal expenditures: 2013 vs. 2019	62	0.45	0.315	1.429	0.16			
Afraid to travel	69	-0.023	0.038	0.597	0.552			
Rating of services: Services of the Armed Forces (Army, Navy, Air Force)	69	0.007	0.116	0.058	0.954			
Rating of services: Non-State services (legal private security services, etc.)	69	0.296	0.155	1.906	0.061 *			
Rating of services: The formal justice services (prosecution, judges, attorneys, etc.)	69	-0.061	0.09	0.678	0.5			
Rating of services: Alternative justice services (Peace Judges, Equity Conciliator, etc.)	69	0.051	0.087	0.582	0.562			
Drugs and coca response: Should not be grown because they are illegal	69	0.183	0.218	0.84	0.404			
Drugs and coca response: Production positively affects families who grow them	69	-0.035	0.109	0.318	0.751			
Drugs and coca response: Only way for some families to survive	69	0.296	0.154	1.923	0.059 *			
Drugs and coca response: Family and friends think cultivation is bad	69	0.009	0.135	0.064	0.949			
Royalties (new): 2015 vs. 2018	54	6E+09	4E+09	1.446	0.154			
Veedurias: Do you know about citizen oversight?	69	-0.009	0.021	0.454	0.651			
Veedurias: Do you participate in citizen oversight?	69	-0.027	0.065	0.414	0.681			
Veedurias: Does the board help with the oversight of public works?	69	0.097	0.058	1.681	0.097 *			
Veedurias: How much does citizen oversight help fight corruption?	69	0.064	0.129	0.498	0.62			
Municipal performance index (MDM): 2016 vs. 2018	69	-5.539	2.775	1.996	0.05 *	*		

Annex 2e: Arauca Specific Dynamics: Endline Relative to Midline

The department of Arauca faced especially severe security concerns in recent years—a critical contextual factor in understanding broader success. We analyzed Arauca separately because of the limited ability to collect information at the baseline, in particular. We therefore consider midline results as a baseline for Arauca from which changes through the endline are assessed.¹ Given the limited number of observations and waves of data, the results here are limited, and we only report some select results.

While Arauca has been plagued by ELN violence, the midline results suggest views in Arauca do not differ much from the rest of the RGA treatment and control municipalities in Colombia. As Figure A2.25 shows, support in Arauca for the Army and formal justice institutions was moderate and compares favorably to other regions of the country. In the treatment group, support for non-state services and alternative justice is a little higher on average but a little lower for the Army and formal justice.

Figure A2.25: Ratings of State and Non-State Services in Arauca Alone

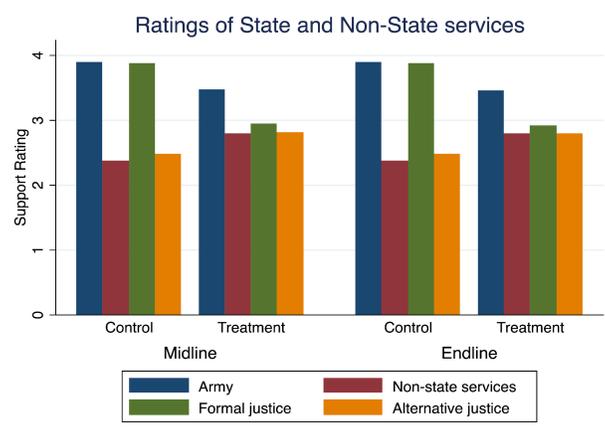


Figure A2.26 illustrates a similar pattern related to crime and victimization as compared to other parts of Colombia. Over 20% of the respondents in the treatment group report being the victim of a crime but that number drops to a little over 15% in the control group.

¹ We could pool the control and treatment municipalities in Arauca with the other control and treatment municipalities in Colombia for purposes of cross-sectional comparisons.

Figure A2.26: Victim of Crime Last Year in Arauca Alone

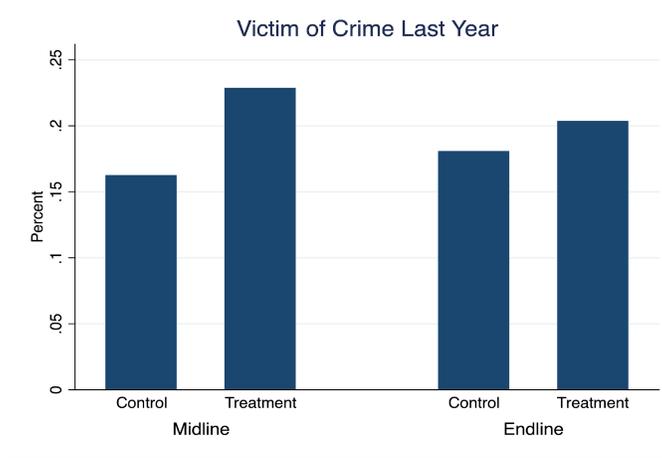
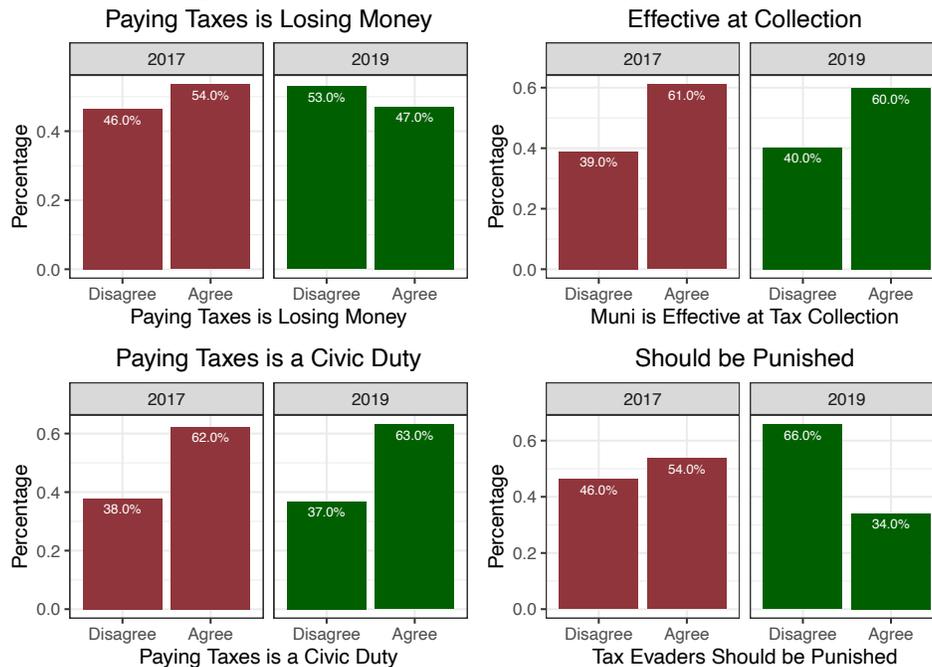


Figure A2.27, which shows several different views on paying taxes, reflects similar results to those for the rest of Colombia. This suggests Arauca may not be substantially different even given the high levels of violence. Views on taxes in Arauca are largely stable. When asked whether the municipality was effective at tax collection or whether paying taxes was a civic duty, views were stable and positive. When asked whether paying taxes is losing money, there was a 7 percentage point reduction from midline to endline. A disappointing result was a 17 percentage point reduction in people who felt tax evaders should be punished. Given Arauca's delicate security situation, we did not conduct any focus groups or interview municipal elites.

Figure A2.27: Attitudes about Taxes in Arauca



Annex 2f: Other Region-Specific Dynamics

We report on each of the following four indicators at other places in the report, but elsewhere by treatment/control and wave status, or by Department. Here we aggregate to the regional level and report the results.

Figure A2.28: Mean Municipal Performance Index by Region and Wave

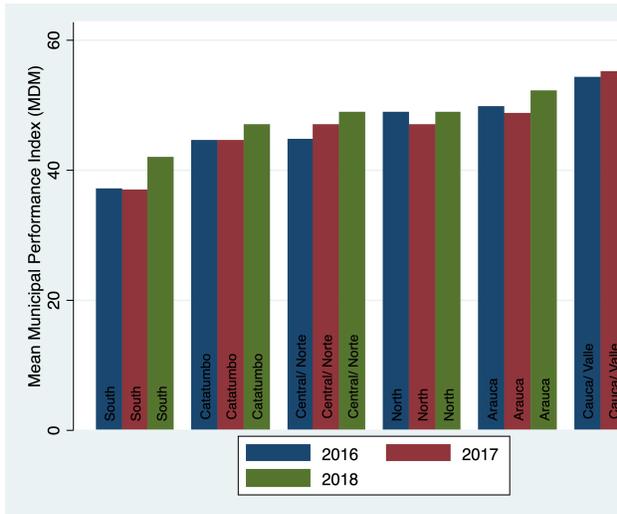


Figure A2.29: Mean Tributary Revenues by Region and Wave

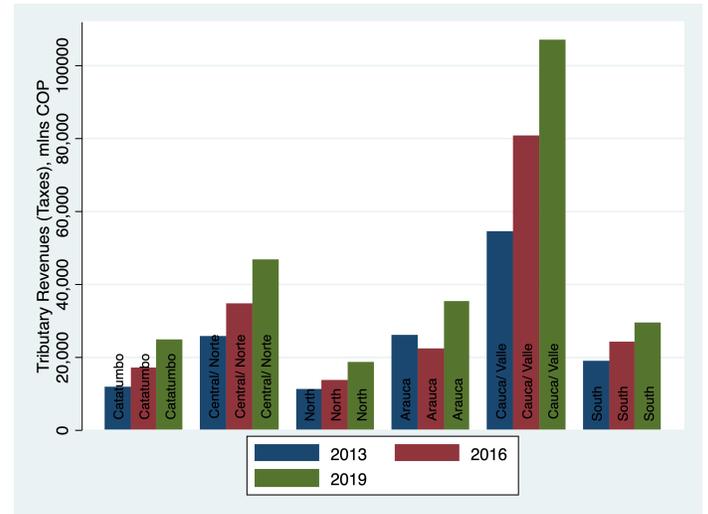


Figure A2.30: Mean Participation in Participatory Budgeting by Region & Wave

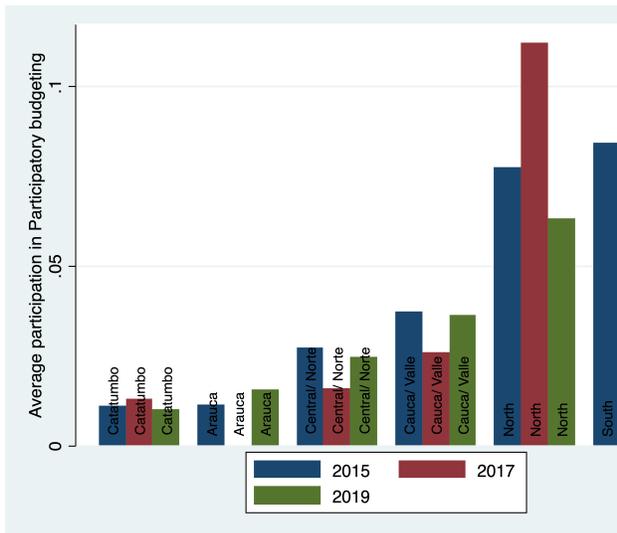
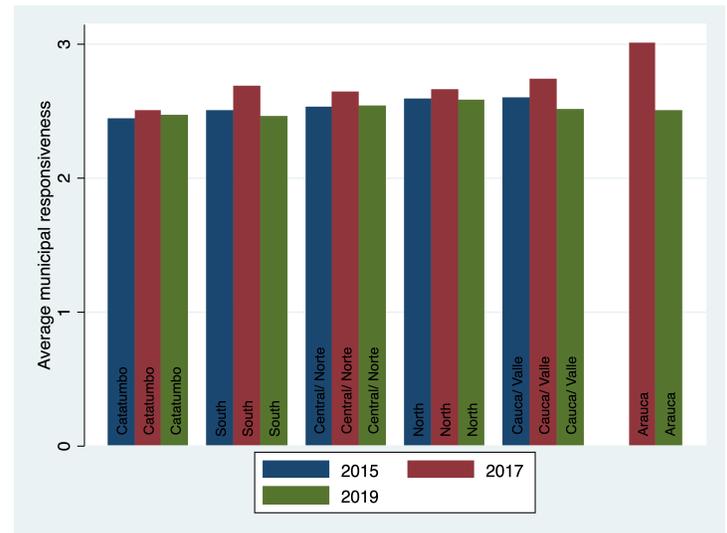


Figure A2.31: Municipal Responsiveness by Region and Wave



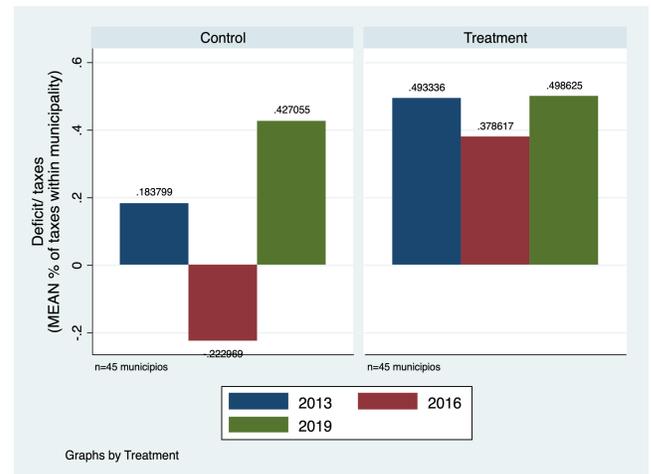
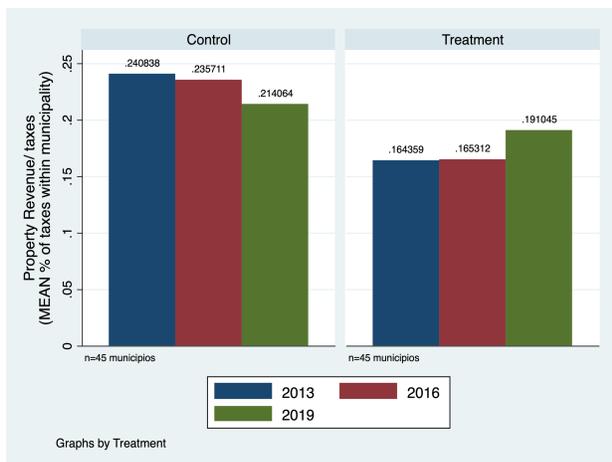
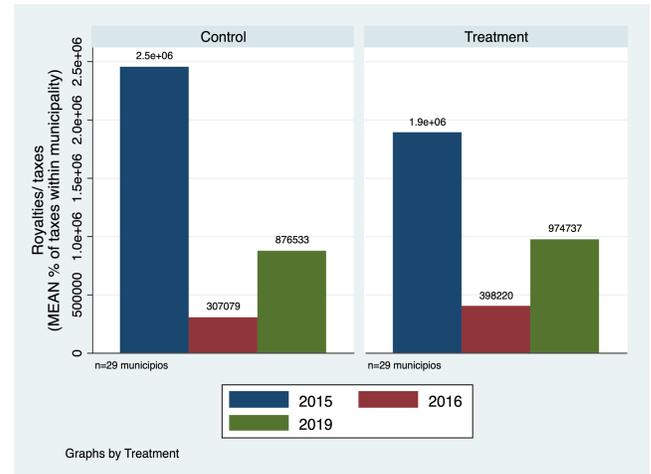
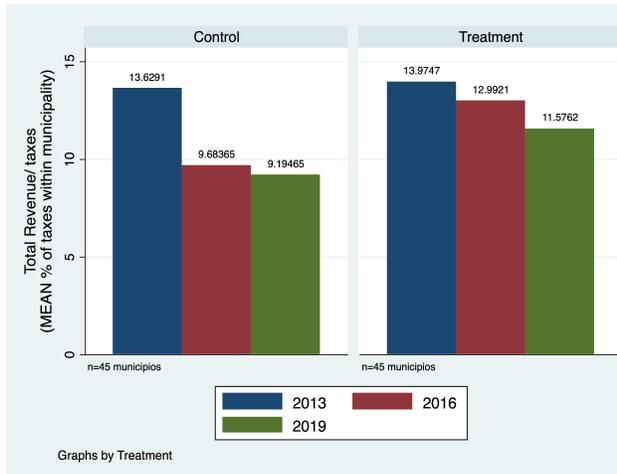
Annex 2g: Rescaled Financial Indicators

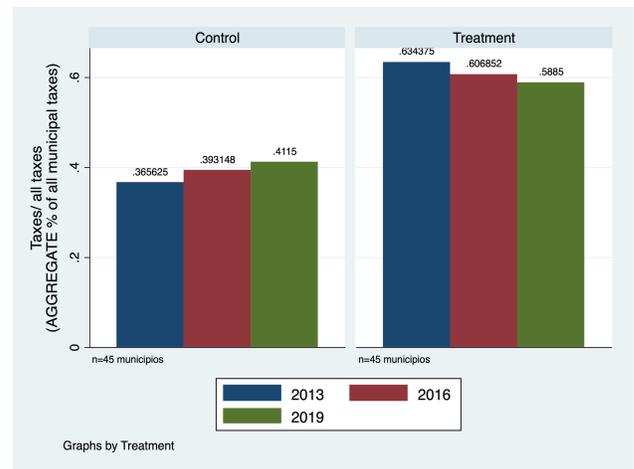
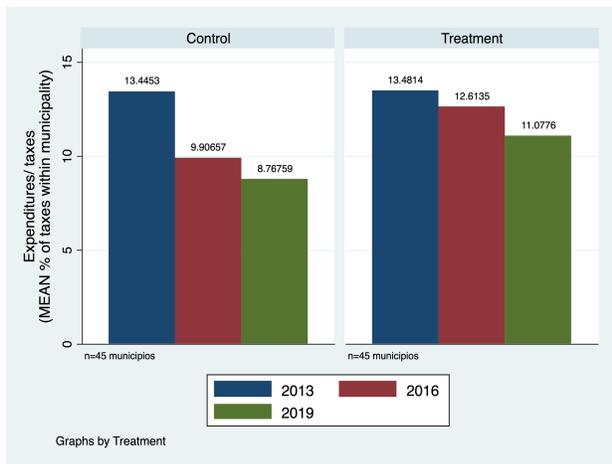
As a Proportion of Taxes

The following figures (A2.32-A2.37) represent each financial indicator as a proportion of local (municipal) taxes. *Total Revenue* and *Total Expenditures* outpace the amount of taxes gathered in each municipality by as much as 13 times, i.e., for each peso of tax gathered, 13 pesos are gathered from other municipal revenue sources. However, this amount declines over time in both Control and Treatment municipalities. This shows both that municipalities are gaining more Total Revenue (as seen in the next section as well), but also that they are gathering more local taxes. And that taxes are increasing at a slightly higher rate.

Note that these graphs show the *averages* of these indicators within years and Control and Treatment groups, i.e., average 2013 Control Revenues outpaced average 2013 Control Taxes by about 13:1, which goes down to about 9:1, on average, by 2019.

Figures A2.32-A2.37



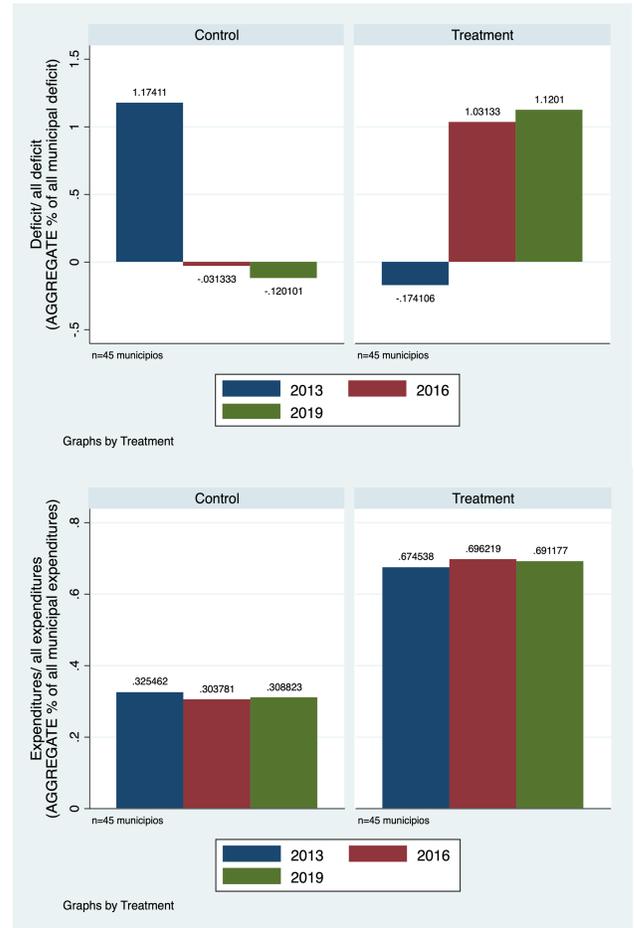
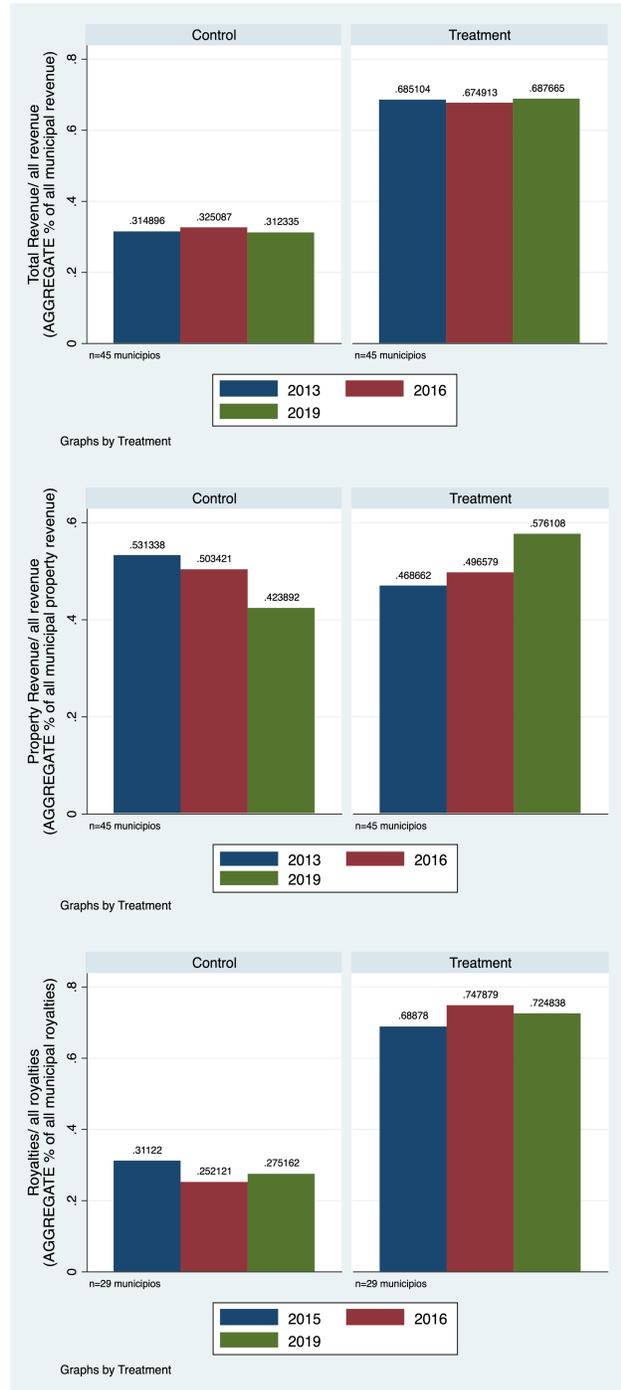


Other Proportion Graphs

This section shows each financial metric as a proportion of the total of that metric, e.g., Revenues as a proportion of Total Revenues. See Figures A2.38-A2.42. Here it is interesting to note that Treatment municipalities make up about 2/3 of the Revenues, Expenditures, Royalties, and Taxes found in the entire dataset (45 municipalities). However, as evidenced by the first section, financial metrics in proportion to taxes are still comparable.

Another interesting trend is that, although the Control municipalities represent the majority of deficits in 2013, this trend flips to the Treatment municipalities in years 2016 and 2019 – this pairs with the high levels of deficits in Treatment municipalities in 2016 (shown below) and the increasing levels of Total Expenditures in Treatment municipalities from 2016 to 2019 (also shown below).

Figures A2.38-A2.42

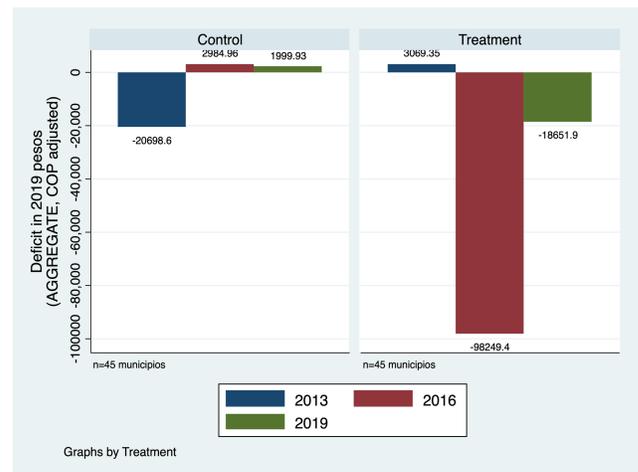
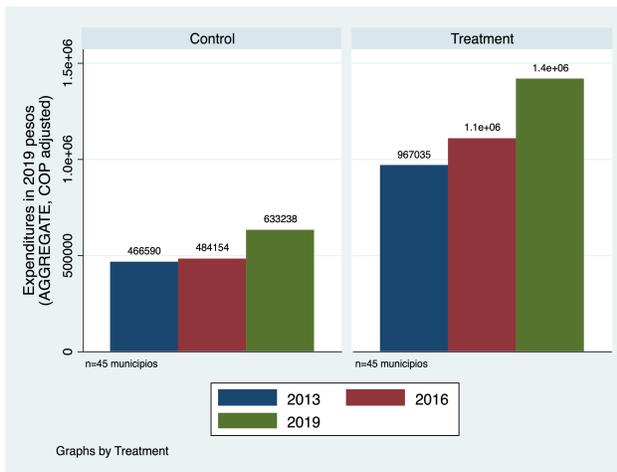
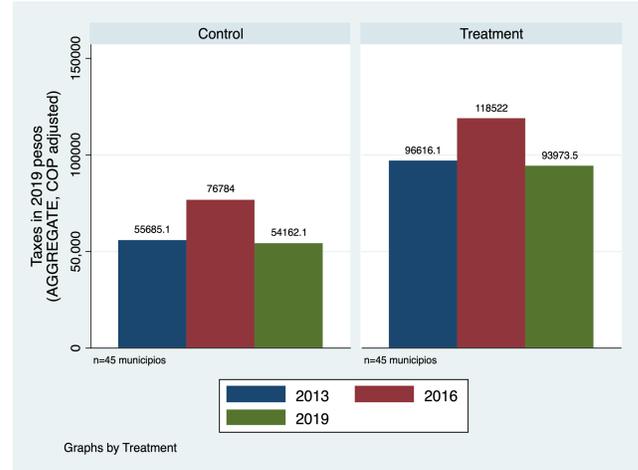
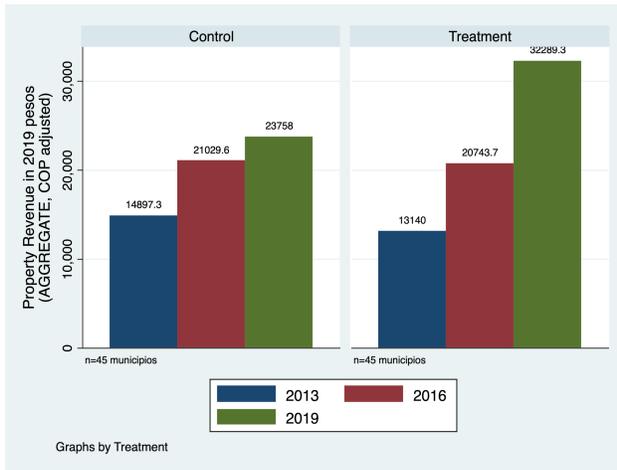
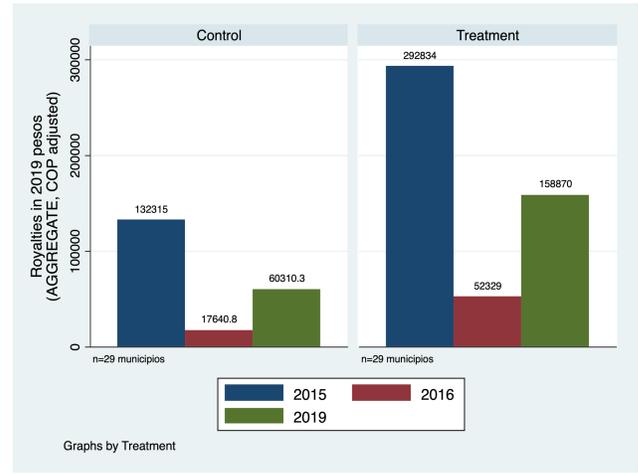
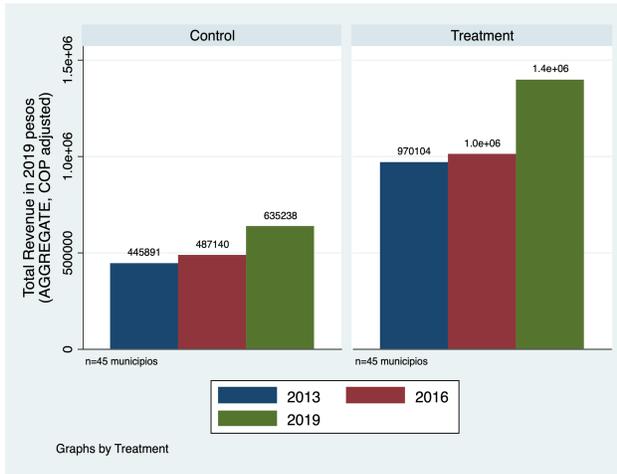


In 2019 Pesos

The following are *aggregate* graphs to show the total financial indicators (in 2019 pesos) by municipality across Control and Treatment, and across years. Here we can see what was hinted at above – Total Revenues and Expenditures are increasing by year, but only slightly; the only exception to this being the jump in Expenditures and Revenues in the Treatment municipalities from 2016 to 2019.

These changes are outpaced by the increases to Taxes, which saw an impressive increase in both Treatment and Control municipalities from 2013 to 2016 (also note that taxes are characteristically higher, on average, in Treatment municipalities). These taxes go back down in 2019, *but* this couples with the spike in Revenue and Expenditures, meaning that these metrics in proportion to taxes still go down, as evidenced by the last section. Also of note: there were considerably more deficits in 2016 in the Treatment group of municipalities.

Figures A2.43-A2.48



Annex 3: Findings, Conclusions, and Recommendations Table

Table A3.1 disaggregates the various categories to give a more detailed perspective of changes in context or impact over time and across treatment and control municipalities. Each row captures aspects of context and the two components and the caption explains exactly which indicators were used to proxy the concepts.

Table A3.1: Findings, Conclusions, and Recommendations Summary Table

CATEGORIES	FINDINGS	OU	IM	CONCLUSIONS	RECOMMENDATIONS
FINANCIAL, PERFORMANCE					
A. REVENUES					
1. Total municipal revenues (\$)	There are greater overall revenues across municipalities including between treatment and control	+	+	There are increases in overall revenues, property tax revenues, and royalty revenues.	Property and royalty revenues should be emphasized early to improve trust and help achieve later downstream outcomes, such as combatting <i>corruption</i>
2. Property tax (\$)	There are greater property tax revenues across municipalities and in treatment relative to control	+	+		
3. Tributary Income (\$)	There are overall increases in tributary revenue, but not statistically different between treatment and control	+	=		
4. Royalties (\$)	RGA municipalities received more net royalty revenues (RGA > Control)	+	+		
5. National Transfers (\$)					
B. INVESTMENTS					
6. Municipal investment (\$)	Fewer residents believe the budget is well invested	=	=	RGA generated increases in expenditures on revenues, though there has not yet been a related increase in perceptions of responsible investment	Devote efforts to communication with citizens about investment. Devote resources to the best-managed localities to maximize proper use of resources.
7. Investment with royalties (\$)	There were some increases in royalty expenditures, though not significantly different	+	=		
C. SERVICE PROVISION					
8. Access to public services	Citizen perceptions indicate steady access to services, few or modest changes Roughly 25 percent of veedurías were in each of the areas of education, health, and road projects	=	=	There was little change in access to or ratings of public services	Encourage realistic expectations among the public about the timing of improvements in quality of life.
9. Ratings of public services					
D. MUNICIPAL PERFORMANCE					
10. Fiscal Performance (IDF, %)	Some improvements in Municipal Fiscal Performance pre midline, but effects attenuated at end line.	=	=	There was minimal improvement in Fiscal Performance and signs of decreasing impact of RGA on Municipal Performance	Identify why trainings of municipal staff members may not lead to improved municipal performance
11. Municipal Performance (MDM, %)	Municipal Performance increased across the board, but in relative terms declined in RGA compared to Control	+	-		
12. Performance of local government	Declining performance of mayors Community elites confirm poor perceptions of mayors	-	-	Perceptions of mayors poor	Continue tracking performance indicators

CATEGORIES	FINDINGS	OU	IM	CONCLUSIONS	RECOMMENDATIONS
13. Financial Contextual Factors	Reform royalty distribution formulas (Sep, 2020) ²	NA	NA		
PARTICIPATION, TR., ACCOUNT.					
E. CITIZEN PARTICIPATION					
14. Easy to participate in political life	No changes in perceptions of ease of participation in politics Increasing security threats impeded participation Perceptions of closed political systems where one needs a political patron and a certain economic status to run for office	=	=	Political participation may have been stimulated by program activities but remains low due to security risks	Improve mayoral transparency and address security risks to community leaders to incentivize participation
15. Political participation	Participation levels did not increase (RGA = control)	=	=		
16. Participatory budgeting	No changes in participation in participatory budgeting	=	=		
F. CIVIC OVERSIGHT					
17. Citizen Oversight	-RGA encouraged many new veedurias -Participation in citizen oversight committees (veedurias) was mixed -Perceptions that the JAC board improved oversight increased -There is some <i>mismatch</i> between oversight activities and the location of new royalty-based projects -Some mayors' office staff perceive the veedurias as "politicized" and impediments to the mayor's work	+ = + = -	= = + = -	Some oversight committees were activated to ensure funds were well-spent and projects well-managed	<u>Continue</u> to prioritize citizen oversight and participation to ensure accountable public investments; educate officials about the utility of citizen oversight
18. Motivations to do oversight	The most important motivations for participating in oversight groups were ensuring projects meet community needs and municipal budgets are well spent	NA	NA		
19. Discouragement to do oversight	Some officials view veedurias as threatening to local politicians Security threats may inhibit participation	- -	- -	Security risks hurt engagement in oversight activities	Ensure a secure environment for oversight
G. MAYORS ACCOUNTABILITY					
20. Public accountings	Number of public accountings decreased and fell more precipitously in RGA areas Citizens report that their mayors do not adequately inform communities about their activities Citizens reported receiving little information on investment of public funds, tax collection. "We have no idea about financial activity of the municipality"	- - -	- - -	Public accountings are still seen as ineffective, and decreased over time.	Do substantially more to encourage mayoral public accountings; current efforts not working
21. Municipal responsiveness	Citizens feel they have <i>limited impact</i> on municipal management	=	=		
22. Municipality takes community into account	Community leaders complain that <i>municipalities are not resolving</i> their most urgent needs	-	-		
POLITICAL CONTEXT	Local elections 2015 and 2019	NA	NA		

² [DNP. Reform to the General System of Royalties. Sep., 2020](#)

CATEGORIES	FINDINGS	OU	IM	CONCLUSIONS	RECOMMENDATIONS
	New presidential administration 2018				
TARGETING					
H. REGIONS					
23. Regions: Property tax (\$)	Most regions increased by endline, with Cauca increasing the most	+	+	Taxes and royalties changed somewhat uniformly across regions	Consider the region-specific gaps in outcomes as well as the unique contextual challenges of implementing programming
24. Regions: Royalties (\$)	Most regions experienced decreases in royalty revenues	-	-		
25. Municipal responsiveness	RGA associated with municipalities being perceived as more responsive in Catatumbo and Central regions	+	=		
I. WOMEN					
26. Who can be a leader	There is high acceptance of women as leaders across municipalities	+	+	RGA helped elevate some underrepresented individuals as possible leaders in communities	Continue oversight trainings both for purposes of oversight but also for empowering underrepresented groups
27. Women and Oversight	Oversight trainings were also opportunities to build new leadership in municipalities	+	+		
28.					
J. ETHNIC COMMUNITIES					
29. Ethnic C. and Oversight	Participation increased in Indigenous Mingas, but this was not a result of RGA	=	=		
30. Ethnic C. and Participatory budgeting	Participation increased but not in participatory budgeting	=	=		
TRUST AND LEGITIMACY					
K. TRUST					
31. Level of trust in state institutions	Trust in institutions remained steady	=	=	Lack of transparency remains highly salient for participation and trust in local government	To boost trust and legitimacy, local elites should be encouraged to play a greater role in raising awareness about RGA-style programs with the general public
32. (Perception of) Corruption	Perceived levels of corruption remain high but are decreasing (RGA = Control)	+	=		
L. LEGITIMACY					
33. Confidence in Mayor's office	Insecurity and being a recent victim of a crime is correlated with less confidence in the mayor's office	-	=	Legitimacy: almost no change	Efforts need to be given not only to greater tax collection, but to changing the culture surrounding taxes
34. Perception of municipal performance					
35. Attitudes about payment of taxes	Little change in attitudes about payment of taxes General erosion of positive cultures of taxation Modest evidence in support of a better tax culture	=	=	Two issues continue to hinder tax collection: informality of land ownership and "tax-evasion" culture	
36. Paying taxes is a civic duty	Community elites suggest tax collection is not effective	-	-		
SECURITY CONTEXT					
M. OUTCOMES C1, C2 AND SECURITY					
37. Participation - Security	Security issues posed obstacles to participation and programming Security conditions are correlated with increased participatory budgeting	NA	NA	Security continues to be a concern, and participation in program-related activities were perceived by some as increasing risks	The security risks of local social leaders should be addressed so they can more safely and fully participate
38. Oversight - Security	Residents were still too afraid to participate	NA	NA		

CATEGORIES	FINDINGS	OU	IM	CONCLUSIONS	RECOMMENDATIONS
	Veedurias may sometimes be perceived as dangerous Stigmatization of leadership and involvement with oversight activities was an obstacle				
N. IMPACT ON SECURITY (short term)					
39. Change in security conditions	Citizens and community elites see the decline in <i>security conditions</i> as the result of a failed peace process	NA	NA	Program does have discernable effects on security	Continue tracking security indicators to see if they improve as public works activities increase
40. Change in feelings of safety	Perceptions of safety declined at endline; RGA not associated with feelings of safety	NA	NA		
41. Impact of armed groups on Mayor's	Perceptions that armed groups impact decrease over time across control and treatment	NA	NA		
CONTEXT OF SECURITY					
42. Peace process support	Steady skepticism about the peace process Most citizens feel the peace process does not take into account their views	NA	NA	The peace process alleviated fears in some municipalities, creating greater openings for civic organizing and political participation	To boost sentiments of inclusion, communicate the links between programming, communication channels with local officials, and the larger peace process. Withhold aid from the Colombian government unless citizen security is prioritized
43. Armed groups free to move in municipality	Presence and activity of armed groups remains a concern	NA	NA		
44. Afraid to travel	Citizens remain fearful of traveling within their municipalities				
45. Victim of crime last year	Victimization rates decreased slightly in RGA areas relative to controls	NA	NA		
46. Homicide (rate)	Homicide rates declined slightly, like a result of the peace agreement. Homicides have spiked in some areas, especially in the South and Cauca regions	NA	NA		
47. Attitudes towards drugs and coca	There are few identifiable shifts in attitudes towards drugs, coca, or legality	NA	NA		

Annex 4: Detailed Methodology

The technical approach in this evaluation is extensively detailed in the original statement of work and then updated in the baseline and midline reports (and their annexes). For full methodological discussion, see those documents. Below, we review the core components of the methodology with attention to required adaptations introduced at the endline. We note that municipality sampling was one of the most important considerations as we needed to identify a valid set of counterfactual municipalities that could be compared with the RGA municipalities.

Annex 4a: Municipality Sampling

USAID selected 40 municipalities, from an original pool of 58 that were placed on the government's national territorial consolidation plan, to receive the RGA. The Government of Colombia (GOC) provided the baseline pool, all of which had a history of violence and weak local capacity. GOC at some level

operated in the other 18 municipalities. As the RGA-intervention municipalities were already selected, the impact evaluation was quasi-experimental and focused on the selection of a useful set of “control” municipalities to compare to the “treated” RGA municipalities.³

To make appropriate comparisons, we rigorously defined a counterfactual (what would have occurred had the intervention not been implemented) and balanced possible confounding factors that could, in theory, explain the outcomes of interest. Standard regression techniques attempt to control for alternative claims, but are not well-suited to isolate the causal impact of any particular intervention because some of the control units may not actually be comparable. That is, the set of municipalities in the treated and control sets may vary in too many ways that makes a controlled comparison impossible. Without such control, establishing the causal impact of a particular intervention becomes impossible.

We preprocessed the data such that the RGA municipalities were matched on all possible dimensions to a set of comparison municipalities so that possibly confounding factors were fully balanced. That is, we had a very large pool of possible municipalities (from the approximately 1,100 overall municipalities) and used computational algorithms to identify the similarities with the RGA municipalities so that distributions of values for possible confounding factors were as similar as possible or what is termed *balanced*. The effect was that the sample, at least on observable characteristics, appeared similar to a randomized sample.

To accomplish this, we first divided cases into units with and without USAID planned intervention. Within the untreated units – a large set of possible comparison cases – we identified municipalities comparable to those that received treatment. Choosing municipalities that were not comparable would offer little analytic leverage and, instead, would muddy the inferences that we can make. To properly match treated and untreated units for analysis, we utilized a technique called covariate-balanced propensity score matching.⁴

We utilized covariate balancing propensity score matching (CBPS),⁵ which permitted us to optimize the balance among covariates, such as region, and then modeled assignment to treatment. Propensity scores assigned to an observation a conditional probability of assignment to the treatment group, given a set of covariates.⁶ A logit or probit model (treatment=1, control=0) was estimated to predict this probability, and the propensity score equaled this probability. The covariate-balanced propensity score approach has an advantage over other techniques because it is a single, nonparametric score that should be unbiased as long as important causal factors are observable.

As a general rule, with matching models it is best to include as much information as is available. Thus, there are no requirements for simple or *parsimonious* matching models. Instead, we identified as many factors as possible that could even plausibly influence the results – intermediate outcomes or overall mission objectives – and included them in the matching models.

³ The set of municipalities identified for treatment and control did not change over the life of the project. That said, security conditions sometimes precluded operating in collecting data in all of the originally identified municipalities. Thus, while we made every effort to preserve the integrity of the treatment and control comparisons, the information was sometimes limited. We note in the different evaluation reports – baseline, midline, and endline – where those data collection challenges entered in.

⁴ There are many types of matching, each of which produces fairly similar results, and we could have utilized a different matching approach. But the covariate-balanced propensity score matching approach is a sophisticated approach, minimizing the problems of many past approaches.

⁵ Imai, Kosuke, and Marc Ratkovic. 2014. "Covariate balancing propensity score." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 76 (1): 243-263.

⁶ Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 24:295–313.

In the end, we utilized the covariate information listed in Table A4.1, which included information on past violence, population, region of municipalities, the Government's fiscal performance index, the open government index, and where USAID is already operating currently and formerly. We also included other international cooperation programs including data from the Agencia Presidencial de Cooperación Internacional de Colombia, which produced a "Mapa de Cooperación Internacional". We also identified as much information as possible from the National Administrative Department of Statistics (DANE) website, which posts a sizeable quantity of data.

Table A4.1: Variables employed in the matching

General Outcome	Variables (Concepts)	Indicators / Scale	Source of indicator	Unit of Observ..
Background	Department	Dep't Tag	Colombia Gov't	
	Altitude		Colombia Gov't	Municipality
	Ruralness	Percent rural (0 .002 – 1)	Colombia Gov't	Municipality
	Government consolidation muni	Yes (1) No (0)	USAID	Municipality
Violence	BACRIM Presence	Yes (1) No (0)	Instituto de Estudios para el Desarrollo y la Paz	Municipality
	Fatalities from Attacks	0 – 105	CDMH	Incident
	Paramilitary Assassinations	0 – 335	CDMH	Incident
	Homicide Rate	Number per 100,000 (0 - 508)	CEDE	Municipality
	Guerrilla Attacks	0 – 291	Vargas	Incident
	Clashes	0 – 130	Vargas	Incident
Demographic	Minority Population	Percent of the population that is minority (0-1)	Census	Municipality
	Distance from Dep't Capital	Distance in km (0 – 493)	CEDE	Municipality
	Infant mortality rate	Average per 1,000 live births (7.6 – 91.97)	Colombia Ministry of Health	Municipality
	Education facilities	1 – 2,721	CEDE	Municipality
	Debt magnitude	0 – 99.4	CEDE	Municipality
	Distance from nearest market	Distance in km (0 – 926)	CEDE	Municipality
Economic	Remittances	Total value since 2012 (5.6 – 15.9 billion)	SGR	Municipality
	Dead from mine incidents	Total since 1990 (0 – 68)	Colombia Gov't	Municipality
	Coca Production	Hectares (0 – 5,464)	UNODC	Municipality
Political	Corruption risk	Ordinal risk score (0 – 3)	Fundación Paz y Reconciliación	Municipality
	Fiscal Performance	Effective operating budget (314 – 1,966,462)	Colombia Gov't	Municipality
International Involvement	USAID Activity	Yes (1) No (0)	USAID	Municipality
	World Bank Program Activity	Yes (1) No (0)	World Bank	Municipality

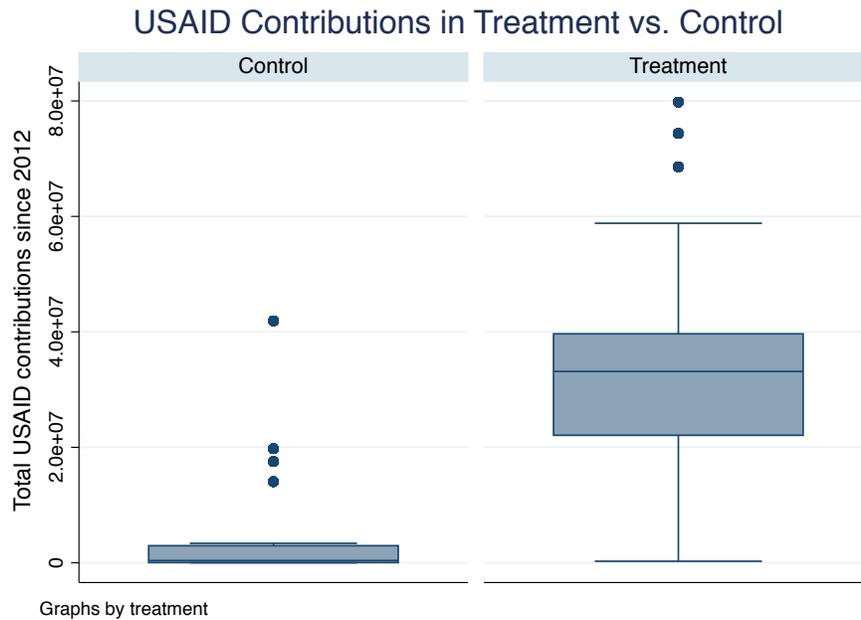
Including all of this information, in brief, we developed a single score for the likelihood that a unit would receive treatment. That score “in expectation” balanced all possible confounding factors that we could observe. Once covariate-balanced propensity scores were assigned to all possible units, we compared pairs of treated and untreated units with similar propensity scores. These comparisons approximated those that would occur in a randomized experiment to the extent that relevant factors were observable.

A naïve strategy would have been to identify every possible municipality in the country and then include all possible information about those municipalities in a matching process that cuts down the number of cases. Such an approach would naturally include control units in conflict-affected areas that we would have cared about comparing. But the matching process also included some municipalities outside the scope and interest of USAID. We thus began with the baseline set of municipalities (greater than the 80 used in the selection process between USAID and GOC) in the departments of interest to USAID. Using that set of possible municipalities, we gathered as much data as possible and estimated propensity scores and created the final set of matches.

We used structured matching to provide a final set that included the 40 RGA-treated municipalities and then another set that experienced neither USAID nor GOC interventions. Such an approach allowed us to make comparisons among municipalities with and without USAID involvement, which was of primary interest.

The statistical matches were largely comparable, but a few exceptions existed. Most critically, the data on historical USAID development assistance across municipalities pointed to a tendency towards continuity of programming. This posed challenges from an evaluation perspective because the data showed that RGA treatment municipalities had historically received much higher levels of USAID programming and financial support than the control municipalities. This previous “priming” of treatment municipalities means that any RGA program effects need to be qualified by acknowledging that the effects were due to the RGA *plus prior programming*. See Figure A4.1, which demonstrates this using data from the USAID MONITOR system. We tried numerous possible ways to balance treatment and control municipalities on this dimension but there were not sufficient possible control municipalities with substantial USAID operations. Thus, the treatment and control groups are mostly imbalanced with respect to this covariate.

Figure A4.1: Prior USAID Activities in Treatment/Control

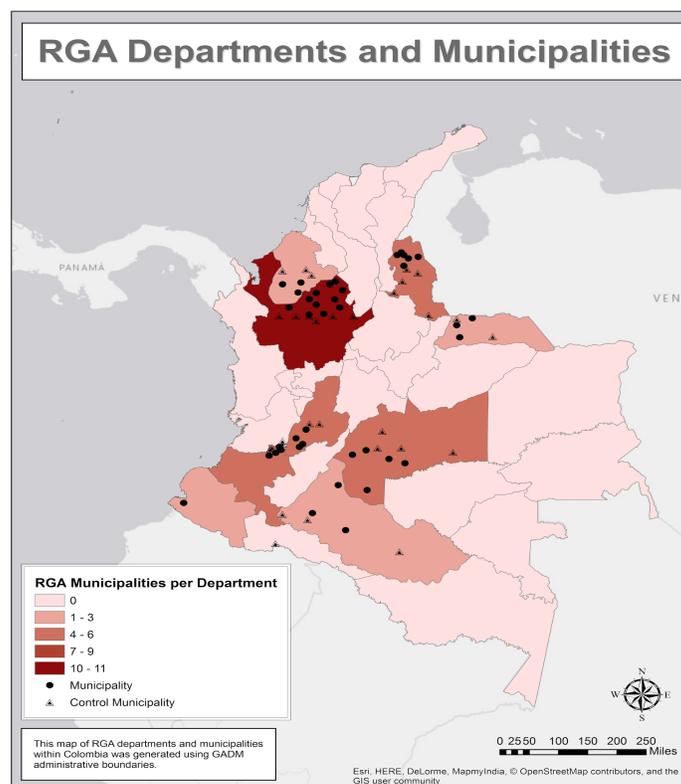


One implication of USAID's past heavy involvement in treatment municipalities is that citizens should be more likely to report that they are aware of current RGA activities. When asked knowledge/awareness questions, citizens often lie and we expected some level of citizens saying they were aware of the RGA. But if past activities shape citizens beliefs, then we would have expected citizens in treatment municipalities to report knowing about the RGA. Curiously, throughout the entire activity, despite past and current USAID work, very few citizens reported knowing much about RGA.

The final set of treatment and control municipalities are illustrated below in Figure A4.2 (See SOW or baseline reports for lists of the treatment and control municipalities.)⁷

⁷ USAID has worked extensively in many of the municipalities previously. Thus, the RGA intervention is not a clean, first-time intervention. While we used matching to identify control municipalities that were similar, USAID has had an overall stronger historical presence in the treatment areas, thus making for a less than optimal balance across these municipalities. We specified USAID history in the matching model different ways and no matter the approach, the matching results were similar.

Figure A4.2: Treatment and Control Municipalities



See attached document (“A.4.Sampling-Procedures-English.pdf”) for more details about the survey sampling procedures. Also see Annex 6 below for more information on sources of information.

Annex 4b: Analysis

To interpret the results, we plotted the quantitative data in figures and provide visual interpretation. In most cases, this involves comparing changes in control from baseline to midline to endline, and then changes in treatment from baseline to midline to endline, and then comparing across control and treatment. To contextualize these comparisons, we also estimated statistical regression models comparing

outcome indicators in treatment municipalities to outcomes in control municipalities at the endline (which essentially produces a difference-in-means estimate), and then control for the midline and baseline values (Glennester and Takavarasha 2013), clustering survey respondents, where appropriate, on municipalities. Throughout, we refer to these results as “statistically meaningful” (or not) based on whether the regression analyses suggest that the results could not have occurred by random chance. Given the small number of municipalities, these regressions provide a difficult test, and so the visual interpretation is all the more important as well as supplementing with the qualitative data. (The full set of regression results is reported in Annex 2d, Table A2.9.) Given that not many of the relationships are “statistically meaningful” we primarily draw attention to the analyses in the cases where they are.

We note two challenges that shape our analysis, discussed further in the baseline and midline reports. First, due primarily to security concerns, the implementing partner was not able to begin operations in all 40 treatment municipalities at the same time. This was especially true at baseline, but persisted through midline to some extent. As such, we attempted to quantify the varying implementation efforts across municipalities by considering variation in the data on MSI’s activities as well as through USAID’s MONITOR tracking data. All regression results are weighted by RGA effort according to these measures. Second, the security conditions prevented us from conducting all evaluation activities in all treatment and control municipalities. Because of worsening security conditions, especially at baseline and midline, but to some extent at endline, data coverage was mostly comprehensive but with some gaps that we have discussed at least qualitatively along with some targeted analysis for Arauca.

Annex 5: Instruments

See attached document (“A.5.Instruments.zip”) for the survey, interview, and focus group instruments.

Annex 6: Sources of Information

We successfully collected data in most municipalities. Security concerns prevented us from collecting baseline data in Arauca, though we were able to collect midline and endline data for Arauca with a shortened survey instrument, and have included a separate subsection in the main report, and more information below dedicated to the Arauca findings. At endline, we did not have any major security incidents, though we did find that many people were concerned about reporting and did not want to respond to some questions. Thus, there was a general atmosphere of insecurity that likely shaped reporting on many survey questions of interest.

Most of the data collection occurred at the municipality level and entailed conducting household perception surveys, interviews with elites, focus groups, document review, and collecting other observational data. The survey, interview, and focus group questionnaires are located in Annex 5. The survey firm, Cifras y Conceptos, carried out the data collection. Ahead of the endline collection, we revised the instruments and worked with Cifras y Conceptos to develop all training procedures. Before baseline, midline, and endline, we conducted pilot surveys based to refine questions for the final data collection. The evaluation team also made site visits to several municipalities during each of the baseline, midline, and endline waves. For the endline, site visits occurred in Tolima Department (RGA municipality of Chaparral), where the evaluation team observed focus group and survey data collection.

The within-municipality sampling procedures are detailed in Annex “A.4.Sampling-Procedures-English.pdf”. We originally planned to conduct 7,000 surveys at the baseline, midline, and endline surveys for a total of 21,000 surveys. In the baseline, security challenges prevented us from operating in all municipalities (see

discussion in baseline report), so the final count of surveys ended up being 6,389. In the midline, we attempted to compensate for the previous shortage of baseline surveys. As such, in the midline, we surveyed additional respondents in Arauca, and in total surveyed 7,007 respondents for the midline. For the endline, we completed surveys with 6,692 for an overall total of 20,088 with a final set of 20,002 useable observations.⁸

Surveys were conducted in the municipality centers (cabeceras) as well as other smaller population centers (centros poblados) around the municipality, which capture a rural component, though the surveys are not fully representative of the rural areas as we were constrained in the number of rural areas we could capture.⁹ For the endline, we sought to generate a panel of repeat respondents through several sampling approaches. We first attempted to survey the same individuals we had previously surveyed in both the baseline and midline. If that was not possible, we attempted to resurvey individuals who previously participated in either the baseline or the midline surveys, but not both. If that was not possible, then we checked whether the individual previously surveyed had moved to another municipality or department and tried to survey them there if possible. If that was not possible, we asked to survey someone else in the same household as in one of the previous surveys. Failing that, we resampled new respondents from the same municipalities but in different houses. The full distribution of sampling and resampling, along with percentages in each category, is reported in Table A6.1.

Table A6.1: Distribution of endline respondents

Panel 1	Panel 2	Panel 3	Panel 4	Panel 5	Panel 6	Total
1,914	1,320	754	582	2,117	5	6,692

1. Same person in both baseline and midline measurements
2. Same person who responded in one of the measurements
3. Same household, different person
4. Same house structure, new family/person
5. Same street/block within a municipality
6. Same person who responded in baseline or midline, but not in different municipality

As with the baseline and midline, in the endline we also collected data from other sources.¹⁰ In addition to the citizen survey, we also attempted to conduct interviews with three government officials in each

⁸ Although there were minor security incidents at baseline and midline (see those reports), the endline data collection did not face any significant security incidents. However, the security environment had deteriorated across the country. Respondents were fearful and there was a widespread refusal to answer questions that were even borderline sensitive. Arauca was a consistently challenging department in which to access and collect information, and we adapted the survey instrument accordingly since respondents would not answer a variety of sensitive questions.

⁹ The issue of urban/rural representativeness within municipalities was discussed with USAID ahead of the baseline, and it was clear that USAID could not support a larger evaluation that would be representative of the rural areas within municipalities. As such, USAID encouraged the evaluation team to sample from some centros poblados, though not in a representative way.

¹⁰ Measuring perceptions is a standard technique, and improving perceptions is a central desired outcome of USAID's programming, and the theory of change articulates its importance. Based on the initial tasking: "The purpose of the RGA is to improve sub-national governance in 40 conflict-affected municipalities of Colombia. Municipal capacity to effectively deliver services to citizens plays a key role in building the legitimacy of the state at the local level. Given that the lack of state legitimacy at this level contributes to conditions that foster or

mayoral office, six elites within the community (e.g., within social accountability organizations), and focus groups. We were not always able to achieve that coverage given that some leaders would not agree to meet or were otherwise difficult to contact. In total, we were able to conduct 290 elite interviews. The mission and implementing partners determined not to target specific areas, but rather pursue broad community effects. As such, we could not identify specific neighborhoods, centros poblados, social and government leaders, or otherwise, ahead of time to collect baseline data. We were told that those decisions would not be made until much later, so we were left doing random and/or arbitrary selection of local leaders and centros poblados.

We also conducted focus groups to better understand citizen perceptions, including their views on municipal spending and oversight, and general governance and security conditions in the municipalities. The focus group questions were adapted from the citizen questionnaire. Through the survey firm Cifras y Conceptos, we conducted 20 focus groups for each of the baseline and midline reports, and repeated those focus groups for the endline, though not with the exact same set of individuals. As such, the focus groups do not constitute their own panel, but rather a new sample albeit within the same set of municipalities at baseline and midline.

Finally, we also collected administrative information from publicly available sources and included those indicators as appropriate to compare to the baseline and midline conditions. Specifically, we collected administrative data on indicators of interest from national government agencies and directly from municipalities, including site visits and on-the-ground observations. Observations for some indicators vary in coverage from about 40 to 70 municipalities since we were not able to acquire data for all the municipalities in the sample.

In sum, we collected an enormous amount of data across a variety of sources and over time. In the findings sections, we analyze the data to provide supporting evidence and illustrative material. All data have been anonymized for security reasons.

See original statement of work, and earlier waves of the evaluation, for longer discussions of sources of information.

Annex 7: Evaluation Statement of Work

See attached document (“A.7.RGA_IE-SOW_27 March 2015 signed.pdf”) for initial statement of work.

Annex 8: Summary Information about Evaluation Team Members

Michael G. Findley is Professor of Government and Public Affairs at the University of Texas at Austin. Findley conducts research on political violence, international development, and international law, with extensive fieldwork in Colombia, Kenya, DRC, Sudan, South Sudan, South Africa, and Uganda. His publications appear in leading scholarly and policy journals and presses. He works extensively in

perpetuate conflict, increased legitimacy will contribute to minimizing conflict.” Note that legitimacy is a perception-based characteristic. And the tasking also asked for due attention to issues of trust. And so on. There are of course potential biases with perceptions data, which is why we included other forms of data. That was the aim of introducing administrative data, interview data, focus groups, etc. Relevant administrative indicators of revenues, spending, and service provision should help to account for inefficiencies in municipal performance, as should the DNP indices in the analysis.

collaboration with various policy organizations including the World Bank, USAID, African Development Bank, UNICEF, UN Peacebuilding Fund, UN Development Program, International Aid Transparency Initiative, and many developing country governments to promote evidence-based policymaking. Such work has resulted in numerous impact evaluations (and reports), policy publications, official policy working papers, and policy reviews.

Oliver Kaplan is Associate Professor at the Josef Korbel School of International Studies at the University of Denver. He is the author of the book, “Resisting War: How Communities Protect Themselves” (Cambridge University Press, 2017), which examines how civilian communities organize to protect themselves from wartime violence. He was a Jennings Randolph Senior Fellow at the U.S. Institute of Peace and previously a postdoctoral Research Associate at Princeton University in the School of Public and International Affairs and at Stanford University. His research has been published in *The Journal of Conflict Resolution*, *Journal of Peace Research*, *Conflict Management and Peace Science*, *Stability*, *The New York Times*, *Foreign Affairs*, *Foreign Policy*, *CNN*, and *National Interest*. Kaplan received his Ph.D. in political science from Stanford University and completed his B.A. at UC San Diego.

Ana Marrugo is a graduate student of anthropology at the University of Pittsburgh. Her research focuses on the workings of transitional law in Colombia and its implications for movements for social justice in the country. She worked at the Victims’ Unit in Colombia strengthening victim’s organizations abroad and in the creation of strategic alliances with international partners to assist victims living outside of Colombia. She has also worked as a consultant for projects of national and international agencies in political participation and development in Colombia.

Alejandro Ponce de Leon is Ph.D. candidate at the University of California, Davis. Ponce de Leon conducts research on culture, political violence and agricultural policies in contemporary Colombia.

Danny Walker is a data specialist who has worked with USAID, UNICEF, the IMF, and the World Bank on a wide variety of quantitative initiatives. These have ranged from ethnic fractionalization in Indonesian, to civil service wage bill in Brazil, to crowd-sourced SMS messages in Uganda, to an index of municipal characteristics in Colombia for this present project. In each case, Danny’s career interest is in how data and analytical insights can aid in international development, reducing civil unrest and violence, and alleviating poverty. He currently resides in West Hartford, Connecticut.

Joseph K. Young is Professor of Public Affairs and International Service at American University in Washington, DC. Young conducts research on political violence, governance, and public policy with field work in Latin America and the Middle East. His publications have appeared in top journals in political science, economics, international studies, criminology, and Latin American Studies. He has been invited to speak to organizations in the defense community and has consulted on a Department of Defense initiative focusing on countering violent extremism. The National Science Foundation and the National Consortium for the Study of Terrorism and Responses to Terrorism (START) have funded his research.

Annex 9: Statement of Differences (when applicable)

Not applicable