Online Appendix

Local Elites As State Capacity: How City Chiefs Use Local Information To Increase Tax Compliance In The D.R. Congo

By Pablo Balán, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel

A1 Additional Exhibits for the Main Analysis **45**
A1.1 Additional Exhibits for Paper Section 2 — Setting . . . 45
A1.2 Additional Exhibits for Paper Section 3 — Design . . . 46
A1.3 Additional Exhibits for Paper Section 5 — Estimation . . . 52
A1.4 Additional Exhibits for Paper Section 6 — Main Results . . . 53
A1.5 Additional Exhibits for Paper Section 7 — Mechanisms . . . 65
A1.6 Additional Exhibits for Paper Section 8 — Distributional Impacts . . . 80
A1.7 Additional Exhibits for Paper Section 9 — Policy Implications . . . 90

A2 Additional Details on the Tax Campaign and its Evaluation **94**
A2.1 Block-Randomized Design . . . 94
A2.2 Tax Letter Message Treatments . . . 94
A2.3 Chief Jurisdiction Mapping . . . 95
A2.4 Logistics Pilot . . . 96
A2.5 Time Imbalance . . . 96
A2.5.1 Preferred Specification and Robustness Tests . . . 97
A2.6 Detailed Survey-based Variable Descriptions . . . 98

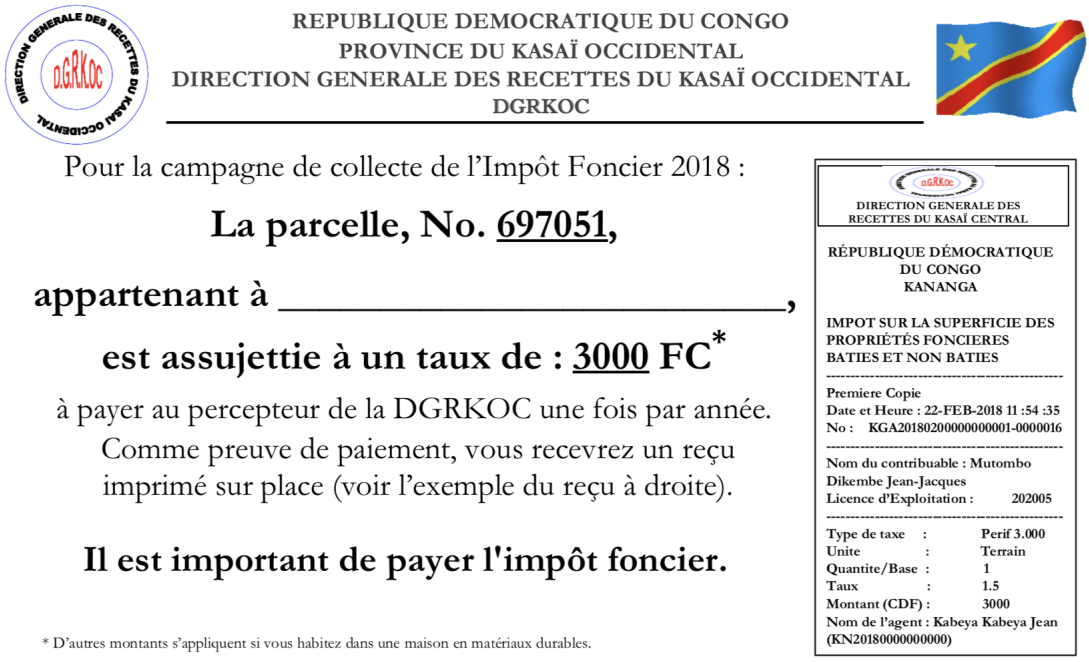
A3 Further Analysis **105**
A3.1 Conceptual Framework . . . 105
A3.1.1 Setup . . . 105
A3.1.2 Collector objective functions . . . 107
A3.1.3 The Government's Decision . . . 108
A3.1.4 Discussion . . . 109
A3.1.5 External Validity . . . 111
A3.2 Combined Team — Central X Local . . . 114
A3.3 State Collector Team Composition and Performance . . . 116
A3.4 Collector Exhaustion and Demoralization . . . 117
A3.5 Quantifying the Knowledge Gap between Chiefs and State Agents . . . 122
A3.6 The Limits to Codifying Local Information . . . 123
A3.7 Cost-Effectiveness . . . 125

A4 Ethical Considerations **129**

A1 Additional Exhibits for the Main Analysis

A1.1 Additional Exhibits for Paper Section 2 — Setting

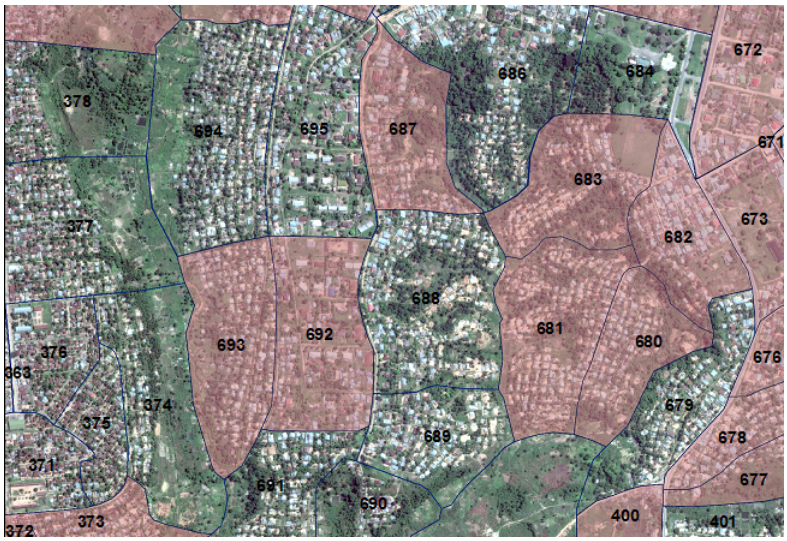
FIGURE A1: SAMPLE TAX NOTICE



Notes: This figure displays a sample tax notice, discussed in Section IA. The flier reads: "For the 2018 property tax collection campaign: the property 697051 belonging to [name of owner] is subject to a tax rate of 3000 CF to be paid to a DGRKOC collector once per year. As proof of payment, you will receive a receipt printed on the spot (see example to the right). It is important to pay the property tax." The footnote reads "Other amounts apply if you live in a house built of durable materials." This flier contains the Control message ("It is important to pay the property tax"), discussed in the text in Section VC and in detail in Section A2.2. A version of the flier in Tshiluba, the primary local language, was printed on the opposite side. Fliers were identical across treatment arms.

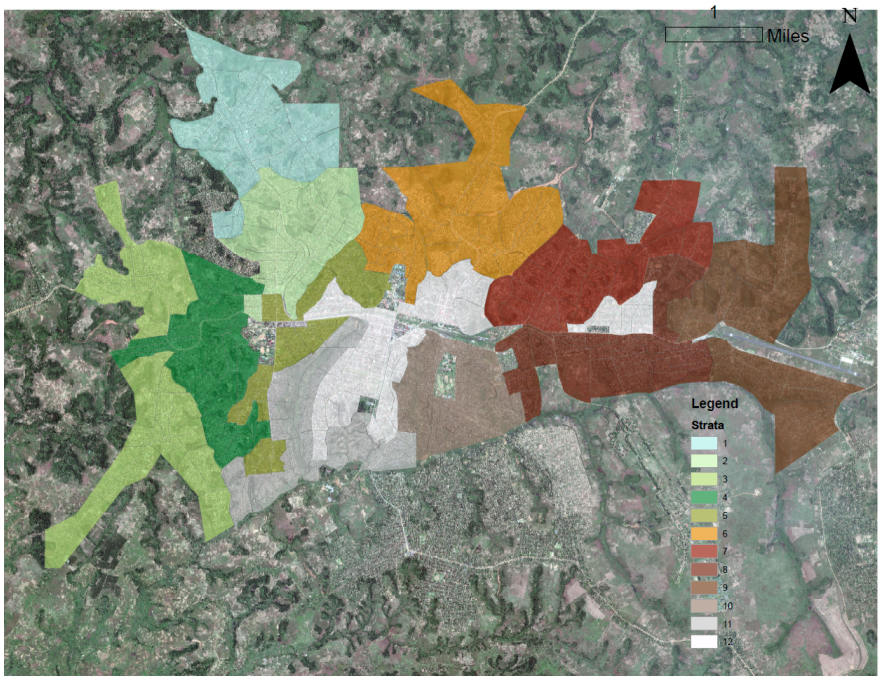
A1.2 Additional Exhibits for Paper Section 3 — Design

FIGURE A2: THE UNIT OF RANDOMIZATION: NEIGHBORHOODS OF KANANGA



Notes: This figure displays a sample of neighborhood divisions in Kananga, which are discussed in Section A2.1.

FIGURE A3: GEOGRAPHIC STRATA



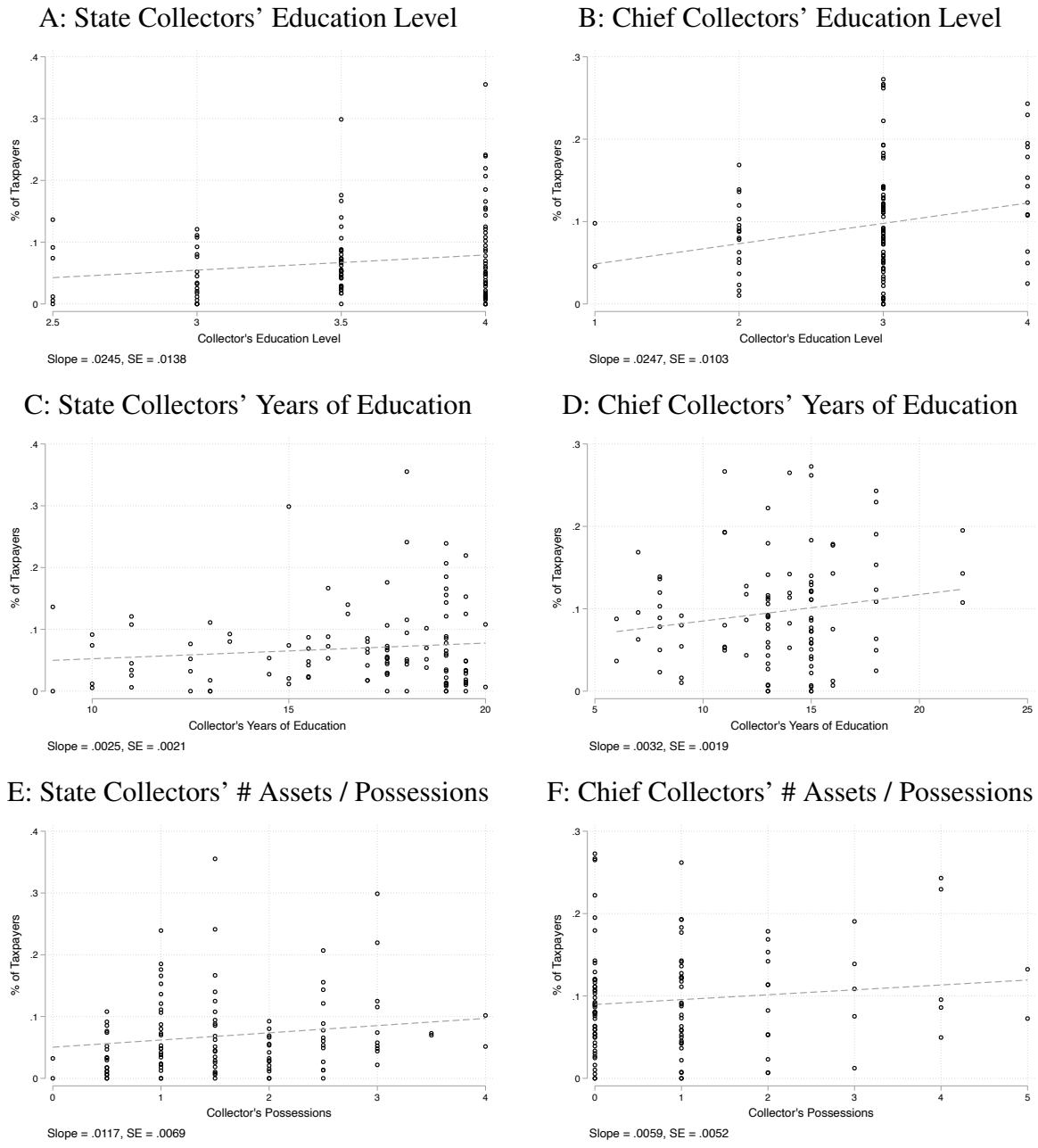
Notes: This figure displays the geographic strata of Kananga, which are discussed in Section A2.1.

TABLE A1: LOCAL V. CENTRAL: COLLECTOR CHARACTERISTICS

Variable	State Collectors	Chief Collectors	Difference
	(1)	(2)	(3)
Age	30.760	58.712	27.952***
	(8.098)	(11.031)	(1.740)
% Female	0.060	0.045	-0.015
	(0.240)	(0.208)	(0.037)
Born in Kananga	0.480	0.607	0.127
	(0.505)	(0.491)	(0.085)
Log Monthly Income	4.238	4.045	-0.192
	(0.969)	(1.153)	(0.189)
Number of Possessions	1.820	1.044	-0.776***
	(1.320)	(1.263)	(0.218)
Years of Education	16.940	13.266	-3.674***
	(3.413)	(3.487)	(0.592)
Works Other Job	0.682	0.761	0.079
	(0.471)	(0.428)	(0.078)
Test Maths (Mean)	0.745	0.743	-0.002
	(0.234)	(0.258)	(0.043)
Reading Ability (Mean)	1.770	1.838	0.068
	(0.612)	(0.779)	(0.124)
Trust in Government (Mean)	3.033	2.716	-0.317*
	(0.732)	(1.051)	(0.165)
Government Capacity (Mean)	150.178	158.660	8.482
	(73.893)	(99.387)	(15.690)
Poor Priority (Mean)	2.680	2.758	0.078
	(0.563)	(0.588)	(0.099)
Progressivity (Mean)	2.584	2.470	-0.114**
	(0.285)	(0.308)	(0.051)
Observations	50	113	163

Notes: This table compares baseline characteristics of state collectors in neighborhoods assigned to the Central treatment arm (Column 1) and chiefs in neighborhoods assigned to the Local treatment arm (Column 2). Column 3 reports a simple difference-in-means test. The data come from surveys conducted with tax collectors before the 2018 campaign. The first seven variables are the respondent's age, a sex indicator, an indicator for being born in Kananga, log monthly income, wealth (defined as the number of possessions: motorbike, car, radio, TV, generator and sewing machine), years of education, and an indicator for working another job during the tax campaign. *Math Ability* and *Reading Ability* are collectors' average score on a series of quiz-type questions. The last four measures concern attitudes about the government and redistribution, measured through survey questions with Likert-scale response options. These comparisons are discussed in Section IIA.

FIGURE A4: COLLECTOR PERFORMANCE AND EDUCATION / WEALTH



Notes: This figure shows the relationship between tax compliance in the neighborhood and tax collectors' education levels (Panels A and B), years of education (Panels C and D), and wealth (Panels E and F). Wealth is defined as number of possessions among the following: motorbike, car, radio, TV, generator, and sewing machine. The correlations are reported separately for neighborhoods assigned to the Central and CLI treatment arms where tax collection was conducted by state agents (Panels A, C, and E) and for neighborhoods assigned to the Local treatment arm where tax collection was conducted by city chiefs (Panel B, D, and F). These comparisons are discussed in Section IIA

TABLE A2: RANDOMIZATION BALANCE: BILATERAL TREATMENT COMPARISONS

	Local	CLI	CXL
	(1)	(2)	(3)
Panel A: Baseline Characteristics			
Years of Education	-0.003	0.001	-0.003
	(0.003)	(0.003)	(0.003)
Electricity	0.008	0.021	0.030
	(0.027)	(0.031)	(0.031)
Log HH Monthly Income	0.006	-0.003	-0.006
	(0.005)	(0.005)	(0.005)
Trust of Chiefs	0.012	0.026**	0.026**
	(0.012)	(0.012)	(0.013)
Trust of National Government	-0.015	-0.010	-0.002
	(0.014)	(0.012)	(0.013)
Trust of Provincial Government	0.026	0.018	-0.001
	(0.016)	(0.015)	(0.016)
Trust of Tax Ministry	-0.001	-0.008	-0.003
	(0.010)	(0.012)	(0.012)
Observations	2117	1768	1501
Clusters	221	187	159
F, p	1.08, 0.37	1.12, 0.34	1.15, 0.33
Panel B: Midline Characteristics			
Sex	-0.001	-0.027**	-0.005
	(0.009)	(0.011)	(0.010)
Age	0.0004	-0.0003	-0.0003
	(0.0004)	(0.0004)	(0.0004)
Majority Tribe	0.001	-0.013	0.001
	(0.018)	(0.014)	(0.011)
Employed	-0.002	0.008	0.004
	(0.013)	(0.014)	(0.010)
Salaried	0.003	-0.032**	-0.025*
	(0.016)	(0.014)	(0.015)
Works for Government	-0.029	0.029	-0.019
	(0.025)	(0.024)	(0.024)
Relative Works for Government	0.036	0.024	0.043*
	(0.025)	(0.023)	(0.024)
House Quality	-0.001	0.002	0.005
	(0.022)	(0.018)	(0.019)
Distance to State Buildings and City Center	0.061	-0.470**	0.048
	(0.158)	(0.156)	(0.199)
Distance to Health Institutions	0.064	0.257	-0.066
	(0.201)	(0.222)	(0.187)
Distance to Education Institutions	0.445*	0.387	0.179
	(0.267)	(0.250)	(0.310)
Distance to Roads	-0.171	0.035	0.197
	(0.145)	(0.133)	(0.133)
Distance to Eroded Areas	0.157	0.026	0.458
	(0.262)	(0.297)	(0.303)
Observations	10666	8500	7542
Clusters	172	141	123
F, p	0.98, 0.47	2.37, 0.01	1.00, 0.46
Panel C: Neighborhood Characteristics			
Per Capita Property Tax Revenues in 2016	0.0001	-0.0001	-0.0001
	(0.0002)	(0.0001)	(0.0002)
Affected by Conflict in 2017	0.132	-0.131	0.444**
	(0.289)	(0.362)	(0.215)
Observations	221	190	160
Clusters	221	190	160
F, p	0.39, 0.68	0.41, 0.67	2.46, 0.09
Stratum FE	Yes	Yes	Yes

Notes: This table summarizes balance tests for bilateral treatment comparisons. Each column compares the noted treatment arm to Central. The bottom row of each panel contains the statistics for tests of the omnibus null hypothesis that the treatment effects for the covariates studied in Table 3 are all zero using parametric F tests. As usual, regressions include stratum fixed effects and cluster standard errors at the neighborhood level. We run separate tests for variables drawn from baseline survey, midline survey, and neighborhood-level data to maximize the number of observations included in each regression. Midline characteristics include the distance characteristics from registration reported in Table 3. We discuss these results in Section IIC.

TABLE A3: RANDOMIZATION BALANCE: INCLUDING CONTROL GROUP

	N	Control Mean	Central	Local	CLI	CXL
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Property Owner Characteristics						
Years of Education^B	3667	9.75	0.81	0.71	0.40	0.41
			(1.50)	(1.50)	(1.51)	(1.52)
Electricity^B	3680	0.19	-0.06	-0.05	-0.07	-0.04
			(0.09)	(0.09)	(0.09)	(0.09)
Log HH Monthly Income^B	3646	10.64	-0.11	0.07	-0.15	-0.25
			(0.29)	(0.29)	(0.30)	(0.34)
Trust of Chief^B	3666	2.91	0.16	0.23	0.31	0.36
			(0.35)	(0.35)	(0.35)	(0.35)
Trust of National Government^B	3488	2.33	0.19	0.22	0.18	0.20
			(0.18)	(0.18)	(0.18)	(0.19)
Trust Provincial Government^B	3511	2.25	0.16	0.25	0.19	0.18
			(0.20)	(0.20)	(0.21)	(0.21)
Trust of Tax Ministry^B	3474	2.37	-0.01	0.02	-0.04	-0.08
			(0.15)	(0.15)	(0.16)	(0.16)
Sex^M	22699	0.84	-0.07***	-0.05**	-0.07**	-0.08***
			(0.02)	(0.02)	(0.03)	(0.03)
Age^M	20269	53.85	0.50	0.24	0.17	0.64
			(1.16)	(1.14)	(1.27)	(1.30)
Majority Tribe^M	23014	0.81	-0.04	-0.02	-0.01	-0.02
			(0.08)	(0.08)	(0.08)	(0.08)
Employed^M	24764	0.78	-0.04	-0.03	-0.03	-0.05
			(0.03)	(0.03)	(0.03)	(0.03)
Salaried^M	24765	0.23	0.02	0.01	0.01	0.01
			(0.05)	(0.05)	(0.05)	(0.05)
Works for Government^M	24765	0.16	-0.01	-0.01	-0.003	-0.01
			(0.05)	(0.05)	(0.05)	(0.05)
Relative Works for Government^M	27497	0.26	-0.03	-0.04	-0.02	-0.03
			(0.05)	(0.05)	(0.05)	(0.05)
Panel B: Property Characteristics						
House Quality^M	28957	-0.14	0.14	0.05	0.28	0.14
			(0.31)	(0.31)	(0.31)	(0.32)
Distance to State Buildings and City Center^R	44899	1.86	-0.36**	-0.31*	-0.28	-0.36*
			(0.18)	(0.18)	(0.19)	(0.20)
Distance to Health Institutions^R	44899	0.38	-0.05	-0.02	0.01	-0.06
			(0.08)	(0.08)	(0.08)	(0.08)
Distance to Education Institutions^R	44899	0.78	-0.13	-0.10	-0.04	-0.13
			(0.21)	(0.21)	(0.21)	(0.21)
Distance to Roads^R	44280	0.38	0.04	0.08	0.03	0.04
			(0.14)	(0.14)	(0.14)	(0.15)
Distance to Eroded Areas^R	44280	0.12	-0.003	0.01	0.002	0.02
			(0.02)	(0.02)	(0.02)	(0.02)
Panel C: Neighborhood Characteristics						
Per Capita Property Tax Revenues in 2016^B	356	176.48	-31.11	-22.08	-78.00	-47.53
			(162.29)	(162.58)	(160.37)	(161.13)
Affected by Conflict in 2017^B	356	0.20	-0.18	-0.16	-0.16	-0.14
			(0.18)	(0.18)	(0.18)	(0.18)
Panel D: Attrition:						
Baseline to Endline	4246	0.13	-0.02	-0.04	-0.05	-0.06
			(0.05)	(0.05)	(0.05)	(0.05)
Baseline Replacement	3483	0.17	-0.015	-0.004	-0.007	0.002
			(0.04)	(0.04)	(0.04)	(0.05)
Registration to Midline	45162	0.26	-0.05	-0.03	-0.06	-0.10*
			(0.06)	(0.06)	(0.06)	(0.06)

Notes: This table reports the coefficients from balance tests estimated by regressing characteristics for property owners (Panel A), properties (Panel B), and neighborhoods (Panel C) on treatment indicators, clustering standard errors at the neighborhood level. Panel D shows differences in attrition from baseline to endline surveying, replacement at endline of baseline respondents, and attrition from registration to midline surveying. The Control arm is the excluded category. Randomization stratum fixed effects are not included because Control neighborhoods do not exist in every strata. Superscripts B, M, and R denote variables from baseline, midline, and registration, respectively. Variables are described in Section A2.6. Joint orthogonality tests for specific treatment comparisons are shown in Table A2. We discuss these results in Section IIC.

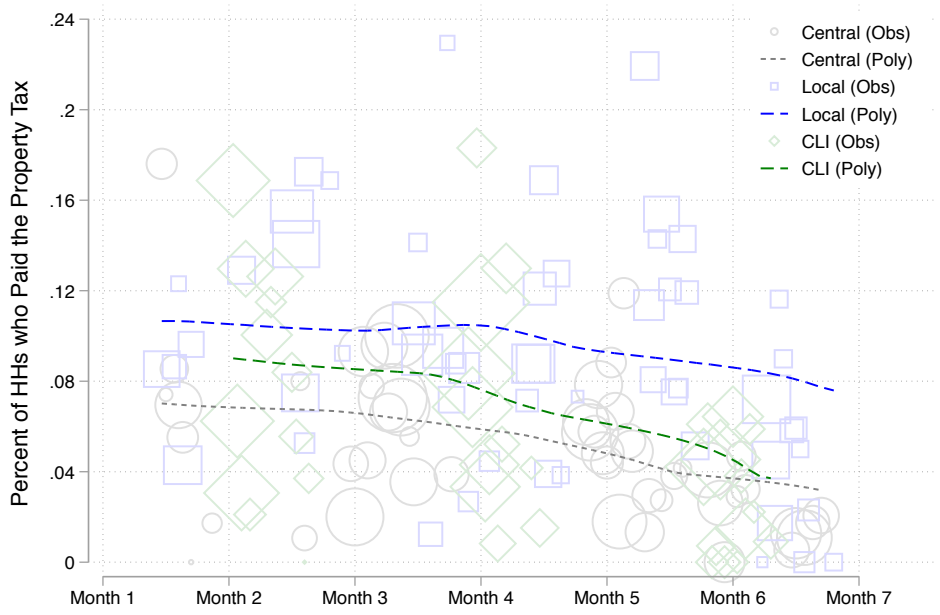
TABLE A4: MIDLINE NON-RESPONSE ACROSS TREATMENTS

	Local	CLI	CXL
	(1)	(2)	(3)
Sex Missing	0.181	-0.081**	0.388*
	(0.317)	(0.035)	(0.214)
Age Missing	-0.304	-0.090**	-0.460**
	(0.319)	(0.038)	(0.214)
Majority Tribe Missing	0.026	0.035	0.025
	(0.024)	(0.025)	(0.023)
Employed Missing	-0.348	0.097*	0.044
	(0.220)	(0.057)	(0.040)
Salaried Missing	0.368*	-0.060	-0.031
	(0.217)	(0.051)	(0.032)
Relative Works for Government Missing	-0.002	0.002	0.002
	(0.032)	(0.033)	(0.030)
Observations	22533	18927	16494
Clusters	221	189	160
F, p	1.58, 0.15	1.54, 0.17	0.95, 0.46
Stratum FE	Yes	Yes	Yes

Notes: This table summarizes tests for differential midline non-response. Each column compares the noted treatment arm to Central. The bottom row of each panel contains the statistics for tests of the omnibus null hypothesis that the treatment effects for all the variables listed are zero using parametric F-tests. Regressions include stratum fixed effects and cluster standard errors at the neighborhood level. The *Works for Government* variable is omitted as it is constructed using the same underlying variable as *Salaried* and thus collinear.

A1.3 Additional Exhibits for Paper Section 5 — Estimation

FIGURE A5: DECREASING COMPLIANCE OVER TIME — CENTRAL, LOCAL, CLI



Notes: This figure shows the decrease in compliance for Central, Local, and CLI over the 2018 tax campaign. Blue squares represent Local observations, gray circles represent Central observations, and green diamonds represent CLI observations, with size proportional to the number of observations. Lines — dashed blue for Local, dotted gray for Central, and dashed green for CLI — are local linear polynomials estimated separately by treatment. This figure is discussed in Section IV.

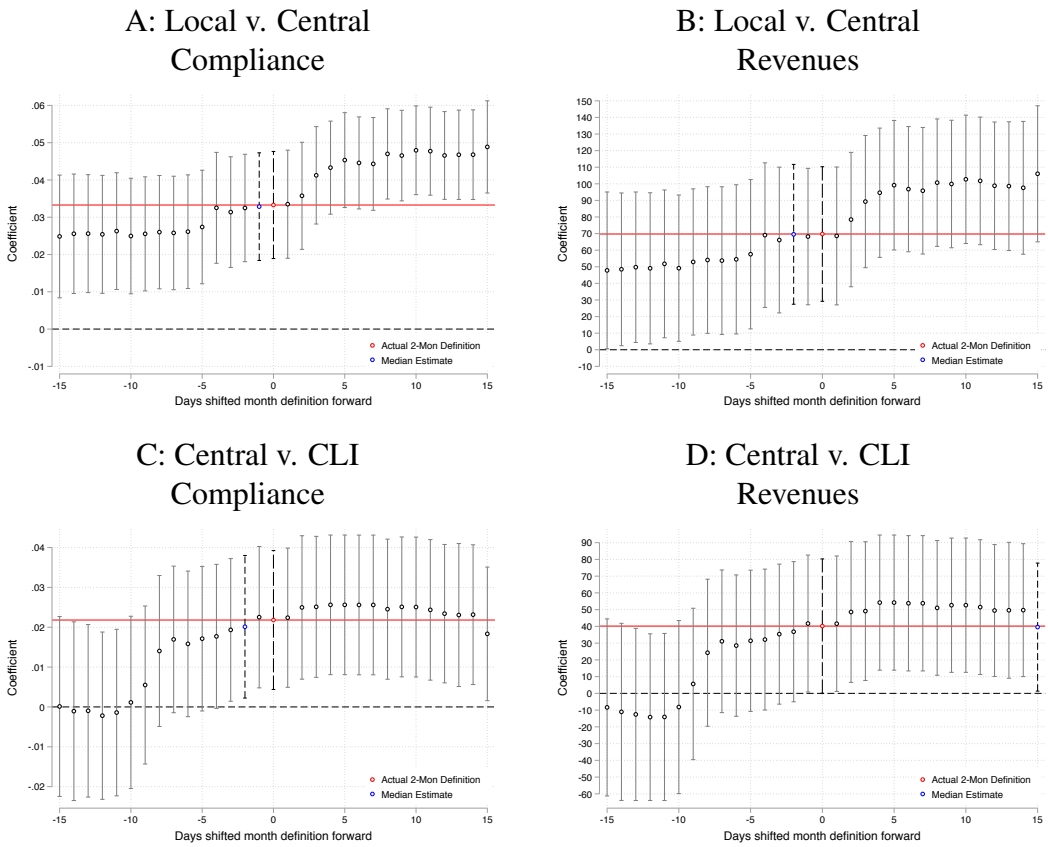
A1.4 Additional Exhibits for Paper Section 6 — Main Results

TABLE A5: LOCAL V. CENTRAL ROBUSTNESS: DIFFERENT APPROACHES TO TIME IMBALANCE

	No Adjustment (1)	Two Month Fixed Effects (2)	Shift Median Two Month Fixed Effects (3)	Interaction Weighted Estimator (4)	One Month Fixed Effects (5)	Time Restriction (6)	Coarsened Exact Matching (7)
Panel A: Compliance							
Local	0.023**	0.033***	0.033***	0.031***	0.032***	0.042***	0.032***
	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Observations	28872	27764	27506	37186	28872	25912	26637
Clusters	221	213	211	221	221	199	203
Central Mean	.068	.063	.064	.063	.068	.053	.068
Panel B: Revenues							
Local	46.362**	69.744***	69.558**	73.775***	68.695**	91.176***	77.966**
	(23.068)	(20.695)	(21.493)	(18.343)	(21.901)	(20.199)	(30.905)
Observations	28872	27370	27664	36792	28872	25912	26637
Clusters	221	210	212	221	221	199	203
Central Mean	192.891	184.394	185.422	184.394	192.891	158.855	192.891
One Month FE	No	No	No	No	Yes	No	No
Two Month FE	No	Yes	Yes	Yes	No	No	No
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table displays alternative approaches for addressing time imbalance in the comparison of the Local arm to the Central arm, the excluded category, as noted in Section IV and discussed at length in Section A2.5. Panel A reports impacts on compliance, and Panel B reports impacts on revenues. Column 1 makes no adjustments. Column 2 includes the time period fixed effects described in Section IV. Column 3 includes time period fixed effects defined by selecting the median estimate among all permutations of the start date (Figure A6). Column 4 implements an interaction-weighted estimator, following Gibbons et al. (2018), in which time periods defined as in Column 2 are not included as fixed effects but interacted with the treatment indicator and the estimate is the average of the coefficient on the interaction terms, weighted by the number of observations in each period. Column 5 includes one-month fixed effects. Column 6 trims the sample to periods when both treatment arms were in operation. Column 7 implements coarsened exact matching (Iacus et al., 2012). All regressions include fixed effects for house type and randomization strata and cluster standard errors at the neighborhood level. We discuss these results in Section IVA.

FIGURE A6: SHIFTING TWO-MONTH FIXED EFFECT START DATE



Notes: This figure shows robustness to shifting the start date when constructing two-month fixed effects 15 days forward and backwards from the start date in our preferred specification. Panels A and B report estimates for Local compared to Central collection for compliance and revenues, respectively. Panels C and D report estimates for Central + Local Information (CLI) compared to Central. The long-dashed red estimate comes from the preferred definition of time periods; the short-dashed blue estimate is the median among all shifted estimates. All regressions include fixed effects for house type and randomization strata and cluster standard errors at the neighborhood level. We discuss these results in Section IVA and report the median estimate in Table A5.

TABLE A6: LOCAL V. CENTRAL ROBUSTNESS: FULLY-SATURATED MODEL WITH CROSS-RANDOMIZED TREATMENTS

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Compliance						
Local	0.033***	0.033***	0.051***	0.039***	0.036***	0.057***
	(0.007)	(0.007)	(0.011)	(0.008)	(0.010)	(0.014)
Observations	27764	27764	27764	23618	23618	23618
Clusters	213	213	213	213	213	213
Central Mean	.063	.063	.063	.068	.068	.068
Panel B: Revenues						
Local	68.855***	68.923***	75.294**	81.797***	72.674**	76.723**
	(20.560)	(20.562)	(23.838)	(23.626)	(22.119)	(28.482)
Observations	27764	27764	27764	23618	23618	23618
Clusters	213	213	213	213	213	213
Central Mean	182.236	182.236	182.236	196.263	196.263	196.263
Tax Rate FE	No	Yes	Yes	No	No	Yes
Tax Rate FE X Local	No	No	Yes	No	No	Yes
Col. Bonus FE	No	No	No	Yes	Yes	Yes
Col. Bonus FE X Local	No	No	No	No	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates from Equation 1, comparing property tax outcomes in Local and Central (the excluded category). The panels show the estimates from separate regressions with the outcome an indicator for compliance (Panel A) and revenues (Panel B), respectively. All regressions include fixed effects for house, time period, and randomization strata, and they cluster standard errors at the neighborhood level. Column 1 shows the preferred specification, including no additional controls. Column 2 includes dummies for tax rate abatement groups. Column 3 adds interactions between the abatement group dummies and the Local indicator. Column 4 includes dummies for collector bonus type. Column 5 adds interactions between the collector bonus type dummies and the Local indicator. Column 6 includes abatement and collector bonus dummies and interactions with the Local indicator. Bergeron et al. (2020b) provides details on abatement and collector bonus treatment groups. We discuss these results in Section IVA.

TABLE A7: LOCAL V. CENTRAL ROBUSTNESS: INCLUDING CONTROLS, PILOT NEIGHBORHOODS, EXCLUDING MISASSIGNED NEIGHBORHOOD, AND TOP-CODING

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Compliance						
Local	0.032***	0.030***	0.031***	0.031***	0.033***	0.029***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Observations	27751	27751	27751	28784	27658	219
Clusters	213	213	213	219	212	
Central Mean	.063	.063	.063	.064	.063	.061
Panel B: Revenues						
Local	65.349**	60.651**	62.026**	64.469**	68.718***	64.366**
	(20.531)	(21.003)	(20.817)	(19.917)	(20.486)	(19.999)
Observations	27751	27751	27751	28766	27658	219
Clusters	213	213	213	219	212	
Central Mean	182.3	182.3	182.3	181.525	182.3	182.416
Controls:						
Age, Age^2, Sex, Education	Yes	Yes	Yes	No	No	No
Distance to Schools (Imbalanced)	No	Yes	Yes	No	No	No
Employed, Salaried	No	No	Yes	No	No	No
Government Job (Self & Family)	No	No	Yes	No	No	No
Majority Tribe	No	No	Yes	No	No	No
Adjustments:						
Includes Pilot Neighborhoods	No	No	No	Yes	No	No
Excludes Misassigned Neighborhoods	No	No	No	No	Yes	No
Top-Code 10% Neighborhoods	No	No	No	No	No	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	No
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates from Equation 1, comparing property tax outcomes in Local and Central (the excluded category). The panels show the estimates from separate regressions with the outcome an indicator for compliance (Panel A) and revenues (Panel B), respectively. All regressions include fixed effects for house, time period, and randomization strata, and they cluster standard errors at the neighborhood level. Column 1 includes controls for age, age-squared, and sex, measured at midline. Column 2 controls for distance from schools (the one imbalanced covariate when comparing Local to Central in Table A2). Column 3 adds controls for having any job, a salaried job, and a government job, a family member with a government job, and belonging to the majority tribe. When including controls, we replace missing values in control variables with the mean for the entire sample and include a separate dummy (for each control variable) for the value being missing. Column 4 includes pilot neighborhoods, with time period and stratum values that reflect its implementation several months before the campaign and in a remote neighborhood. Column 5 excludes the neighborhood misassigned from CXL to Local during the campaign. Column 6 displays estimates from a regression on mean outcomes at the neighborhood-level, winsorizing the top 10% of neighborhoods, using robust standard errors, and assigning the minimum value for time period fixed effects to a neighborhood. We discuss these results in Section IVA.

TABLE A8: LOCAL V. CENTRAL: CONTROLLING FOR COLLECTOR CHARACTERISTICS

	Paid Property Tax								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Local	0.0331***	0.0378**	0.0374***	0.0435***	0.0340***	0.0338***	0.0328***	0.0358***	0.0456**
	(0.0072)	(0.0166)	(0.0073)	(0.0081)	(0.0078)	(0.0085)	(0.0091)	(0.0082)	(0.0207)
Age		-0.0002							0.0001
		(0.0005)							(0.0005)
Number of Possessions			0.0059						0.0058
			(0.0042)						(0.0048)
Years of Education				0.0032**					0.0032**
				(0.0013)					(0.0014)
Trust in Government (mean)					0.0030				0.0050
					(0.0049)				(0.0053)
Taxes Important						-0.0009			-0.0039
						(0.0078)			(0.0080)
Tax Ministry Important							-0.0000		0.0003
							(0.0062)		(0.0069)
Progressivity (mean)								0.0052	0.0104
								(0.0120)	(0.0130)
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.023	0.023	0.024	0.025	0.024	0.024	0.023	0.024	0.025
Observations	27764	26497	27453	27031	26489	27152	26361	27152	25443
Clusters	213	203	210	207	203	208	202	208	194
Control Mean	.075	.075	.075	.075	.075	.075	.075	.075	.075

Notes: This table reports estimates from Equation 1, comparing property tax outcomes in Local and Central (the excluded category), while additionally controlling for collector characteristics for which state and chief collectors have statistically significant differences in Columns 2–8. The value of collector characteristics are those of the chief in Local and the mean of those of the assigned collectors in Central. All regressions include fixed effects for randomization strata, time periods, and cluster standard errors at the neighborhood level. We discuss these results in Section IVA.

TABLE A9: LOCAL V. CENTRAL: EXEMPTION CATEGORIES

	Exempted	Incorrect Exemption	Senior	Widow	Government Pension	Handicapped	Exempt (by Coethnic)	Exempt (by Know Col.)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local	0.039*	-0.012	0.041***	-0.006	0.005	0.003**	0.041	-0.026
	(0.021)	(0.007)	(0.014)	(0.012)	(0.003)	(0.001)	(0.032)	(0.024)
Local X Coethnic							0.041	
							(0.040)	
Coethnic							-0.080***	
							(0.030)	
Local X Knows Collector								0.067*
								(0.038)
Knows Collector								0.064**
								(0.031)
Observations	13772	13771	13772	13772	13772	13772	7288	13772
Clusters	213	213	213	213	213	213	207	213
Central Mean	.264	.956	.126	.112	.013	.004	.314	.031
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows differences in the exemption rates of properties by chief and state collectors. Column 1 examines treatment effects on official exemptions. Column 2 reports whether third-party evaluations of exemption status diverged with the official designation. Columns 3–6 correspond to the different exemption categories: being senior (age 65+) in Column 3, being a widow in Column 4, receiving a government pension in Column 5 and being handicapped in Column 6. Columns 7 and 8 report exemptions by treatment and coethnicity between collectors and property owners and whether the collector and property owner know each other, respectively. All regressions include randomization stratum fixed effects and house fixed effects as well as the time fixed effects described in Section IV and standard errors are clustered at the neighborhood level. These results are discussed in Section IVA.

TABLE A10: LOCAL V. CENTRAL: AWARENESS OF OTHER TREATMENTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local	0.033***	0.030**	0.035***	0.029*	0.033***	0.033***	0.034**	0.034**
	(0.007)	(0.013)	(0.009)	(0.016)	(0.008)	(0.010)	(0.012)	(0.012)
Local X # Adjacent in Other Treatment (Strict)		0.003						
		(0.008)						
# Adjacent in Other Treatment (Strict)	0.004	0.003						
	(0.005)	(0.008)						
Local X # Adjacent in Other Treatment (Broad)				0.003				
				(0.006)				
# Adjacent in Other Treatment (Broad)			-0.001	-0.003				
			(0.004)	(0.005)				
Local X Length of Border Shared with Other Treatment (Strict)						-0.002		
						(0.030)		
Length of Border Shared with Other Treatment (Strict)					0.007	0.008		
					(0.015)	(0.029)		
Local X Length of Border Shared with Other Treatment (Broad)							0.004	0.004
							(0.018)	(0.018)
Length of Border Shared with Other Treatment (Broad)							-0.012	-0.012
							(0.020)	(0.020)
# Adjacent (Total)	-0.001	-0.001	0.001	0.001				
	(0.003)	(0.003)	(0.003)	(0.003)				
Length of Border (Total)					0.002	0.001	0.006	0.006
					(0.008)	(0.008)	(0.009)	(0.009)
Observations	27764	27764	27764	27764	27764	27764	27764	27764
Clusters	213	213	213	213	213	213	213	213
Central Mean	.068	.068	.068	.068	.068	.068	.068	.068
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table analyzes potential spillovers due to awareness of other types of tax collectors working in adjacent neighborhoods. The specifications follow Miguel and Kremer (2004) in controlling for the number of adjacent neighborhoods in different treatments (as well as the total number of adjacent neighborhoods). We evaluate two definitions of alternative treatments: the "strict" version codes adjacent neighborhoods as being in the alternative treatment if in Central (for a Local neighborhood) or Local (for a Central neighborhood); the "broad" version codes these (adjacent neighborhoods) as Central, CLI, or CXL (if Local) and Local or CXL (if Central). Due to campaign staggering across neighborhoods, we only consider exposure to treatments in adjacent neighborhoods in which collectors had already worked or were currently working, rather than neighborhoods that had been assigned to a different treatment but had not yet received tax collectors. Columns 1 and 3 report estimates of the effect of Local, controlling for the number of adjacent neighborhoods in the alternative treatment arm and total adjacent neighborhoods, for the strict and broad definitions, respectively. Columns 2 and 4 report estimates of the impact of Local collection with an interaction term for the number of adjacent neighborhoods assigned to the alternative treatment arm, controlling for the total number of adjacent neighborhoods, for strict and broad, respectively. Columns 5 and 7 report estimates of the impact of Local, controlling for length of neighborhood borders (in kilometers) shared with the alternative treatment and total length of borders, for strict and broad respectively. Columns 6 and 8 report estimates of the impact of Local collection with an interaction term for the length of neighborhood borders shared with neighborhoods assigned to the alternative treatment arm, controlling for length of neighborhood borders shared with the alternative treatment and total length of borders, for strict and broad, respectively. We include fixed effects for house type, randomization strata and time periods described in Section IV, and cluster standard errors at the neighborhood level. We discuss these results in Section IVA.

TABLE A11: LOCAL V. CENTRAL: FISCAL EXTERNALITIES

Dependent variable	$\hat{\beta}$	SE	R^2	N	$\bar{x}_{Central}$
Panel A: Informal Labor Taxes					
Salongo Extensive (Midline)	-0.031	0.032	0.057	13952	0.376
Salongo Intensive (Midline)	-0.240	0.247	0.025	13568	1.659
Salongo Extensive (Endline)	0.005	0.028	0.063	2413	0.404
Salongo Intensive (Endline)	0.459	0.445	0.051	2358	3.996
Panel B: Other Formal Taxes					
Vehicle Tax	0.013	0.008	0.049	2405	0.031
Market Vendor Fee	0.057***	0.017	0.046	2409	0.128
Business Tax	0.008	0.010	0.044	2409	0.043
Income Tax	0.037***	0.014	0.031	2406	0.095
Obsolete Tax	0.003	0.005	0.025	2387	0.014

Notes: Each row summarizes an OLS estimation of Equation 1, comparing Local and Central, with the dependent variable noted in the first column. $\hat{\beta}$ is the coefficient on the treatment indicator, followed by the cluster-robust standard error, R^2, number of observations, and $\bar{x}_{Central}$, the Central group mean. In Panel A, rows 1 and 2 (3 and 4) report *salongo* contributions along the extensive margin and intensive margin of hours, respectively, at midline (endline). In Panel B, the outcomes are self-reported payment of other formal taxes at endline. Obsolete tax is a poll tax, which existed in the past but not currently, to test the reliability of self-reports. All regressions include fixed effects for randomization strata, and cluster standard errors at the neighborhood level. Regressions using midline data include house type fixed effects, while those using endline data do not, as discussed in Section IV, because this affords analysis in a larger endline sample. The number of observations varies across regressions due to (*i*) outcomes being drawn from different surveys, and (*ii*) non-response for specific survey questions. We discuss these results in Section IVA.

TABLE A12: LOCAL V. CENTRAL: INFORMAL LABOR TAX SUBSTITUTION

	Salongo (Midline) (1)	*Salongo* Hours (Midline) (2)	*Salongo* (Endline) (3)	*Salongo* Hours (Endline) (4)
Panel A: Taxpayers				
Local	-0.026	-0.207	0.000	0.490
	(0.032)	(0.254)	(0.030)	(0.454)
Local X Paid Tax	-0.075**	-0.262	-0.051	-1.387
	(0.035)	(0.226)	(0.070)	(1.039)
Paid Tax	0.061**	-0.128	0.038	0.757
	(0.029)	(0.167)	(0.052)	(0.796)
Observations	13953	13569	2330	2278
Clusters	206	205	221	221
Central Mean (No Pay)	.372	1.685	.406	4.008
Panel B: Predicted Compliers				
Local	-0.014	-0.106	0.022	1.927
	(0.042)	(0.666)	(0.081)	(1.324)
Local X Predicted Complier	-0.035	-0.253	-0.097	-2.186
	(0.041)	(0.650)	(0.096)	(1.538)
Predicted Complier	0.090***	0.157	0.178**	2.651**
	(0.023)	(0.664)	(0.077)	(0.941)
Observations	9835	9726	583	568
Clusters	195	190	150	150
Central Mean (Pred. Non-Complier)	.355	1.697	.372	4.382
Time FE	Yes	Yes	No	No
House FE	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes

Notes: This table shows estimates from versions of Equation 1, comparing the Local arm to the Central arm (excluded group), where we include an interaction with verified property tax payment (Panel A) and predicted compliance (Panel B). Predicted compliance is defined as belonging to the top 25th percentile of values for the mean of predicted ease of payment and predicted willingness to pay, generated through the exercise described in Section VB. The outcome is informal labor tax (*salongo*) participation as measured in the midline and endline surveys. Columns 1 and 2 report *salongo* contributions along the extensive margin and intensive margin (hours contributed), respectively, at midline. Columns 3 and 4 report analogous estimates measured at endline. All regressions include fixed effects for house type and randomization strata and cluster standard errors at the neighborhood level. Columns 1 and 2 include time period fixed effects because they analyze midline data, as discussed in Section IV. We discuss these results in Section IVA.

TABLE A13: LOCAL V. CENTRAL: "TOTAL" TAX BURDEN (TAXES, BRIBES, *Salongo*)

	Paid Tax or Bribe		Paid Tax, Bribe, or *Salongo*	
	(extensive)	(intensive)	(extensive)	(intensive)
	(1)	(2)	(3)	(4)
Local	0.029*	0.127***	0.033***	0.105**
	(0.017)	(0.027)	(0.007)	(0.038)
Time FE	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes
Observations	27764	27138	27764	27138
Clusters	213	213	213	213
Control Mean	.234	.045	.063	.055

Notes: This table shows treatment effects on household payment of property taxes, bribes, and/or *salongo* labor contributions. Columns 1 and 3 show the extensive margin (i.e. dummies for paying taxes or bribes, or for paying taxes, bribes, or doing *salongo*). Columns 2 and 4 show the intensive margin of contributions, i.e. the total number of contributions (max = 3). These intensive-margin outcomes are standardized to facilitate interpretation of magnitudes.

TABLE A14: INVESTIGATING HAWTHORNE EFFECTS: AWARENESS OF MONITORING AND BRIBE-TAKING BEHAVIOR IN LOCAL

	Endline Collector Survey			*Midline Household Survey*		
	Chief Perception of Monitoring / Punishment for Bribe-Taking			Household Bribe Payment		
Chief characteristic	(1)	(2)	(3)	(4)	(5)	(6)
Knows Deposed Chiefs	-0.458**			0.0009		
	(0.228)			(0.0065)		
Knows 2016 Campaign		0.065			-0.0055	
		(0.288)			(0.0056)	
Neighborhood in 2016 Campaign			-0.040			-0.0014
			(0.230)			(0.0278)
House FE	No	No	No	Yes	Yes	Yes
Stratum FE	No	No	No	Yes	Yes	Yes
R^2	0.049	0.001	0.002	0.017	0.018	0.017
Observations	81	81	80	6393	6393	6492
Dep. Var. Mean	.043	.043	.043	.018	.018	.018

Notes: This table shows correlations between baseline chief/neighborhood characteristics and outcomes related to the acceptance of bribes in Local (i.e. neighborhoods with chief tax collection). Columns 1–3 examine an outcome drawn from a survey with chiefs conducted after the 2018 tax campaign, in which chiefs were asked to estimate the probability that collectors accepting bribes during the campaign would be sanctioned. Columns 4–6 examine bribe payment reported by citizens in the midline survey. *Knows Deposed Chiefs* and *Knows 2016 Campaign* come from a baseline survey conducted with chiefs, who reported whether they had ever heard of (*i*) a chief being deposed, and (*ii*) the 2016 property tax campaign, respectively. *Neighborhood in 2016 campaign* indicates neighborhoods randomly assigned to the 2016 property tax campaign, as measured in administrative data. We discuss these results in Section IVB.

TABLE A15: HETEROGENEOUS TREATMENT EFFECTS BY CHIEF CHARACTERISTICS: BRIBES

Chief Characteristic	$\hat{\beta}_1$	SE	$\hat{\beta}_2$	SE	$\hat{\beta}_3$	SE	N	$\bar{y}_{Control}$
Panel A: Demographics								
Age $> p50$	-0.003	0.004	0.010*	0.006	-0.006	0.005	12129	0.018
Wealth $> p50$	-0.001	0.003	0.015	0.009	-0.005	0.006	12129	0.017
Education $> p50$	0.002	0.004	-0.002	0.006	0.003	0.005	12129	0.015
Minority ethnic	0.001	0.004	0.004	0.009	-0.007	0.005	11978	0.018
Panel B: Chief Power / Experience								
Locality chief	0.005	0.005	-0.005	0.006	0.010*	0.006	10669	0.015
Customary chief	0.003	0.004	-0.002	0.006	-0.017**	0.007	12129	0.017
Chief for over 10 years	-0.001	0.004	0.005	0.006	-0.004	0.005	11978	0.016
Dynastic succession	0.001	0.004	-0.001	0.007	-0.003	0.005	11869	0.018
Remote neighborhood	0.011***	0.004	-0.015***	0.005	0.001	0.005	12129	0.014
Panel C: Political Ties								
Political party member	0.008**	0.004	-0.019***	0.007	0.010*	0.006	11978	0.014
Ruling party member	0.006	0.004	-0.022**	0.009	0.013*	0.008	11978	0.015
Opposition party member	0.003	0.004	-0.014***	0.009	0.002	0.006	11978	0.016
Has other gov position	0.002	0.004	-0.002	0.007	0.004	0.006	11978	0.018
Panel D: Views of Government								
Gov. trust $> p50$	0.007	0.005	-0.013*	0.007	0.005	0.005	12129	0.016
Tax ministry trust $> p50$	0.007	0.004	-0.015**	0.007	0.008	0.005	12129	0.015
Gov. performance $> p50$	0.001	0.004	0.001	0.007	-0.001	0.005	12129	0.016
Gov. responsiveness $> p50$	0.002	0.005	-0.000	0.006	-0.004	0.005	12129	0.017
Gov. integrity $> p50$	0.004	0.004	-0.005	0.006	0.003	0.006	12129	0.015
Panel E: Salience of Monitoring								
Knows deposed chiefs	0.001	0.005	-0.000	0.007	0.006	0.005	11978	0.017
Knows 2016 campaign	0.004	0.006	-0.004	0.008	-0.002	0.005	11869	0.014
Nbhd in 2016 campaign	0.003	0.004	-0.002	0.006	0.013	0.017	12077	0.015
Panel F: Citizens' Perceptions of Chief								
Trusted by citizens	0.004	0.005	-0.005	0.007	0.006	0.005	12129	0.017
Accessible to citizens	0.007	0.005	-0.010	0.007	0.002	0.005	12129	0.016
Active in chief role	0.003	0.005	-0.003	0.007	0.003	0.005	12129	0.018

Notes: This table shows heterogeneous treatment effects by a range of chief characteristics measured before the tax campaign. Specifically, each row summarizes the results from estimating the equation $y_{ijkt} = \beta_0 + \beta_1 Local_{jkt} + \beta_2 Local_{jkt} * Z_{jk}^{Chief} + \beta_3 Z_{jk}^{Chief} + \alpha_k + \theta_t + \varepsilon_{ijkt}$, where Z_{jk}^{Chief} indicates the corresponding characteristic of the neighborhood chief shown in the first cell of each row. y_{ijkt} is bribe payment, α_k are stratum fixed effects, and θ_t are time fixed effects. Standard errors are clustered at the neighborhood level (213 in total). All chief characteristics are dichotomized to maximize power for estimating heterogeneous treatment effects. Continuous variables are transformed into indicators to report above-median values of the characteristics (denoted by $> p50$). We discuss these results in Section IVB.

TABLE A16: THE COUNTERFACTUAL TO CHIEF BRIBE COLLECTION: PREDICTING CHIEF BRIBE PAYERS IN CENTRAL

	Compliance (Admin) (1)	Compliance (Admin) (2)	Paid Bribe (Endline) (3)
Panel A: Predicted Bribe Payment > 75th Percentile			
Predicted Bribe Payer	0.013	0.032	0.000
	(0.024)	(0.054)	(0.018)
Mean (Predicted Bribe Non-Payer)	.089	.089	.013
Panel B: Predicted Bribe Payment > 90th Percentile			
Predicted Bribe Payer	0.069*	0.146*	-0.020**
	(0.041)	(0.085)	(0.008)
Mean (Predicted Bribe Non-Payer)	.086	.086	.016
Visited Post-Registration Only	No	Yes	No
Time FE	No	No	No
House FE	No	No	No
Stratum FE	Yes	Yes	Yes
Observations	847	329	414
Clusters	109	95	102

Notes: This table provides evidence on the counterfactual of the increase in bribes in Local relative to Central. Specifically, it shows correlations between predicted chief bribe payment and tax and bribe payments in the Central treatment arm. Predicted bribe payment is constructed by regressing bribe payment (at endline) in Local on baseline household and property characteristics, retaining the variables with significant coefficients, and using these variables to predict bribe payment (at endline) in the Central treatment arm. This exercise simulates the likely bribe payers if Central neighborhoods had in fact been assigned to Local. The variables used in the final prediction exercise include: age of property owner, whether a family member of household members works for the government, whether the household possesses a radio, trust in the provincial government, whether the respondent knows the neighborhood chief, has the chief's phone number, and attends the same church as the chief. The Predicted Bribe Payer variable is an indicator for the predicted value being greater than the 75th (Panel A) or 90th percentile (Panel B). Tax compliance is measured using administrative tax data, post-registration visits using the midline survey, and bribe payment using the endline survey (our preferred measure using a local code for bribes, discussed in Section IVB). All regressions include fixed effects for randomization strata and cluster standard errors at the neighborhood level. Column 2 restricts the sample to households that received visits from tax collectors after registration. The number of observations for bribe outcomes is smaller than the full endline sample because this variable was only collected among households who reported at least one visit from tax collectors after property registration. We discuss these results in Section IVB.

A1.5 Additional Exhibits for Paper Section 7 — Mechanisms

TABLE A17: LOCAL V. CENTRAL: TAX VISITS — NO HOUSE FIXED EFFECTS

	Visited by Collector	Number of Visits by Collector	Other Contact with Collector	Instances of Other Contact
	(1)	(2)	(3)	(4)
Local	-0.008	0.016	0.008	0.019
	(0.026)	(0.046)	(0.007)	(0.013)
Time FE	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes
Observations	18162	18151	3513	3513
Clusters	209	209	206	206
Mean	.417	.552	.025	.039

Notes: This table reports estimates from Equation 1, comparing the tax visits collectors made after registration in Local and Central (the excluded category). All regressions include fixed effects for randomization strata and time periods described in Section IV, and cluster standard errors at the neighborhood level. Columns 1 and 2 report differences in tax visits by collectors — after the registration visit — by the extensive and intensive margins, respectively. Columns 3 and 4 report differences in other contact with collectors outside of the tax campaign, as reported by citizens, by the intensive and extensive margins, respectively. We discuss these results in Section VA.

TABLE A18: CENTRAL V. CENTRAL + LOCAL INFORMATION ROBUSTNESS: DIFFERENT APPROACHES TO TIME IMBALANCE

	No Adjustment (1)	Two Month Fixed Effects (2)	Shift Median Two Month Fixed Effects (3)	Interaction Weighted Estimator (4)	One Month Fixed Effects (5)	Time Restriction (6)	Coarsened Exact Matching (7)
Panel A: Compliance							
Central Plus Local Info	-0.001	0.024**	0.019**	-0.004	0.024**	0.019**	0.041**
	(0.011)	(0.009)	(0.009)	(0.008)	(0.010)	(0.009)	(0.016)
Observations	23911	20636	19767	32754	23911	18834	8575
Clusters	190	165	161	190	190	150	72
Central Mean	.068	.051	.057	.051	.068	.055	.024
Panel B: Revenues							
Central Plus Local Info	-10.315	40.178*	39.558**	-30.749	57.325**	37.204*	52.277
	(26.089)	(20.481)	(19.509)	(21.816)	(20.830)	(20.452)	(34.337)
Observations	23911	20176	20507	31963	23911	18834	8575
Clusters	190	162	160	190	190	150	72
Central Mean	192.891	155.747	138.945	155.747	192.891	156.774	61.224
One Month FE	No	No	No	No	Yes	No	No
Two Month FE	No	Yes	Yes	Yes	No	No	No
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table displays alternative approaches for addressing time imbalance in the comparison of the Central + Local Information (CLI) arm to the Central arm, the excluded category. Panel A reports impacts on compliance, and Panel B reports impacts on revenues. Column 1 makes no adjustments. Column 2 includes the time period fixed effects described in Section IV. Column 3 includes time period fixed effects defined by selecting the median estimate among all permutations of the start date (Figure A6). Column 4 implements an interaction-weighted estimator, following Gibbons et al. (2018), in which time periods defined as in Column 2 are not included as fixed effects but interacted with the treatment indicator and the estimate is the weighted average of the coefficient on the interaction terms, weighted by the number of observations in each period. Column 5 includes one-month fixed effects. Column 6 trims the sample to periods when both treatment arms were in operation. Column 7 implements coarsened exact matching (Iacus et al., 2012). All regressions include fixed effects for house type and randomization strata and cluster standard errors at the neighborhood level. We discuss these results in Section VB.

TABLE A19: CENTRAL V. CENTRAL + LOCAL INFORMATION ROBUSTNESS: CONTROLLING FOR IMBALANCED MIDLINE COVARIATES

	Compliance	Revenues	Visited	Visits	Compliance	Compliance
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Including Imbalanced Midline Covariates						
Central Plus Local Info	0.024**	52.263**	-0.008	-0.021	0.021	0.030**
	(0.011)	(25.449)	(0.034)	(0.055)	(0.016)	(0.011)
Local						0.065***
						(0.009)
Controls for Imbalanced Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Visit Control	No	No	No	No	Yes	No
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10064	10064	10051	10048	3864	16436
Clusters	155	155	155	155	150	253
Central Mean	.059	159.808	.393	.51	.1	.059
Test CLI=Local p-value						0.002
Panel B: Excluding House Fixed Effects						
Central Plus Local Info	0.024**	24.241	-0.018	-0.029	0.027*	0.023**
	(0.009)	(23.476)	(0.028)	(0.044)	(0.014)	(0.009)
Local						0.045***
						(0.007)
Visit Control	No	No	No	No	Yes	No
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	No	No	No	No	No	No
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20629	20629	13879	13872	5281	33731
Clusters	165	165	163	163	161	267
Central Mean	.051	150.714	.387	.497	.097	.052
Test CLI=Local (p-value)						0.01

Notes: This table compares the Central + Local Information (CLI) arm to the Central arm, the excluded category, controlling for the characteristics imbalanced at midline — sex of property owner, whether property owner is salaried, and distance to government buildings and market — as shown in Table A2 (Panel A) and excluding house type fixed effects (Panel B). Columns 1, 5, and 6 report impacts on compliance. Column 2 reports effects on revenues. Columns 3 and 4 report differences in tax visits by collectors after registration by the extensive and intensive margins, respectively. All regressions include fixed effects randomization strata and time periods, and cluster standard errors at the neighborhood level. Column 5 restricts to the subsample of properties that received any tax visits after registration. Column 6 includes a dummy for the Local treatment in the regression. The bottom row reports the p-value from a test for equality between the CLI and Local. We discuss these results in Section VB.

TABLE A20: LOCAL V. CENTRAL + LOCAL INFORMATION: HETEROGENEOUS TREATMENT EFFECTS BY OWNER PRESENT AT REGISTRATION

	(1)	(2)	(3)	(4)
Local	0.021**	0.027**	0.028**	0.029**
	(0.009)	(0.009)	(0.009)	(0.009)
Local X Owner Present at Registration	0.001	0.005	0.005	0.014
	(0.009)	(0.009)	(0.009)	(0.010)
Owner Present at Registration	0.039***	0.033***	0.035***	0.051***
	(0.007)	(0.006)	(0.007)	(0.008)
Observations	28875	27767	27767	23805
Clusters	221	213	213	213
Central Mean (Owner Not Present)	.036	.035	.035	.035
Time FE	No	Yes	Yes	Yes
House FE	No	No	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes
Exempt Excluded	No	No	No	Yes

Notes: This table reports estimates from Equation 1, comparing property tax compliance in Local and Central (the excluded category), where we include an interaction with an indicator for the owner of the property being present at registration. All regressions include fixed effects for randomization strata and cluster standard errors at the neighborhood level. Column 1 regressions do not include time period fixed effects described in Section IV, while those in other columns include them. Regressions in Columns 1–2 do not include house fixed effects while Column 3 includes them. Regressions in Column 4 exclude exempt properties. The data include all properties registered by tax collectors merged with the government's property tax database. We discuss these results in Section VB.

TABLE A21: THE VALUE OF CHIEFS' INFORMATION — NO HOUSE FIXED EFFECTS

	Visited (1)	Compliance (2)	Visited (3)	Compliance (4)	Visited (5)	Compliance (6)	Visited (7)	Compliance (8)
Panel A: Ease of payment								
Ease of payment	0.046***	0.055***	0.029**	0.043***				
	(0.012)	(0.007)	(0.014)	(0.008)				
Predicted ease of payment					0.054**	0.045***	0.013	0.040***
					(0.017)	(0.012)	(0.017)	(0.007)
Wall quality			0.027**	0.017**	0.017*	0.008	0.021**	0.011**
			(0.012)	(0.007)	(0.010)	(0.006)	(0.010)	(0.004)
Roof quality			0.005	0.000	0.003	-0.004	0.018**	-0.010
			(0.006)	(0.002)	(0.006)	(0.004)	(0.008)	(0.006)
Erosion threat			0.017	-0.003	0.002	-0.007	-0.000	-0.005
			(0.011)	(0.004)	(0.011)	(0.007)	(0.010)	(0.004)
Observations	5574	8135	4551	5150	5748	5763	4998	5004
Clusters	79	80	66	66	93	93	80	80
Mean	.376	.072	.352	.065	.435	.103	.41	.059
Panel B: Willingness to pay								
Willingness to pay	0.035**	0.037***	0.033**	0.038***				
	(0.011)	(0.007)	(0.012)	(0.008)				
Predicted willingness to pay					0.045**	0.036***	0.007	0.032***
					(0.016)	(0.010)	(0.015)	(0.009)
Wall quality			0.025*	0.017**	0.018*	0.009	0.021**	0.011**
			(0.013)	(0.008)	(0.009)	(0.006)	(0.010)	(0.005)
Roof quality			0.011	0.001	0.004	-0.003	0.018**	-0.010
			(0.008)	(0.002)	(0.006)	(0.004)	(0.008)	(0.006)
Erosion threat			0.016	-0.004	0.002	-0.006	-0.000	-0.005
			(0.012)	(0.005)	(0.011)	(0.007)	(0.010)	(0.005)
Observations	3933	5521	3929	4461	5748	5763	4998	5004
Clusters	50	50	50	50	93	93	80	80
Mean	.357	.062	.357	.066	.435	.103	.41	.059
Treatment	CLI	CLI	CLI	CLI	Local	Local	Central	Central
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

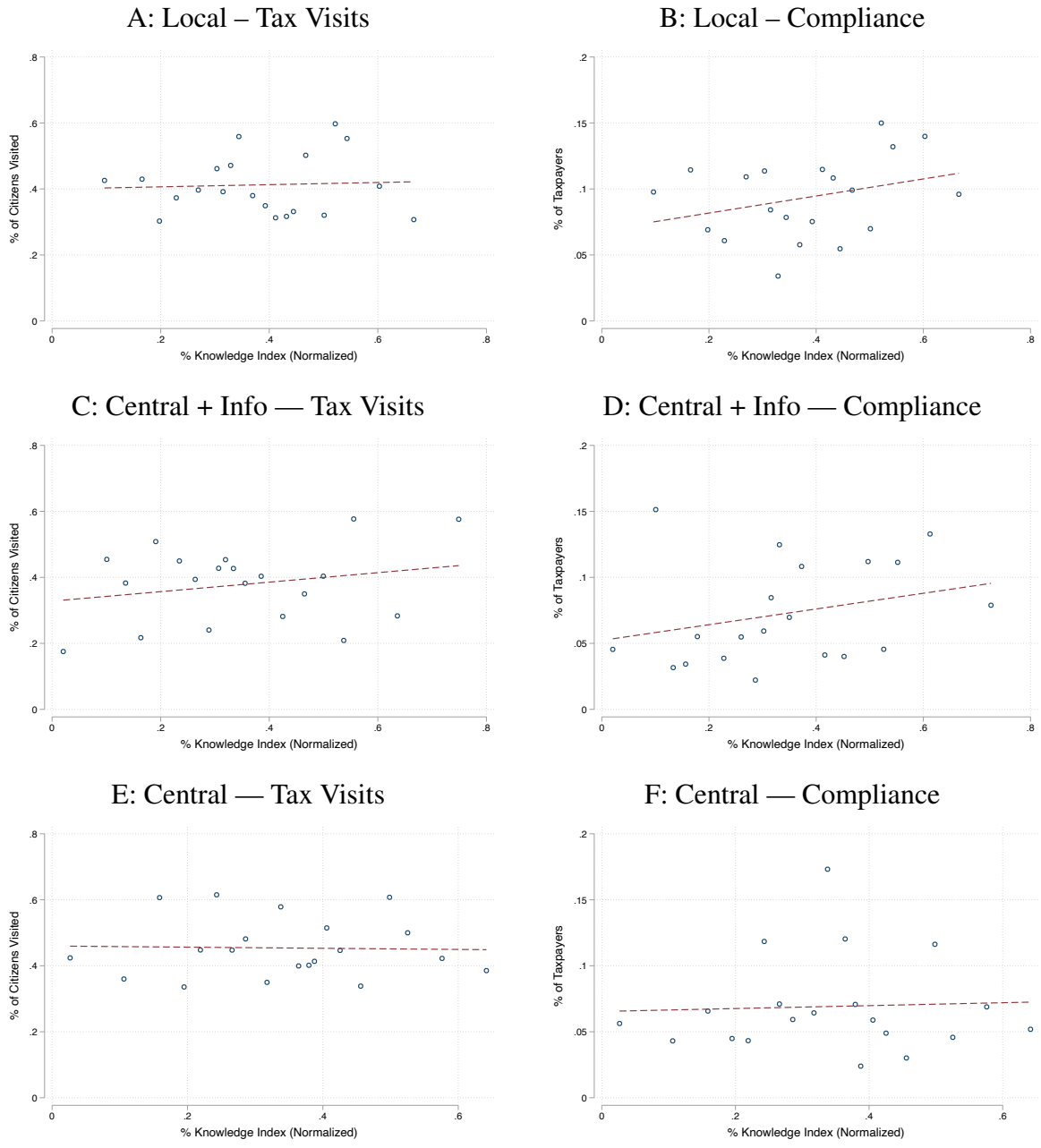
Notes: This table explores the extent to which chiefs' recommendations in Central + Local Information (CLI) predict tax visits after registration and tax payment, while excluding house fixed effects as a robustness check. Columns 1–4 show correlations in CLI between chiefs' recommendations and outcomes. Columns 5–8 report correlations between predicted propensity measures described in Section VB and outcomes in the Local (Columns 5 and 6) and the Central (Columns 7 and 8) arms. Columns 1, 3, 5, and 7 show correlations between propensity and visits; Columns 2, 4, 6, and 8 show correlations between propensity and compliance. All regressions include randomization stratum fixed effects and cluster standard errors at the neighborhood level. Columns 3, 4, and 6–8 include controls for visible household characteristics. We discuss these results in Section VB.

TABLE A22: THE VALUE OF CHIEFS' INFORMATION — COMPARING TREATMENTS

	CLI vs. Central		CLI vs. Local	
	Visited	Compliance	Visited	Compliance
	(1)	(2)	(3)	(4)
Panel A: Predicted ease of payment				
Predicted ease of payment	0.016	0.024**	0.037*	0.038**
	(0.020)	(0.009)	(0.021)	(0.011)
CLI X Predicted ease of payment	0.081**	0.028**	0.059*	0.011
	(0.037)	(0.012)	(0.030)	(0.015)
CLI	-0.061*	-0.002	-0.084**	-0.047**
	(0.034)	(0.012)	(0.033)	(0.016)
Observations	8396	8407	8556	8575
Clusters	139	139	144	144
Comparison Group Mean	.413	.061	.449	.112
Panel B: Predicted willingness to pay				
Predicted willingness	0.025	0.026**	0.033	0.025**
	(0.021)	(0.008)	(0.020)	(0.010)
CLI X Predicted willingness	0.027	0.013	0.036	0.021
	(0.041)	(0.011)	(0.029)	(0.014)
CLI	-0.031	0.006	-0.079*	-0.059**
	(0.044)	(0.017)	(0.042)	(0.018)
Observations	8396	8407	8556	8575
Clusters	139	139	144	144
Comparison Group Mean	.413	.061	.449	.112
House Char. Controls	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes

Notes: This table explores the extent to which the content of chiefs' recommendations in Central + Local Information (CLI) predict tax visits after registration and tax payment differentially across treatments. Columns 1–2 compare CLI to Central, regressing outcomes of receiving a post-registration visit and paying the tax, respectively, on the predicted ease of payment measure (Panel A) and willingness to pay measure (Panel B) described in Section VB, an indicator for the CLI treatment arm, and their interaction. Columns 3–4 repeat the same exercise comparing CLI to Local. All regressions include house type and randomization stratum fixed effects, controls for observable household characteristics (wall quality, roof quality, and erosion threat), and cluster standard errors at the neighborhood level. We discuss these results in Section VB

FIGURE A7: TAX VISITS AND COMPLIANCE BY CHIEF KNOWLEDGE OF CITIZENS



Notes: This figure shows the relationship between chiefs' knowledge of the inhabitants of the neighborhood and (*i*) the percent of property owners who received a tax visit after registration (Panels A, C, and E), and (*ii*) the level of tax compliance (Panels B, D, and F). Chiefs' knowledge of the inhabitants of the neighborhood is measured by the percentage of correct answers when asked to provide the name, education level, and occupation of a randomly selected group property owners. We show these relationships for neighborhoods assigned to Local in Panels A and B as well as for neighborhoods assigned to CLI and Central tax collection in Panels C and D, and E and F, respectively. Table A23 analyzes these relationships in a regression framework. We discuss these results in Section VB.

TABLE A23: TAX VISITS AND COMPLIANCE BY CHIEF KNOWLEDGE OF CITIZENS

	CLI		Central		Local	
	Visits	Compliance	Visits	Compliance	Visits	Compliance
	(1)	(2)	(3)	(4)	(5)	(6)
Chief Info > Median	0.010	0.028*	-0.020	-0.007	-0.016	0.024*
	(0.043)	(0.017)	(0.041)	(0.012)	(0.034)	(0.012)
Observations	79	80	110	110	111	111
Mean	.377	.073	.454	.069	.412	.093

Notes: This table shows the relationship between city chiefs' knowledge of the inhabitants of the neighborhood and (*i*) the percent of property owners who received a tax visit after registration (Columns 1, 3, and 5), and (*ii*) the level of tax compliance (Columns 2, 4, and 6). Chiefs' knowledge of the inhabitants of the neighborhood is measured by the percentage of correct answers when asked to provide the name, education level, and occupation of a randomly selected group property owners. We show these relationships for neighborhoods assigned to (*i*) Central + Local Information (Columns 1–2), where state collectors did consult with chiefs, (*ii*) Central (Columns 3–4), where state collectors did not consult with chiefs — a placebo check — and (*iii*) Local (Columns 5–6), where chiefs themselves collected taxes. We discuss these results in Section VB.

TABLE A24: COLLECTOR OUTCOMES AS A FUNCTION OF DISTANCE TO THEIR OWN NEIGHBORHOODS

	State Collectors		Chief Collectors	
	Compliance	Revenue (in CF)	Compliance	Revenue (in CF)
	(1)	(2)	(3)	(4)
Distance (state collector)	-0.006***	-12.584**		
	(0.002)	(5.894)		
Distance (chief collector)			-0.005	3.071
			(0.019)	(62.236)
Time FE	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes
Stratum FE	No	No	No	No
Observations	22398	22398	13880	13880
Clusters	183	183	107	107
Mean	.066	172.966	.0940	251.686

Notes: This table estimates the relationship between tax compliance (Columns 1 and 3) or tax revenue (Columns 2 and 4) and the distance between collectors' houses and the neighborhoods in which they worked. We estimate this relationship for state collectors in Central and CLI by calculating the average distance for the two randomly assigned collectors (Columns 1 and 2). The relationship for chief collectors is reported in Columns 3 and 4 for completeness, though there is little variation for chief collectors who hailed from the neighborhoods in which they taxed. All regressions include house type fixed effects as well as the time fixed effects described in Section IV. We cluster standard errors at the neighborhood level. We discuss these results in Section VB.

TABLE A25: LOCAL V. CENTRAL: STATE COLLECTORS WORKING NEAR THEIR HOMES

	State Collectors Working Near Home		State Collectors Working Far from Home	
	Compliance	Revenue (in CF)	Compliance	Revenue (in CF)
	(1)	(2)	(3)	(4)
Panel A: Chiefs v. State Collectors in Central				
Local	0.027**	63.062**	0.034***	66.977***
	(0.012)	(31.702)	(0.009)	(24.605)
Observations	17225	17225	24635	24635
Clusters	142	142	199	199
Central Mean	.069	202.237	.062	176.298
Panel B: Chiefs v. State Collectors in Central and CLI				
Local	0.031**	73.158**	0.038***	86.362***
	(0.013)	(33.833)	(0.007)	(18.763)
Observations	17448	17448	28874	28874
Clusters	153	153	237	237
Central Mean	.055	178.929	.051	141.706
Time FE	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes
Strata FE	No	No	No	No

Notes: This table estimates Equation 1 using as the dependent variable whether households paid the property tax (Columns 1 and 3) and the amount of revenues collected (Columns 2 and 4). It includes state collectors in Central (Panel A) and in Central and CLI (Panel B) as the comparison group. We include Panel B, lumping Central and CLI, to increase the number of state collectors randomly assigned to work near their homes in the analysis. Columns 1 and 2 compare chief collection to state tax collection in cases where at least one assigned state collector lived nearby. We define "near" as the maximum distance between a chief's house and the neighborhood in which they taxed, which is 1.59 km in the data. Columns 3 and 4 compare chief collection to state tax collection in cases where no assigned state collector lived nearby. All regressions include house type and the time fixed effects described in Section IV and cluster standard errors at the neighborhood level. We do not include fixed effects for randomization strata as a large share of strata do not contain a neighborhood from each comparison group (49% of strata include only one treatment when comparing Local to Central near home, 30% include only one when comparing Local to Central and CLI near home). We discuss these results in Section VB.

TABLE A26: LOCAL V. CENTRAL: COLLECTION DURING PROPERTY REGISTRATION

	Collection Outcomes during Registration Visit					
	Compliance			Revenues		
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-0.001	-0.001	-0.000	-2.564	-2.850	-1.593
	(0.002)	(0.002)	(0.001)	(4.278)	(4.334)	(4.059)
Time FE	No	No	Yes	No	No	Yes
House FE	No	Yes	Yes	No	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	28872	28872	27764	28872	28872	27764
Clusters	221	221	213	221	221	213
Central Mean	.006	.006	.006	16.116	16.116	15.657

Notes: This table estimates Equation 1 using as the dependent variable whether households paid the property tax during the property registration (Columns 1–3) and the revenue collected (Columns 4–6). As described in the text, collectors were instructed to solicit the tax at the end of each registration visit with households. During property registration, collectors followed a linear property-by-property route through neighborhoods, as demonstrated in Figure A8, meaning that collectors could not selectively target taxpayers at this stage of the campaign. All regressions include randomization stratum fixed effects and cluster standard errors at the neighborhood level. Columns 2, 3, 5, and 6 include house type fixed effects. Columns 3 and 6 include time fixed effects described in Section IV. We discuss these results in Section VC.

FIGURE A8: COLLECTORS' ROUTE THROUGH SAMPLE NEIGHBORHOOD DURING PROPERTY REGISTRATION.

Notes: This map shows the linear, property-by-property route taken by collectors in a sample neighborhood in the Quartier of Malanji. Due to error in GPS measures, some points appear slightly outside of the neighborhood (or across the street). This figure is discussed in Section VC.

TABLE A27: HETEROGENEITY BY CHIEF CHARACTERISTICS: TAX COMPLIANCE

	$\hat{\beta}_1$	SE	$\hat{\beta}_2$	SE	$\hat{\beta}_3$	SE	N	$\bar{y}_{Control}$
Household Survey Data								
Panel A: Citizens' Perceptions of Chief								
Trusted by citizens	0.033***	0.009	-0.001	0.014	0.014	0.011	27764	0.056
Accessible to citizens	0.025**	0.011	0.016	0.016	0.006	0.013	27764	0.062
Active in chief role	0.024***	0.009	0.027*	0.016	0.017	0.012	27764	0.057
Chief Survey and Administrative Data								
Panel B: Demographics								
Age $> p50$	0.038**	0.012	-0.011	0.015	-0.005	0.012	27764	0.064
Wealth $> p50$	0.036***	0.008	-0.017	0.020	0.014	0.014	27764	0.064
Education $> p50$	0.017	0.011	0.036**	0.015	-0.023*	0.012	27764	0.073
Minority ethnic	0.042***	0.009	-0.041**	0.021	0.012	0.018	27453	0.059
Panel C: Chief Power / Experience								
Locality chief	0.043***	0.012	-0.005	0.016	0.002	0.013	24695	0.057
Customary chief	0.041***	0.007	-0.043	0.026	0.024	0.025	27764	0.061
Chief for $>$ 10 years	0.021**	0.009	0.023	0.016	0.002	0.011	27453	0.051
Dynastic succession	0.044***	0.008	-0.047*	0.024	0.046**	0.021	27323	0.056
Remote neighborhood	0.028***	0.010	0.009	0.015	-0.013	0.012	27764	0.069
Panel D: Political Ties								
Political party member	0.030***	0.010	0.009	0.017	-0.014	0.012	27453	0.070
Ruling party member	0.028***	0.008	0.023	0.019	-0.026*	0.014	27453	0.068
Opposition party member	0.034***	0.008	-0.008	0.025	0.003	0.019	27453	0.064
Has other gov. position	0.036***	0.008	-0.011	0.016	0.012	0.014	27453	0.066
Panel E: Views of Government								
Gov. trust $> p50$	0.031***	0.009	0.005	0.017	-0.006	0.013	27764	0.061
Tax ministry trust $> p50$	0.037***	0.009	-0.010	0.017	0.002	0.012	27764	0.062
Gov. performance $> p50$	0.038***	0.009	-0.017	0.016	0.009	0.013	27764	0.061
Gov. responsiveness $> p50$	0.031***	0.011	0.010	0.017	-0.030**	0.013	27764	0.063
Gov. integrity $> p50$	0.040***	0.010	-0.012	0.015	-0.020*	0.011	27764	0.070
Panel F: Salience of Monitoring								
Knows deposed chiefs	0.029**	0.011	0.010	0.017	-0.016	0.013	27453	0.059
Knows 2016 campaign	0.035**	0.016	-0.003	0.019	0.025*	0.014	27323	0.052
Nbhd in 2016 campaign	0.027**	0.013	0.011	0.016	0.047	0.046	27626	0.065

Notes: This table shows heterogeneous treatment effects by a range of chief characteristics measured before the tax campaign. Specifically, each row summarizes the results from estimating the equation $y_{ijkt} = \beta_0 + \beta_1 Local_{jkt} + \beta_2 Local_{jkt} * Z_{jk}^{Chief} + \beta_3 Z_{jk}^{Chief} + \alpha_k + \theta_t + \varepsilon_{ijkt}$, where Z_{jk}^{Chief} indicates the corresponding characteristic of the neighborhood chief shown in the first cell of each row. y_{ijkt} is tax compliance, α_k are stratum fixed effects, and θ_t are time fixed effects. Standard errors are clustered at the neighborhood level (213 in total). All chief characteristics are 0-1 to maximize power for estimating heterogeneous treatment effects. Continuous variables are transformed into indicators to report above-median values of the characteristics (denoted by $> p50$). Panel A includes variables derived from household baseline survey questions about the neighborhood chief. Panels B–F include variables derived from pre-campaign surveys with chiefs as well as administrative data (on customary zones, remoteness, and the 2016 tax campaign). This table is discussed in Section VC.

TABLE A28: FLIER MESSAGE EFFECTS ON TAX COMPLIANCE

	Tax Compliance			Tax Revenue (in CF)		
	(1)	(2)	(3)	(4)	(5)	(6)
Local	0.036***			107.822***		
	(0.008)			(31.185)		
Central Deterrence		0.013*	0.014*		42.705	43.318*
		(0.007)	(0.007)		(25.976)	(25.713)
Local Deterrence		0.010	0.012*		12.997	16.819
		(0.007)	(0.007)		(20.260)	(20.118)
Central Public Goods		0.005	0.005		7.552	7.263
		(0.007)	(0.007)		(20.788)	(20.351)
Local Public Goods		0.006	0.008		30.102	34.208
		(0.007)	(0.007)		(25.280)	(24.843)
Trust		0.010	0.011		28.547	30.866
		(0.007)	(0.007)		(22.949)	(22.850)
Observations	4783	6796	6796	4783	6796	6796
Mean	.012	.024	.024	30.326	59.64	59.64
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	No	No	Yes	No	No
Strata FE	Yes	No	No	Yes	No	No
Neighborhood FE	No	No	Yes	No	No	Yes

Notes: This table reports estimates from a regression of tax compliance (Columns 1–3) and tax revenue (Columns 4–6) on indicators for assignment to the Local treatment or the Central arm (Columns 1 and 4), or on indicators for the randomized messages printed on the tax letters distributed at registration (Columns 2–3 and 5–6). Section A2.2 provides descriptions of the central deterrence, local deterrence, central public goods, local public goods, and trust treatment messages. The excluded category in all regressions analyzing fliers is the control message "It is important to pay the property tax." All regressions include type of house fixed effects. Columns 1 and 4 include geographic randomization stratum fixed effects and the time fixed effects described in Section IV. Columns 3 and 6 include neighborhood fixed effects (tax message treatment randomization strata). The data are restricted to the subsample of properties subject to randomized messages on tax letters, which were introduced toward the end of the property tax campaign. We discuss these results in Section VC.

TABLE A29: LOCAL V. CENTRAL: INTERACTIONS WITH FLIER MESSAGES

	Tax Compliance		Tax Revenue	
	(1)	(2)	(3)	(4)
Panel A: Central Deterrence Message				
Local	0.052**	0.054**	179.273**	196.565**
	(0.017)	(0.018)	(53.603)	(60.449)
Central Deterrence	0.008	0.008	17.214	16.158
	(0.007)	(0.007)	(13.942)	(14.137)
Local X Central Deterrence	0.008	0.010	44.815	51.255
	(0.015)	(0.016)	(66.115)	(71.207)
Observations	1675	1580	1675	1580
Mean	.034	.035	95.343	98.544
Panel B: Local Deterrence Message				
Local	0.034**	0.032*	69.613**	66.327*
	(0.016)	(0.018)	(30.153)	(32.933)
Local Deterrence	0.008	0.008	14.513	14.541
	(0.008)	(0.008)	(13.338)	(13.326)
Local X Local Deterrence	0.007	0.010	0.444	6.039
	(0.015)	(0.016)	(34.416)	(36.918)
Observations	1682	1585	1682	1585
Mean	.033	.035	77.170	80.631
Panel C: Central Public Goods Message				
Local	0.043**	0.043**	89.392**	89.044**
	(0.013)	(0.015)	(25.733)	(28.054)
Central Public Goods	0.008	0.008	21.771**	21.797**
	(0.005)	(0.005)	(9.730)	(9.695)
Local X Central Public Goods	-0.011	-0.010	-45.274	-43.619
	(0.013)	(0.014)	(35.695)	(38.435)
Observations	1674	1581	1674	1581
Mean	.027	.028	64.695	67.236
Panel D: Local Public Goods Message				
Local	0.035**	0.037**	65.192*	81.790**
	(0.014)	(0.015)	(35.734)	(37.007)
Local Public Goods	0.012	0.012	66.663	65.890
	(0.008)	(0.008)	(47.133)	(47.163)
Local X Local Public Goods	-0.010	-0.008	-53.038	-48.424
	(0.017)	(0.018)	(65.423)	(68.030)
Observations	1674	1579	1674	1579
Mean	.03	.031	87.336	91.324
Panel E: Trust Message				
Local	0.041**	0.040**	95.835**	95.705**
	(0.017)	(0.018)	(33.016)	(35.821)
Trust	0.011	0.011	29.969	30.158
	(0.009)	(0.009)	(21.096)	(21.255)
Local X Trust	-0.004	-0.002	-13.603	-9.882
	(0.020)	(0.021)	(50.680)	(53.911)
Observations	1689	1598	1689	1598
Mean	.032	.033	80.403	83.730
House FE	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes
Strata FE	Yes	Yes	Yes	Yes

Notes: This table reports estimates from a version of Equation 1, comparing the Local to the Central arm, including interactions with indicators for flier messages printed on tax letters distributed at registration. Section A2.2 provides descriptions of the central deterrence, local deterrence, central public goods, local public goods, and trust treatment messages. The excluded flier message category is the control message "It is important to pay the property tax." The dependent variable is tax compliance in Columns 1 and 2 and tax revenue in Columns 3 and 4. All columns include house fixed effects and randomization stratum fixed effects and Columns 2 and 4 also include the time fixed effects described in Section IV. The data are restricted to the subsample of properties subject to randomized messages on tax letters. We estimate the effects of flier messages within the Local arm in Table A30. This table is discussed in Section VC.

TABLE A30: LOCAL: INTERACTIONS WITH FLIER MESSAGES

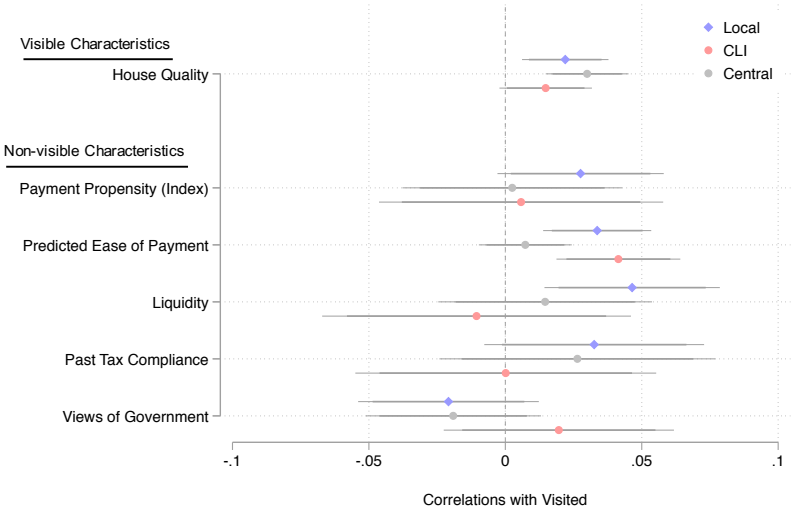
	Tax Compliance		Tax Revenue	
	(1)	(2)	(3)	(4)
Panel A: Central Deterrence Message				
Central Deterrence	0.016	0.018	62.332	66.948
	(0.014)	(0.015)	(63.495)	(67.708)
Observations	1159	1064	1159	1064
Mean	.046	.048	130.889	138.816
Panel B: Local Deterrence Message				
Local Deterrence	0.016	0.020	18.070	24.478
	(0.013)	(0.014)	(32.611)	(35.518)
Observations	1164	1067	1164	1067
Mean	.045	.048	105.928	113.683
Panel C: Central Public Goods Message				
Central Public Goods	-0.003	-0.001	-23.547	-22.095
	(0.012)	(0.013)	(34.274)	(37.016)
Observations	1155	1062	1155	1062
Mean	.036	.039	86.407	92.09
Panel D: Local Public Goods Message				
Local Public Goods	0.002	0.004	14.950	18.376
	(0.015)	(0.016)	(44.532)	(47.726)
Observations	1152	1057	1152	1057
Mean	.039	.042	109.635	117.597
Panel E: Trust Message				
Trust	0.007	0.010	16.087	19.426
	(0.018)	(0.019)	(46.870)	(50.043)
Observations	1173	1082	1173	1082
Mean	.042	.044	106.82	113.956
House FE	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes
Strata FE	Yes	Yes	Yes	Yes

Notes: This table reports estimates from regressions of compliance and revenues on indicators for flier messages printed on tax letters distributed at registration, restricted to the Local arm only. Section A2.2 provides descriptions of the central deterrence, local deterrence, central public goods, local public goods, and trust treatment messages. The excluded flier message category is the control message "It is important to pay the property tax." The dependent variable is tax compliance in Columns 1 and 2 and tax revenue in Columns 3 and 4. All columns include house fixed effects and randomization stratum fixed effects, and Columns 2 and 4 also include the time fixed effects described in Section IV. The data contain only properties subject to randomized messages on tax letters. Table A29 reports estimates from comparisons with the Central arm by flier message.

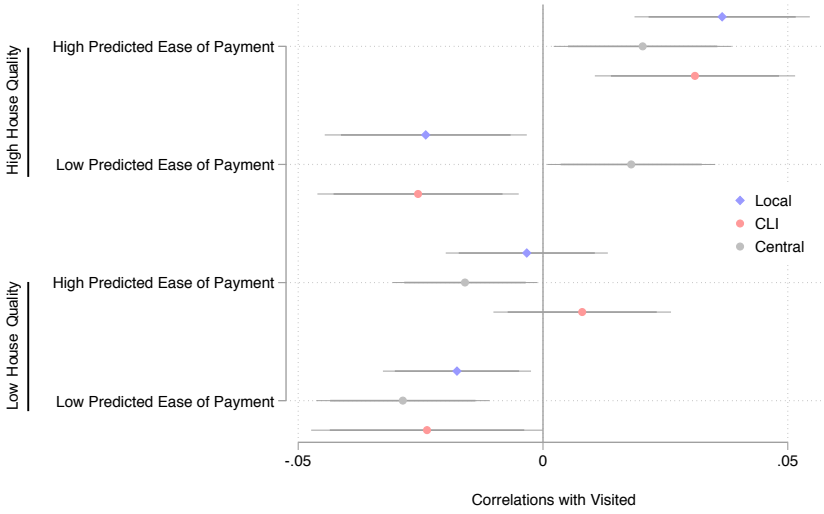
A1.6 Additional Exhibits for Paper Section 8 — Distributional Impacts

FIGURE A9: CHARACTERISTICS OF HOUSEHOLDS VISITED BY TAX COLLECTORS AFTER REGISTRATION WITHIN TREATMENTS

A: Visible and Non-Visible Characteristics



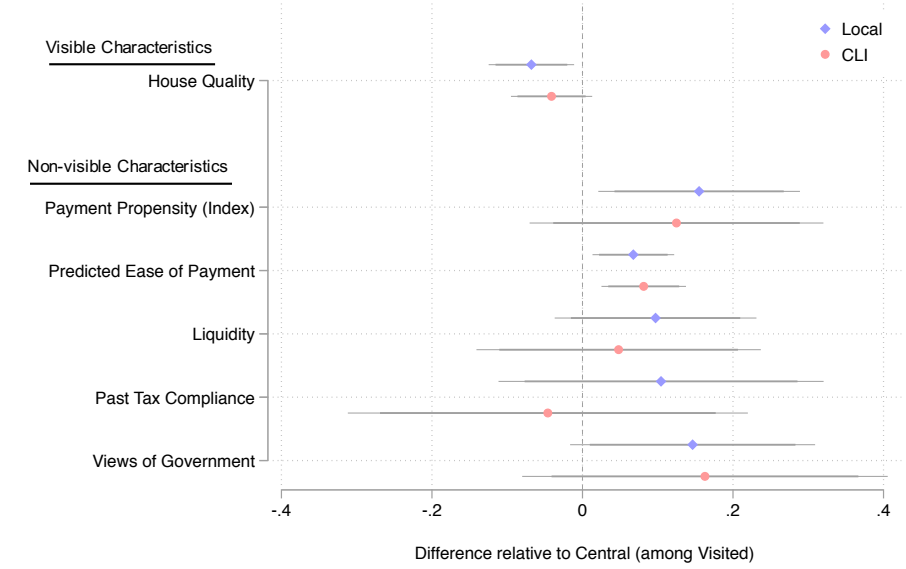
B: Predicted Ease of Payment and House Quality



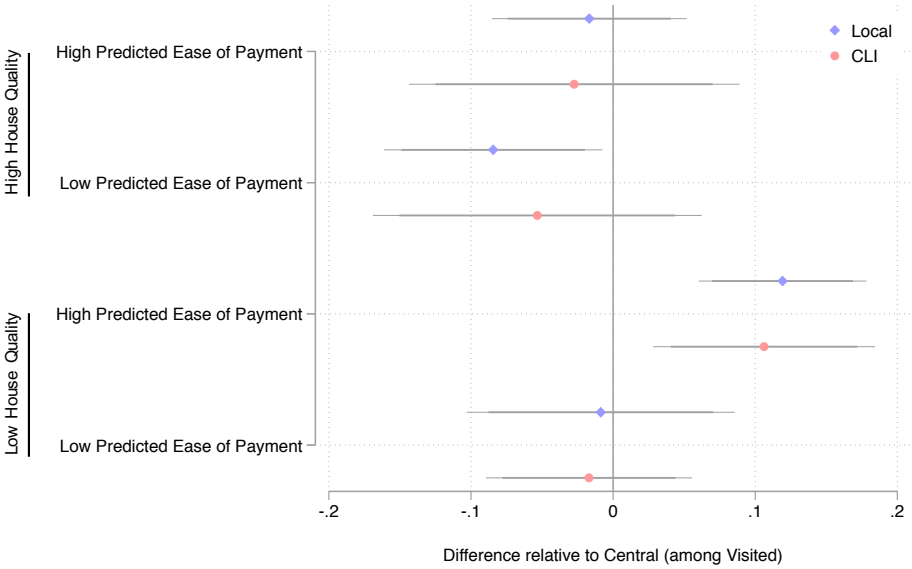
Notes: This figure reports correlations by treatment arm in the characteristics of properties visited by collectors after registration. It supplements the analysis in Figure 1, which examines *differences by treatment* in the characteristics of households that received tax visits after registration. Panel A shows correlations with visible and non-visible characteristics for indices described in Section VIA. Panel B shows correlations with tax visits in the four cells indicated (defined by interactions of high/low dummies for household house quality and predicted ease of payment). Correlations are estimated through separate regressions of characteristics on a treatment indicator among visited properties, controlling for the leave-one-out neighborhood mean of the outcome (Panel A) or the neighborhood mean of house quality and ease of payment (Panel B). We include time period, house type, and stratum fixed effects. We cluster standard errors at the neighborhood level. Households that paid at registration are dropped. This figure is discussed in Section VIA.

FIGURE A10: CHARACTERISTICS OF HOUSEHOLDS VISITED BY COLLECTORS AFTER REGISTRATION ACROSS TREATMENTS — NO HOUSE FIXED EFFECTS

A: Visible and Non-Visible Characteristics



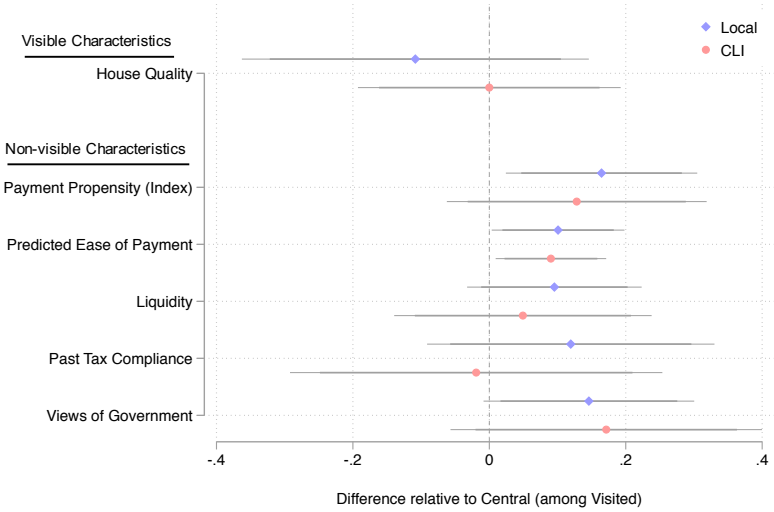
B: Willingness to Pay and House Quality



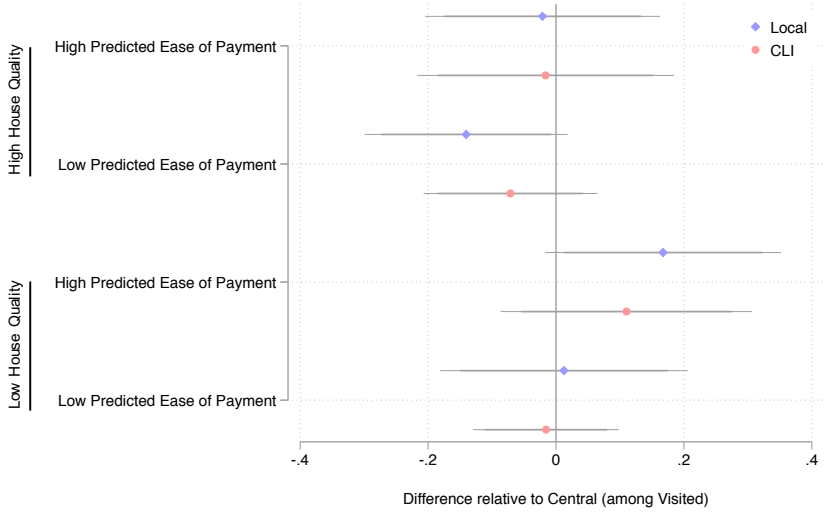
Notes: This figure reproduces the results from Figure 1 but excludes house fixed effects as a robustness check. Specifically, it reports differences by treatment arm in the characteristics of properties visited by collectors after registration, showing differences in characteristics of visited properties in the Local and CLI arms relative to the Central arm. Panel A shows differences in visible and non-visible characteristics for indices described in Section VIA. Panel B shows differences in the probability of receiving a visit in the four cells indicated (defined by interactions of high/low dummies for household house quality and predicted ease of payment). Differences are estimated through separate regressions of characteristics on a treatment indicator among visited properties, controlling for the leave-one-out neighborhood mean of the outcome (Panel A) or the neighborhood mean of house quality and ease of payment (Panel B). We include time period, house type, and stratum fixed effects. We cluster standard errors at the neighborhood level. Households that paid during registration are dropped. We discuss these results in Section VIA.

FIGURE A11: CHARACTERISTICS OF HOUSEHOLDS VISITED BY TAX COLLECTORS AFTER REGISTRATION ACROSS TREATMENTS — OMITTING NEIGHBORHOOD MEAN CONTROLS

A: Visible and Non-Visible Characteristics



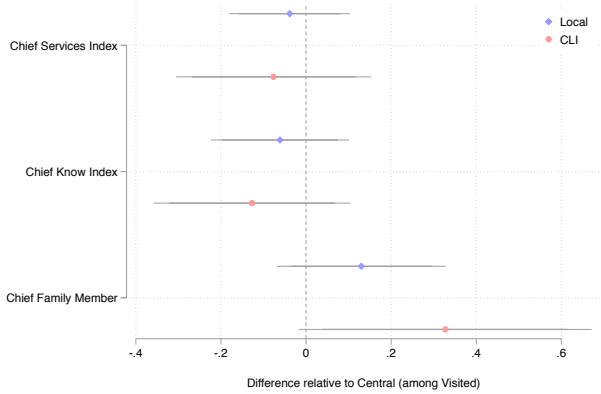
B: Willingness to Pay and House Quality



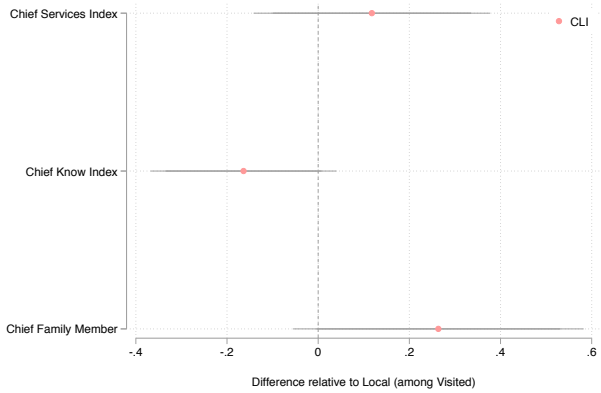
Notes: This figure reproduces the results from Figure 1 but omits the neighborhood mean controls as a robustness check. Specifically, it reports differences by treatment arm in the characteristics of properties visited by collectors after registration, showing differences in characteristics of visited properties in the Local and CLI arms relative to the Central arm. Panel A shows differences in visible and non-visible characteristics for indices described in Section VIA. Panel B shows differences in the probability of receiving a visit in the four cells indicated (defined by interactions of high/low dummies for household house quality and predicted ease of payment). Differences are estimated through separate regressions of characteristics on a treatment indicator among visited properties. We include time period, house type, and stratum fixed effects. We cluster standard errors at the neighborhood level. Households that paid during registration are dropped. We discuss these results in Section VIA.

FIGURE A12: CORRELATIONS BETWEEN TAX VISITS AND CHIEF CONNECTIONS

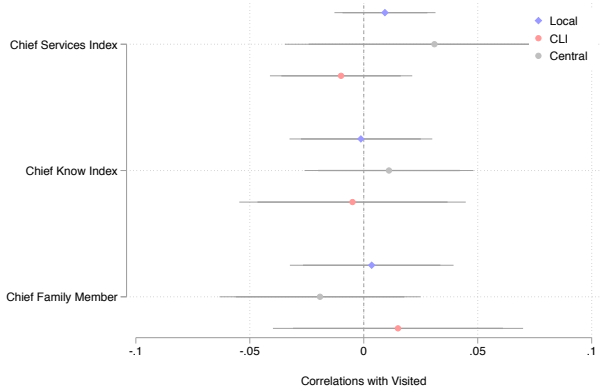
A: Local and CLI v. Central



B: CLI v. Local



C: Correlations by Treatment



Notes: This figure reports differences and correlations by treatment arm in the probability of receiving tax visits after registration and households' connections to the chief. Panel A shows differences in terms of the indices described in Section VIA, comparing Local and CLI to Central. Panel B shows differences comparing CLI to Local. Panel C shows correlations with tax visits by treatment. Differences are estimated through separate regressions of the connection variable on a treatment indicator, controlling for the leave-one-out neighborhood mean. Correlations are estimated through separate regressions of an indicator for receiving a tax visit on a characteristic separately by treatment groups. All regressions control for the leave-one-out neighborhood mean of the connection variable and include time period, house type, and stratum fixed effects and clustering standard errors at the neighborhood level. Households that paid at registration are dropped. We discuss these results in Section VIA.

TABLE A31: LOCAL V. CENTRAL: TAX VISITS AND COMPLIANCE BY COETHNICITY

	Visited Post-Registration			Compliance		
Match with Collector	Tribe	Subtribe	Majority Language	Tribe	Subtribe	Majority Language
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-0.002	0.063	-0.016	0.050***	0.026	0.049**
	(0.031)	(0.044)	(0.039)	(0.011)	(0.019)	(0.017)
Local X Match	0.007	-0.117**	0.020	-0.015	-0.035	-0.003
	(0.040)	(0.058)	(0.045)	(0.016)	(0.044)	(0.019)
Match	-0.010	0.143**	-0.004	0.011	0.051	-0.009
	(0.035)	(0.054)	(0.035)	(0.013)	(0.041)	(0.012)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13628	6457	13628	13752	6491	13752
Clusters	210	114	210	210	114	210
Central Mean (Non-Match)	.438	.297	.432	.072	.052	.074

Notes: This table reports estimates from a version of Equation 1, comparing tax visits and compliance in Local and Central (the excluded category) by whether the collector and property owner are coethnics along a specific dimension. The outcome in Columns 1–3 is whether households reported any tax visits after registration. The outcome in Columns 4–6 is compliance according to administrative data. Match corresponds to an indicator for the chief's or at least one state collector's ethnicity characteristic matching that of the property owner for the characteristics at the top of each column. Columns 1 and 4 show estimates including an interaction with an indicator for a collector's and property owner's tribe matching, Columns 2 and 5 for subtribe, and Columns 3 and 6 for both speaking Tshiluba (the majority language) as the mother tongue. All regressions include fixed effects for time periods described in Section IV, house type, and randomization strata. We cluster standard errors at the neighborhood level. These results are discussed in Section VIA.

TABLE A32: LOCAL V. CENTRAL: THE DISTRIBUTION OF THE TAX BURDEN — NO HOUSE FIXED EFFECTS

Outcome:	Compliance by Property Type		Complier Characteristics		
	Low Band Property	High Band Property	House Quality	Avg. Mon. Income	Liquidity Index
	(1)	(2)	(3)	(4)	(5)
Local	0.037***	0.002	-0.142**	-0.003	-0.079
	(0.008)	(0.013)	(0.056)	(0.038)	(0.161)
Time FE	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes
Observations	24105	3265	1296	224	224
Clusters	206	147	156	120	120
Central Mean	.064	.063	.099	.007	.118

Notes: This table re-estimates the results reported in Table 9 while excluding house fixed effects. Specifically, it reports estimates from a version of Equation 1, comparing property tax compliance in Local and Central (the excluded category). We include fixed effects for randomization strata and time periods, as described in Section IV, and we cluster standard errors at the neighborhood level. Columns 1 and 2 report estimates of the impact of local collection on compliance for low- and high-band households, respectively. Column 3 reports differences in an index of house quality conditional on the property paying the tax. Column 4 reports differences in monthly household income of properties, averaged across baseline and endline values, in Congolese Francs, conditional on paying the tax. Column 5 reports differences in an index of liquidity measures drawn from baseline (except income, which is included and uses information from endline) among payers. Columns 3–5 control for the leave-one-out neighborhood mean of the outcome. We discuss these results in Section VIB.

TABLE A33: LOCAL AND CLI V. CENTRAL: INCIDENCE BY COMPLIER CHARACTERISTICS — NO NEIGHBORHOOD MEAN CONTROLS

Outcome:	Complier Characteristics					
	Local v. Central			CLI v. Central		
	House Quality	Avg. Mon. Income	Liquidity Index	House Quality	Avg. Mon. Income	Liquidity Index
	(1)	(2)	(3)	(4)	(5)	(6)
Local	-0.216	0.002	-0.053			
	(0.156)	(0.041)	(0.174)			
CLI				0.133	0.015	0.183
				(0.127)	(0.053)	(0.211)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1310	228	228	832	141	141
Clusters	157	121	121	118	87	87
Mean	.099	.007	.118	.096	.02	.193

Notes: This table re-estimates the results reported in Columns 3–5 of Table 9 while excluding controls for the neighborhood mean. Columns 1–3 examine the distribution of the noted characteristics among taxpayers in a comparison of Local v. Central, while Columns 4–6 compare CLI v. Central. Column 1 and 4 report differences in an index of house quality conditional on the property paying the tax. Columns 2 and 5 report differences in monthly household income of properties, averaged across baseline and endline values, in Congolese Francs, conditional on paying the tax. Columns 3 and 6 report differences in an index of liquidity measures drawn from baseline (except income, which is included and uses information from endline) among payers. We include fixed effects for house type, randomization strata, and time periods, as described in Section IV, and we cluster standard errors at the neighborhood level. We discuss these results in Section VIB.

TABLE A34: LOCAL V. CENTRAL: THE DISTRIBUTION OF THE TAX BURDEN — PROPERTY VALUE BAND INTERACTIONS

	Compliance	Revenues
	(1)	(2)
Local	0.037***	77.607***
	(0.008)	(17.844)
Local X High-Value Band	-0.029**	-73.065
	(0.013)	(96.963)
High-Value Band	-0.014	398.778***
	(0.009)	(69.284)
Time FE	Yes	Yes
House FE	No	No
Stratum FE	Yes	Yes
Observations	27764	27764
Clusters	213	213
Central Mean (Low-Value Band)	.064	133.152

Notes: This figure reports estimates from a version of Equation 1, showing differences in tax payment in the Local arm relative to the Central arm by heterogeneity in property value band assessed at registration — an interaction-based specification of the analysis in Table 9, Columns 1–2. We include time period and stratum fixed effects and cluster standard errors at the neighborhood level. We discuss these results in Section VIB.

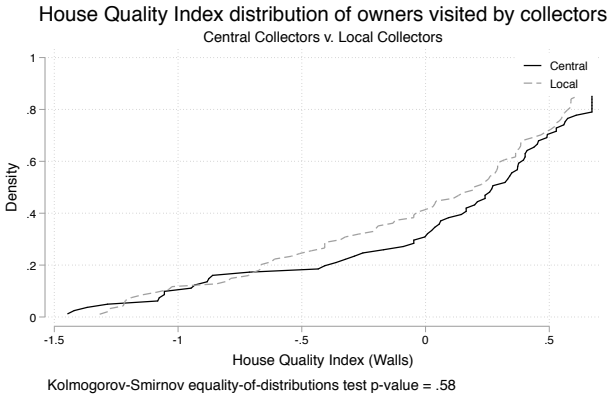
TABLE A35: LOCAL V. CENTRAL: THE DISTRIBUTION OF THE TAX BURDEN — COMPLIER CHARACTERISTICS INTERACTIONS

Heterogeneity Variable:	House Quality	Avg. Mon. Income	Liquidity Index
	(1)	(2)	(3)
Panel A: Compliance			
Local	0.053***	0.017	0.019
	(0.009)	(0.031)	(0.031)
Local X Above Median	-0.004	0.013	0.011
	(0.011)	(0.036)	(0.036)
Above Median	0.034***	0.022	0.023
	(0.006)	(0.025)	(0.025)
Central Mean (Below Median)	.045	.06	.06
Panel B: Revenues			
Local	121.651***	90.690	106.947
	(23.516)	(75.542)	(81.553)
Local X Above Median	-24.687	-38.534	-57.080
	(28.456)	(87.068)	(92.835)
Above Median	87.670***	92.759	102.466*
	(20.038)	(57.610)	(56.979)
Central Mean (Below Median)	96.541	115.385	115.385
Time FE	Yes	Yes	Yes
House FE	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes
Observations	17519	2236	2238
Clusters	174	212	212

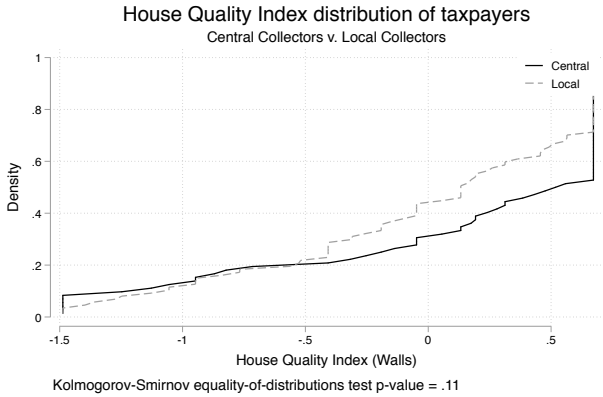
Notes: This figure reports estimates from a version of Equation 1, showing differences in tax payment in the Local arm relative to the Central arm by heterogeneity in the incidence measures described in Table 9, Columns 3–5: house quality, average monthly income, and an index of liquidity, interacting an indicator for the Local arm with indicators for having above-median values for each measure. We include time period, house type, and stratum fixed effects and cluster standard errors at the neighborhood level. We also include controls for the the leave-one-out neighborhood means of the relevant heterogeneity measure. We discuss the interpretation of these results in Section VIB.

FIGURE A13: HOUSE QUALITY, INCOME, AND LIQUIDITY DISTRIBUTIONS AMONG VISITED AND PAYING HOUSEHOLDS BY TREATMENT

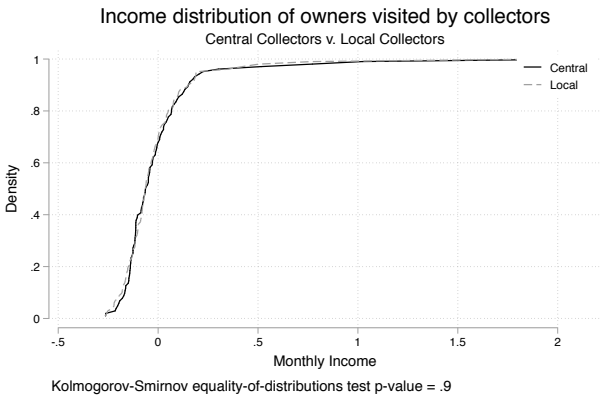
A: House quality - Visited owners



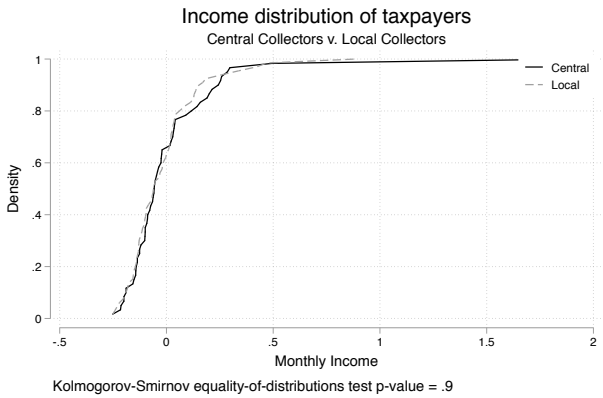
B: House quality - Taxpayers



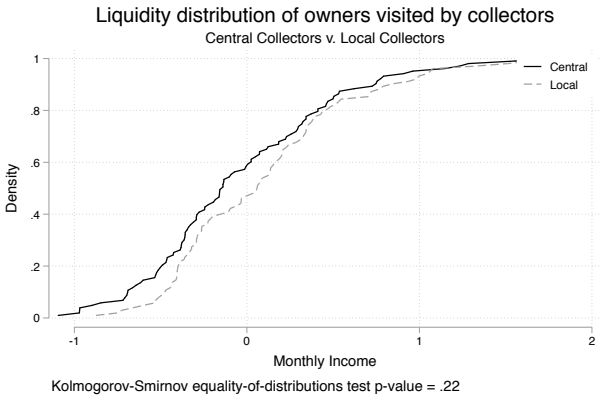
C: Income - Visited owners



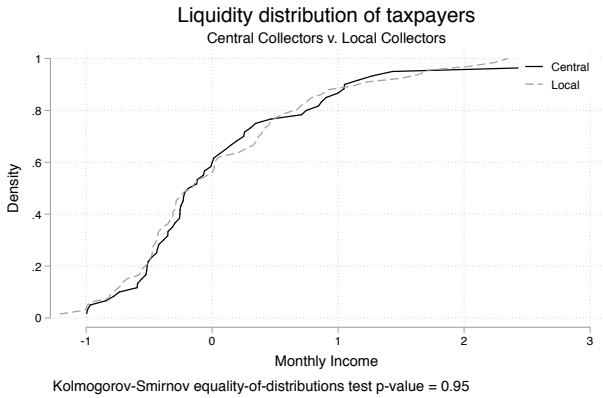
D: Income - Taxpayers



E: Liquidity - Visited owners



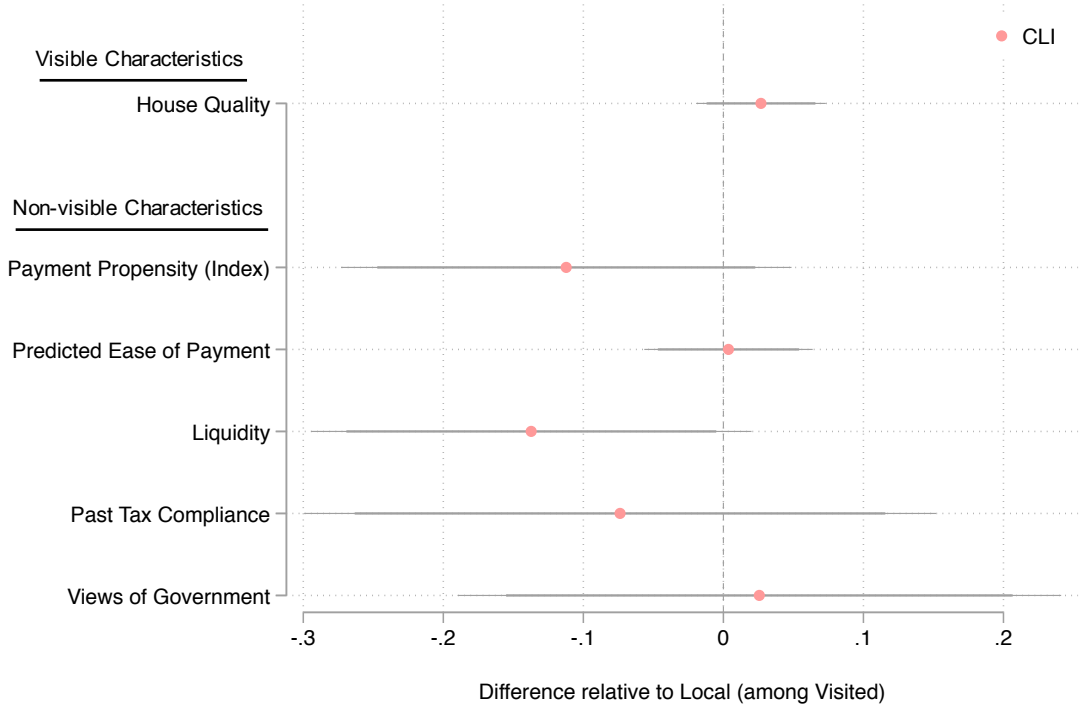
F: Liquidity - Taxpayers



Notes: This figure shows cumulative distribution functions of house quality and income by treatment and separately among households that received tax visits after registration (Panels A, C, and E) and that paid the tax (Panels B, D, and F). In Panel B, the taxpayer distribution has considerable mass at the maximum value of the house quality index in Central, making the CDF somewhat difficult to read. Kolmogorov-Smirnov equality of distributions test p-values are reported at the bottom. We discuss these results in Section VIB.

A1.7 Additional Exhibits for Paper Section 9 — Policy Implications

FIGURE A14: LOCAL V. CENTRAL + LOCAL INFO: DIFFERENCES IN TARGETING OF TAX VISITS BY HOUSEHOLD CHARACTERISTICS



Notes: This figure reports correlations by treatment arm in the characteristics of properties visited by collectors after registration. The figure shows differences in visible and non-visible characteristics for indices described in Section VIA. Correlations are estimated through separate regressions of an indicator for receiving a tax visit on a characteristic separately by treatment groups, controlling for the leave-one-out neighborhood mean of the outcome, including time period, house type, and stratum fixed effects and clustering standard errors at the neighborhood level. Households that paid at registration are dropped. We discuss these results in Section A3.6.

TABLE A36: LOCAL V. CENTRAL: IMPACTS ON HOUSEHOLD WELL-BEING

	Monthly Income	Weekly Transport	Bed hungry last month	Bed hungry last month num. days	Lacks 3000 CF cash today	Lacks 3000 CF cash this month	Lacks 3000 CF cash this month num. days
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Reduced Form							
Local	-2300.525	-37.852	-0.015	-0.017	-0.014	-0.003	0.105
	(7800.918)	(438.961)	(0.023)	(0.077)	(0.023)	(0.027)	(0.176)
Observations	2277	2329	2330	2330	2330	2330	2330
Mean	144789	4456	.516	.993	.675	.652	1.29
Panel B: Instrumenting for Taxes Paid							
$\widehat{\text{Taxes Paid}}$	-1.34e+05	-2574.310	-1.054	-1.181	-0.942	-0.180	7.147
	(4.86e+05)	(30047.563)	(1.954)	(5.270)	(1.802)	(1.827)	(12.946)
Observations	2277	2329	2330	2330	2330	2330	2330
Mean	144789	4456	.516	.993	.675	.652	1.29
Panel C: Instrumenting for Taxes or Bribe Paid							
$\widehat{\text{Taxes or Bribe}}$	33221.221	-1.49e+04	-0.366	0.770	-0.079	-0.115	3.098
	(1.90e+05)	(19209.285)	(0.603)	(1.615)	(0.529)	(0.634)	(3.169)
Observations	1260	1287	1287	1287	1287	1287	1287
Mean	150899	5174	.482	.863	.67	.63	1.1
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates from a version of Equation 1, using endline measures of well-being in Local and Central (the excluded category) as the dependent variable. We include fixed effects for house type and randomization strata and cluster standard errors at the neighborhood level. Columns 1 and 2 report differences in monthly household income and weekly transport (a measure of spending). Columns 3 and 4 report differences in whether the household went to bed hungry at least one day in the last month and how many days, respectively. Columns 5, 6, and 7 report differences in whether the household lacked 3,000 Congolese Francs to be able to make a payment at the date of survey, sometime in the last month, and how many times in the last month, respectively. Panel A reports the reduced form results of a regression of outcomes on an indicator for the Local treatment. Panel B regresses outcomes on an indicator for tax payment instrumented by an indicator for the Local treatment. Panel C regresses outcomes on an indicator for paying a tax or bribe with an indicator for the Local treatment. We discuss these results in Section VII.

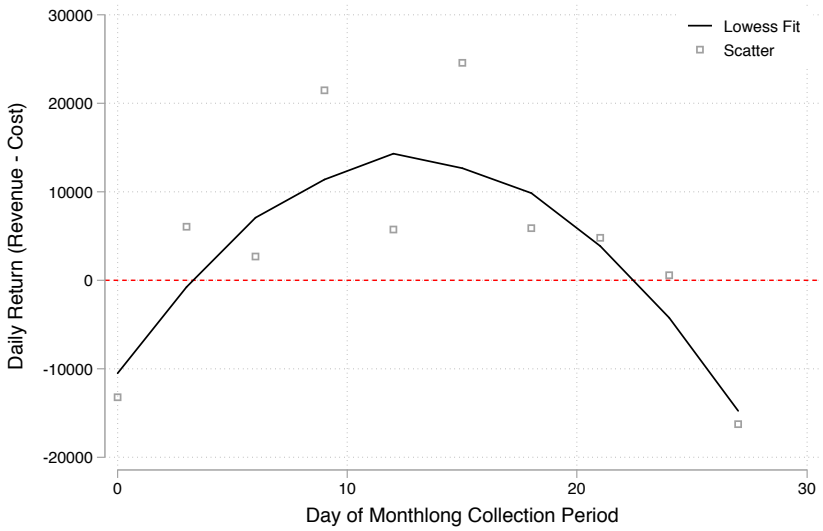
TABLE A37: LOCAL V. CENTRAL: VIEWS OF GOVERNMENT AND CHIEFS BY TAX AND BRIBE PAYMENT

	Provincial Government				Neighborhood Chief			
	Views of govt. (index) (1)	Trust in govt (2)	Responsiveness of govt. (3)	Performance of govt. (4)	Views of chief (index) (5)	Trust in chief (6)	Responsiveness of chief. (7)	Performance of chief. (8)
Panel A: Interaction with Paid Tax								
Local	0.036	0.153**	-0.057	-0.036	0.070	0.057	-0.039	0.085
	(0.052)	(0.060)	(0.046)	(0.055)	(0.053)	(0.056)	(0.059)	(0.063)
Local X Paid Tax	-0.090	-0.288*	0.148	-0.184	-0.155	-0.143	-0.326**	0.057
	(0.118)	(0.151)	(0.137)	(0.138)	(0.132)	(0.136)	(0.150)	(0.120)
Paid Tax	0.082	0.065	-0.101	0.173	0.116	0.028	0.261**	-0.123
	(0.089)	(0.109)	(0.108)	(0.107)	(0.095)	(0.100)	(0.115)	(0.085)
Observations	2329	2207	2205	2102	2303	2291	1637	1302
Central Mean (No Pay)	-.009	.004	-.009	.009	-.01	-.016	.029	-.013
Panel B: Interaction with Paid Bribe (Endline)								
Local	0.082	0.227**	-0.010	-0.121*	0.113*	0.137*	-0.067	0.108
	(0.065)	(0.088)	(0.064)	(0.064)	(0.067)	(0.079)	(0.080)	(0.087)
Local X Paid Bribe	0.321	-0.531	0.842*	0.287	0.154	-0.428	-0.246	0.805
	(0.461)	(0.405)	(0.487)	(0.497)	(0.490)	(0.506)	(0.473)	(0.539)
Paid Bribe	-0.466	0.522*	-0.500	-0.689*	-0.236	0.112	0.235	-0.097
	(0.391)	(0.308)	(0.375)	(0.411)	(0.390)	(0.413)	(0.282)	(0.179)
Observations	1124	1073	1063	1021	1121	1114	789	645
Central Mean (No Pay)	-.081	-.052	-.06	-.047	-.062	-.075	-.021	.01
Baseline Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

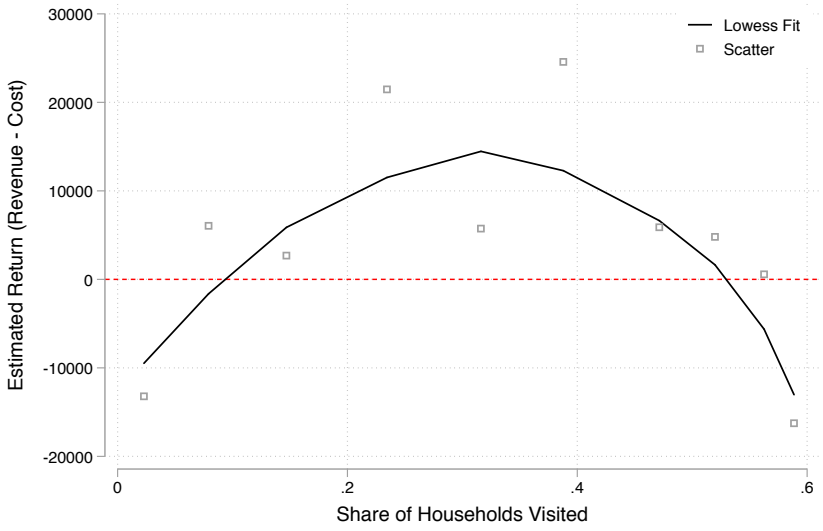
Notes: This table shows estimates from versions of Equation 1, comparing the Local arm to the Central arm (the excluded category). The outcomes are views of chiefs and government as defined in Table 5. Panel A shows estimates by interactions with and indicator for paying the tax according to the administrative data. Panel B shows estimates by interactions with an indicator for paying a bribe to the collector at endline (self-reported). All regressions include fixed effects for house type and randomization strata and cluster standard errors at the neighborhood level. We discuss these results in Section VII.

FIGURE A15: RETURN TO ADDITIONAL VISITS IN CENTRAL

A: Return to Additional Days Worked



B: Return to Additional Visits



Notes: This figure shows the estimated daily return to tax collection in Central (*i*) over the course of the month in which collectors were assigned to a given neighborhood, and (*ii*) as a function of the share of the total households in the neighborhood that were visited. The revenue data come from the handheld receipt printers and the timestamp recorded for each transaction. The cost data come from tax campaign records concerning transportation costs incurred by collectors. We discuss these results in Section VB.

A2 Additional Details on the Tax Campaign and its Evaluation

A2.1 Block-Randomized Design

In the randomization of the main tax collector treatments, we used a block-randomized design, stratifying on three variables.

> **1. Geographic strata.** We use 12 geographic strata corresponding to different city regions (Figure A3). Two encompass the city center; the rest correspond to what the tax ministry calls "the periphery." Blocking on these strata ensures balance on a number of geographic characteristics, including (*i*) the local legitimacy of the chief (higher in the periphery), and (*ii*) the intensity of prior tax enforcement (lower in the periphery).
>
> **2. Treatment status in the 2016 tax campaign.** We also block on treatment status in the 2016 property tax campaign, randomly assigned on the neighborhood level (Weigel, 2020). Treated neighborhoods received visits from tax ministry agents (similar to the Central arm), while control neighborhoods did not (similar to the pure control arm). Stratifying on this variable ensures balance on past door-to-door property tax collection.
>
> **3. Past experience of chiefs in tax collection.** Finally, we block on a measure of whether chiefs reported ever having been involved in tax collection in the past, which was the case for 22% of chiefs.[118] Incorporating this variable into our strata improves balance on this important chief-level characteristic.

We first created strata using the first two variables. Then, for each, we created two substrata based on the third variable.[119]

A2.2 Tax Letter Message Treatments

As shown in Figure A1, the tax letters distributed by collectors during registration in all treatment arms contained cross-randomized messages, as in Blumenthal et al. (2001) and Pomeranz (2015). Collectors read the entire flier out loud to property owners during registration. The tax letters provided basic information about the tax campaign, including the compound number, the compound-specific tax rate for the year, to whom the tax should be paid (state or chief tax collector, or either). In addition, the tax letters contained one of the following messages, randomized at the household level:[120]

[118] Of those who responded affirmatively, 79% reported collecting the property tax, 10% the rental tax (a property tax levied on renters), and the remainder reported having collected other taxes.

[119] We split at the median level of experience with tax collection if a neighborhood had more than one chief, possible in larger neighborhoods with multiple main avenues.

[120] Additionally, some contained an image of a legal receipt along with a phrase noting that the payer should receive a printed receipt. On other letters, there was no copy of the receipt, nor mention of the printed receipts. This treatment, intended for a separate paper on bribe payment, aimed to enable citizens to hold tax collectors accountable to following the protocol of the campaign.

Central deterrence. This message says that refusal to pay the property tax entails the possibility of audit and investigation by the provincial tax ministry.

Local deterrence. The Local version of the deterrence message says that refusal to pay the property tax entails the possibility of audit and investigation by the neighborhood chief (*chef de quartier*). This is the highest-ranking city chief, from whom other city chiefs seek counsel when needing to solve problems.

Central public goods. This message says that the provincial government will be able to improve infrastructure in the city of Kananga only if citizens pay the property tax.

Local public goods. The Local version of this message is exactly the same, except that it mentions each citizen's locality instead of Kananga.[121]

Trust. The trust message reminds citizens that paying the property tax is a way of showing that they trust the state and its agents.

Control. Control letters say "It is important to pay the property tax."

These messages were designed to interact with the main collector treatments to help isolate mechanisms. As noted in our pre-analysis plan, the "Central" ("Local") versions of these messages should have been more credible coming from, and thus complemented the efficacy of, state (chief) collectors. If chiefs collected more taxes because of greater local sanctioning capacity, there should be a more pronounced treatment effect when tax letters contained the Local Deterrence message (rather than Control). Similarly, if chiefs activated tax morale due to their link with local services, we should observe a larger treatment effect when their tax notices contained the Local Public Goods message. Finally, if greater trust in chiefs explains their effectiveness as collectors, then the Trust message should be more potent in Local than Central.

Randomized messages were introduced in the last phase of the campaign. Previously, collectors distributed tax letters identical to those in Figure A1 but without randomized messages. The analysis of the fliers discussed in Section VB thus restricts the sample to the 5,434 properties subject to randomized messages on tax letters. Although this smaller sample reduces our power, an ex-post back-of-the-envelope calculation suggests that we are still powered to reject a flier message main effect of 1.4 percentage points and an interaction effect of about 3 percentage points. We think these are plausible minimum effect sizes given similar studies like Scartascini and Castro (2007), which finds that enforcement messages increased extensive-margin property tax compliance in Argentina by 5 percentage points.

A2.3 Chief Jurisdiction Mapping

The provincial government did not have a precise map of chiefs' jurisdictions. Thus, before the tax campaign, our research team helped the government map out these jurisdictions. In cases where there were multiple chiefs within the same neighborhood, i.e. in charge of two different avenues, chiefs whom the government could choose for the tax campaign were ranked by (*i*) estimating the spatial extent of each chief's domain by calculating a 20-meter

[121] Localities are the smallest administrative unit in Kananga.

buffer around each avenue they were in charge of, and (*ii*) overlaying these domains with geocoded population data. The resulting ranking of chiefs therefore corresponds to chiefs with the largest population-weighted jurisdictions in the neighborhood. In certain cases, top-ranked chiefs were unable to collect due to disability, travel during the campaign, or other reasons. In these cases, the second highest rank chief typically completed the work.

A2.4 Logistics Pilot

A logistics pilot, conducted in March-April 2018, had two main goals. First, it tested a new handheld receipt printer and validated that city chiefs — who are often older and less skilled with technology — would be able to work with the receipt printers. Second, it tested the tax letters and other procedures of the campaign to be sure they could be easily understood by citizens. The pilot was conducted in eight neighborhoods of Kamilabi, a region in northwest Kananga that is isolated from the rest of Kananga by a series of steep ravines. This area was selected strategically due to its remote location to minimize potential informational spillovers. We exclude the pilot neighborhoods from our main estimations. But in Table A7, we show that the main results are robust to including these pilot neighborhoods.

A2.5 Time Imbalance

This section discusses in detail the time imbalance arising from the fact that not all treatments occurred simultaneously (cf. Section IV). As noted, there was a secular decline in compliance over the course of the study (Figure A5). This decline presents a problem for our analysis because the treatments were rolled out in a staggered fashion over time due to logistical constraints at the tax ministry. Although the staggered implementation greatly helps reduce the degree of imbalance by time, there remains imperfect time overlap of treatments. In short, some treatments were implemented in periods with higher compliance, which introduces artificial differences in tax outcomes when comparing treatments. For example, the Central treatment started first and therefore is the only treatment to include observations at a point in time when compliance was highest. Had the Local treatment started at the same time it would have likely registered even higher levels of compliance, according to the trends extrapolated from data collected during the rest of the campaign. Therefore, pooling all data across time would artificially inflate estimates of compliance in treatment arms with (randomly) more observations earlier in time relative to treatment arms with more observations later in time.

Importantly, the decline in compliance over time does not reflect collectors choosing to work in "easy" neighborhoods first because the timing in which they received collectors was random. That is, within the tax ministry's overall schedule alternating between collection treatments, which neighborhoods appeared in different monthlong waves of the campaign was randomly assigned. This decline is also likely unrelated to collector characteristics, as evidenced by the fact that it impacts all treatment arms in a similar fashion. Instead, we suspect the downward trend in compliance reflects growing dissatisfaction with the government as the December 2018 election approached. The unpopular President Joseph Kabila had managed to avoid facing election two years running, and in 2018 protests were erupting across the country, to which the government responded with repression and violence. In survey outcomes we collected, there is a similar decline in attitudes toward the government and tax morale during this time period.

A2.5.1 Preferred Specification and Robustness Tests

In our preferred specification we include fixed effects for two month periods of the tax campaign. This ensures that we are comparing treatments within similar time periods with sufficient overlap in treatment observations to permit valid comparisons within time periods. Because it maximizes time balance on both ends, our preferred fixed effects begin on the midpoint between the first days of the two treatments being compared, and end on the midpoint between their last days. However, strictly speaking, when a two-month period starts (and ends) is arbitrary for the purposes of including time fixed effects, so as a first robustness check, we also run and report our main estimations using fixed effects defined at every possible start date (Figure A6).

We then implement robustness tests using five other approaches: (*i*) including two-month fixed effects defined by shifting the start date of the two-month fixed effect period definition in our preferred specification backwards and forwards 15 days and selecting the shifted version that yields the median estimate among all shifts, (*ii*) using the interaction-weighted estimator from Gibbons et al. (2018), (*iii*) including one-month fixed effects, (*iv*) trimming observations on either end of the campaign if comparison treatments were not also active, (*v*) estimating effects with coarsened exact matching on time to identify clusters of comparable observations within relevant treatment arms (Iacus et al., 2012). Below we describe these alternative tests:

1. Median of Possible Permutations of Two-Month Fixed Effects. Because the start and end points of the two-month fixed effects are arbitrary, we shift these cutoffs 30 times — 15 days backwards and forwards in time — and then redefine two-month intervals and re-estimate Equation 1 for each shifted fixed effect version. Figure A6 shows the results of this exercise and displays the median estimate, which we report as a robustness test for our preferred fixed effect definition. This approach (and the preferred approach) addresses trends over the campaign but not trends within two-month periods.

2. Interaction-Weighted Estimator. This estimator takes the weighted average of estimates from interaction terms of treatment with two-month dummies (defined by the preferred version of two-month fixed effects), weighting by the number of observations in each group. This approach addresses inconsistency in the presence of group-specific heterogeneous treatment effects (Gibbons et al., 2018).

3. One-Month Fixed Effects. One-month rather than two-month fixed effects allow for finer comparisons across time. This approach may better address trends over the campaign though not trends within one-month periods. However, due to staggering, it will also mean many observations do not contribute towards the estimated effect at all because, for a given treatment comparison, there is no overlap with other treatments in time (Figure A5).

4. Trimming Observations. Dropping observations collected before the start (after the end) of other treatment arms, so that only observations collected between the same start and end dates are considered. This addresses problems of overlap at the start and end of the campaign but does not address those in between campaign stages.

5. Coarsened Exact Matching. This approach from Iacus et al. (2012) involves matching on a continuous variable with imperfect overlap across treatments — this matching variable is "coarsened" and then used to match observations across treatments. Such matching offers potentially the best method for dealing with the time confound, as it keeps only the observations closest in time in the treatment groups being compared; however, due to the nature of staggering in the campaign it results in much smaller estimation samples given near matches cannot be found for all observations. This is especially true when comparing the Central and Central + Local Information treatment arms.

Ultimately, we prefer the approach using two-month fixed effects as it addresses the key time imbalance by comparing observations collected relatively close in time, while allowing us to retain the majority of the sample. We use the version defined using the midpoints between the start and end dates of the treatments being compared as it maximizes time balance on both ends and our tests confirm that this choice is robust to shifting the start and end dates. Tables A5 and A18 display the results of these robustness tests for the main estimations. The estimates are remarkably similar across specifications, which we interpret as validation for our preferred approach.

A2.6 Detailed Survey-based Variable Descriptions

This section provides the exact text of the questions used to construct the survey-based variables examined in the paper.[122]

1. *House Quality.* This standardized variable is increasing in the quality of the house of the respondent, as indicated by the quality of its walls. The exact survey prompt to enumerators is as follows:

    - 'Observe the principal material of the walls of the main house.' [Sticks/ Palms, Mud brick - bad condition, Mud brick - good condition, Bricks, Cement]

2. *Average Monthly Income.* This variable is the self-reported (log) income of the respondent averaged over the baseline and endline surveys. It was recorded in both the baseline and the endline surveys in response to the question: 'What was the household's total earnings this past month?'

3. *Education.* This variable measures the years of education of the respondent, standardized to facilitate interpretation of magnitudes. The exact survey questions are as follows:

    - 'What is the highest level of school you have reached? [Never been to school, Kindergarten, Primary, Secondary, University]
    - 'What is the last class reached in that level?' [1, 2, 3, 4, 5, 6, >6]

[122] The main variables, such as payments and views, chief characteristics, and household characteristics, are discussed in the paper in Sections IVB, VC, and VI, respectively.

4. *Erosion Threat.* This standardized variable is increasing in the enumerator's perception that the respondent's property is threatened by ravines, which are caused by erosion and are widespread in Kananga. Properties that lie close to ravines are considerably less valuable. This variable was recorded in the baseline survey in response to the enumerator prompt: 'Is this compound threatened by a ravine?' [Yes - gravely threatened, Yes - somewhat threatened, No]

5. *Has Electricity.* This variable equals 1 if the household reports in the baseline survey that they have access to electricity. The exact question text is: 'Do you have any source of electricity at your home?'

6. *Chief Family Member.* This variable equals 1 if the local chief is a member of the family of the respondent. The exact survey question from baseline is: 'Is the avenue chief a member of your family?'

7. *Chief Know Index.* This is a standardized index increasing in respondents' knowledge of and ties with the local chief. It is composed of the following baseline survey questions:

   - 'Do you know the name of your avenue chief? If yes, what is it?'
   - 'Do you have the phone number of your avenue chief?'
   - 'Do you attend the same church as your avenue chief?'

8. *Chief Services Index.* This is a standardized index increasing in the services and help the respondent has received from the local chief in the past. The exact baseline survey question is as follows: 'In the past six months how many times did your avenue chief helped you deal with any of the following issues?'

   (a) 'Help finding a solution to a problem'

   (b) 'Helped a member of your family get a job'.

9. *Connected to Chief.* This is a standardized index increasing in how close the respondent reports being to the local chief. It is a combination of the Chief Family Member variable and the Chief Know Index.

10. *Trust in Organizations.* This standardized index is increasing in the level of trust the respondent reports having in various organizations. The exact survey question is as follows:

    - Prompt: 'I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence, or none at all?'
    - Organizations:
      (a) 'NGOs'
      (b) 'Local leaders'
      (c) 'The national government (in Kinshasa)'

(d) 'The provincial government'

(e) 'The tax ministry'

(f) 'Foreign research organizations'.

The values were reversed to code this variable.

11. *Liquidity Index.* This is a standardized index increasing in the estimated liquidity of the household. It includes multiple questions about the income, employment, consumption, and possessions of the respondent reported in the baseline survey as well as cash-on-hand reported in the endline survey. As above, the measure of income used is the average of baseline and endline values. The exact survey questions about employment and earnings are as follows:

    - 'Are you the owner of this compound, or do you rent?'
    - 'What type of work do you do now?' [Unemployed-no work, Medical assistant, Lawyer, Cart pusher, Handyman, Driver (car and taxi moto), Tailor, Diamond digger, Farmer, Teacher, Gardner, Mason, Mechanic, Carpenter, Muyanda [bicycle pusher], Military officer/soldier or police officer, Fisherman, Government personnel, Pastor, Porter, Professor, Guard, Work for NGO, Seller (in market), Seller (in a store), Seller (at home), Student, SNCC]
    - 'What was the household's total earnings this past month?'

    The exact survey questions about the household's consumption are as follows:

    - 'How much money have you spent on transport in the past seven days'
    - 'Do you have any source of electricity at home?'

    The exact survey question about the household's possessions is as follows:

    - 'In your household, which (if any) of the following do you own: motorbike, car or truck, radio, television, electric generator, sewing machine, none.'

    The exact survey questions about cash-on-hand from the endline survey are as follows:

    - 'In the past 30 days, has your household had to go to bed hungry because you haven't had enough money on hand?'
    - 'On what dates did you find yourself short of cash for these expenditures?' [1-30, All parts of month were the same]
    - 'Imagine that today you learn that you need to pay an additional 3000 FC for a school fee in order for your child to continue in school. Could you find this money in the next 4 days? '
    - 'In the past 30 days, were there days in which you could not have paid this fee? Which days could you NOT have paid this fee?' [1-30, I could never pay this fee any day]

12. *View of Government Index.* This is a standardized index increasing in the respondent's evaluation of and trust in the government. The underlying survey questions are as follows:

    - 'I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all?'
    - Organizations (values reversed to code these variables):
        (a) 'The national government (in Kinshasa)'
        (b) 'The provincial government'
        (c) 'The tax ministry'
    - 'How would you rate the performance of the provincial government in Kananga?' [Excellent, Very good, Good, Fair, Poor, Very poor, Terrible] The values were reversed to code this variable.
    - 'Now I would like to ask you what you think the provincial government will do with the money it receives from this 2016 property tax campaign. Imagine that the provincial government of Kasaï Central receives $1000 thanks to this campaign. How much of this money will be put to good use, for example providing public goods?'

13. *Past Tax Compliance.* This variable equals 1 if the household reports in the baseline survey that they have paid property tax in the past. The exact question text is: 'Have you ever paid the property tax?'

14. *Payment Propensity Index.* This index is a combination of the Liquidity Index, the View of Government Index, and Past Tax Compliance.

15. *Ease of Payment.* This variable is derived from chief consultations in the CLI arm and equals 1 if the chief believes that the household can very easily afford the payment of the property tax. The exact survey question is as follows: 'Does the household head have the financial means to pay the tax?' [Hardly, Easily, Very easily]

16. *Predicted Ease of Payment.* This is a predicted value of the household's ease of payment using household characteristics, as described in Section VIA. It comprises data collected in the midline survey about the age of the respondent, his sex, his tribe and his employment status from the baseline survey. The exact midline survey questions are as follows:

    - 'Is the owner a man or a woman?'
    - 'How old is the owner?'
    - 'What is his tribe?'

    The exact baseline employment survey questions for employed and, separately, salaried are described in the Liquidity Index entry above. This predicted variable also

takes into account whether or not the respondent works for the government. The exact question is as follows: 'Do you work for the government in any capacity? If yes, please describe the job.' [Teacher, Military/ Police, Construction/ Maintenance of infrastructure, Lawyer, Doctor/ Nurse/ Lab Tech, Secretary, Driver, Functionary, Local chief (avenue, quartier), SNCC (state railroad), Political appointee]

17. *Salongo Contributions.* This is a variable reporting the household's contributions to *salongo*. The exact midline and endline survey questions are as follows:

    - 'Did someone from your household participate in *salongo* in the past 30 days?' (Extensive margin)
    - 'For how many hours in total did they participate in *salongo*? Please add together the time contributed by each member of your household in the past 30 days.' (Intensive margin)

18. *Trust in Government.* This is a variable increasing in the respondent's level of trust in both the provincial and national government. This variable is coded as an average of the answers to the question from the standardized index 'Trust in Organizations' about the national and provincial government.

19. *Responsiveness of Government.* This is a variable reporting the respondent's perception of how responsive the provincial government is. The exact survey question was asked in both the baseline and the endline survey as follows: 'To what degree does the provincial government respond to the needs of your avenue's inhabitants?' [Very responsive, Responsive, A little bit responsive, Not responsive] Values reversed to code this variable.

20. *Performance of Government.* This is a variable reporting the respondent's perception of the overall performance of the provincial government. The exact survey question was asked in both the baseline and the endline survey as follows: 'How would you rate the performance of the provincial government in Kananga?' [Excellent, Very good, Good, Fair, Poor, Very poor, Terrible] Values reversed to code this variable.

21. *Integrity of Government.* This is a variable reporting the respondent's perception of the integrity of the government, i.e. the opposite of corruption. The exact endline survey question is as follows: 'Now I would like to ask you what you think the provincial government will do with the money it receives from this 2018 property tax campaign. Imagine that the provincial government of Kasaï Central receives $1000 thanks to this campaign. How much of that money do you think was misappropriated/wasted?' The integer provided by the respondent was subtracted from 1000 to code the variable.

22. *View of Government (index).* This index is a combination of the following variables: Trust in Government, Responsiveness of Government, Performance of Government, and Integrity of Government.

23. *Trust in Chief.* This is a variable increasing in the respondent's level of trust in the chief. The variable uses the answer to the question from the standardized index 'Trust in Organizations' about the chief.

24. *Responsiveness of Chief.* This is a variable reporting the respondent's perception of how responsive the chief is to the needs of people in the neighborhood. The exact survey question, asked in both the baseline and the endline survey, is as follows: 'To what degree does the chief respond to the needs of your avenue /locality's inhabitants?' [Very responsive, Responsive, A little bit responsive, Not responsive] Values reversed to code this variable.

25. *Performance of Chief.* This is a variable reporting the respondent's perception of the overall performance of the chief. The exact survey question was asked in both the baseline and the endline survey as follows: 'Overall, how would you rate the performance of the chief?' [Excellent, Very good, Good, Fair, Poor, Very poor, Terrible] Values reversed to code this variable.

26. *Integrity of Chief.* This is a variable reporting the respondent's perception of the integrity of the chief. The exact endline survey question is as follows: 'Imagine that the chief is in charge of doing a public project on your avenue. He receives $1000. How much of this money will they put in their pockets?' The integer provided by the respondent was subtracted from 1000 to code the variable.

27. *View of Chief (index).* This index is a combination of variables: Trust in Chief, Responsiveness of Chief, Performance of Chief, and Integrity of Chief.

28. *Perceived Tax Compliance on Avenue.* This is a variable reporting the respondent's perception of what share of their neighbors have paid their property tax in 2018. The exact survey question was asked in the endline survey as follows: 'In your opinion, out of 10 compounds on your avenue, how many actually paid the property tax in 2018?'

29. *Trust in Tax Ministry.* This is a variable increasing in the respondent's level of trust in the tax ministry. The variable uses the answer to the question from the standardized index 'Trust in Organizations' about the tax ministry.

30. *Property Tax Morale.* This is a variable reporting the respondent's perception of how acceptable it is not to pay one's property tax. The exact survey question asked in both baseline and endline surveys is as follows: 'Now, imagine that next week a tax collector from the government comes and visits one of your neighbors. Imagine he absolutely refuses to pay the property tax. In your opinion, how acceptable is this?' [It's acceptable, It could be acceptable under some circumstances, It is not acceptable]

31. *Fairness of Property Taxation.* This is an index increasing in the respondent's evaluation of how fair property taxation is. The underlying endline survey questions are as follows:

    - 'In your opinion, how fair is it that households in your neighborhood must pay the property tax?'
    - 'In your opinion, how fair was the amount demanded for the property tax in 2018?'

- 'In your opinion, how fair were the collectors who worked on the property tax campaign of 2018?' [Very fair, Fair, Unfair, Very unfair] Values reversed to code this variable.

32. *Perception of Enforcement.* This is a variable reporting the respondent's perception of how likely it is that one gets sanctioned for not paying property tax. The underlying baseline and endline survey question is as follows: 'Imagine that next week a tax collector comes and visits one of your neighbors. Imagine he absolutely refuses to pay the property tax. In this case, what is the probability that the government will pursue and enforce sanctions? Choose one of the following options: He will definitely be pursued and punished; He is very likely to be pursued and punished; He is unlikely to be pursued and punished; He is very unlikely to be pursued and punished.' Values were reversed to code this variable.

33. *Paid Bribe.* This is a variable providing the respondent's self-reported bribe payments. It uses a local code for bribes, which is paying the "transport" of the tax collector (or any government agent). The exact midline and endline survey questions are as follows:

    - 'Did you (or a family member) pay the transport of the collector?'
    - 'Apart from the amount that you paid, did the collector ask you for another small sum on the side (for example, for his transport)?'

34. *Other Payments.* This is a variable providing the respondent's self-reported informal payments to officials. The exact midline and endline survey question is as follows: 'Now, I'd like to talk about small payments made to officials such as small amounts paid for transport, water, tea, etc. In the past 6 months, did you make any such payment?'

35. *Vehicle Tax.* This variable equals 1 if the household reports that they have paid the vehicle tax in 2018. The exact question text is: 'Let's discuss the vehicle tax. Did you pay this tax in 2018?'

36. *Obsolete Tax.* This variable equals 1 if the household reports that they have paid the obsolete poll tax in 2018. The exact question text is: 'Let's discuss the poll tax. Did you pay this tax in 2018?'

37. *Market Vendor Fee.* This variable equals 1 if the household reports that they have paid the market vendor fee in 2018. The exact question text is: 'Let's discuss the market vendor fee. Did you pay this tax in 2018?'

38. *Business Tax.* This variable equals 1 if the household reports that they have paid the business tax in 2018. The exact question text is: 'Let's discuss the companies' register. Did you pay this tax in 2018?'

39. *Income Tax.* This variable equals 1 if the household reports that they have paid the income tax in 2018. The exact question text is: 'Let's discuss the income tax. Did you pay this tax in 2018?'

A3 Further Analysis

A3.1 Conceptual Framework

This simple conceptual framework describes a government's decision between collector types in administering a tax collection campaign in a low-compliance setting.[123] We discuss the inputs to the government's choice and the assumptions we make for each. We then discuss how this framework maps to our context and discuss contextual differences and government interventions that could alter the choice between collector types.

A3.1.1 Setup

Property owners: Property owners have intrinsic willingness to comply λ with property tax T (normalized to 1) and encounter costs to non-compliance θ with probability a. λ is a random variable; cost θ may represent the fine (plus tax amount) or punishments such as shaming and a the likelihood of incurring such costs.[124] Taxpayer i complies with the tax if $\lambda_i + a\theta \geq T$.

The probability that a taxpayer pays the tax is $\Pr(\lambda_i \geq 1 - a\theta) = 1 - F(1 - a\theta)$. We assume a and θ are fixed and constant across individuals but can differ by collector type k, and define $\rho_k = 1 - a_k\theta_k$. $\Pr(\lambda_i \geq \rho_k) = 1 - F(\rho_k) = 1 - \int_0^{\rho_k} f(\rho_k)d\rho_k$. The low-enforcement nature of this setting derives from an assumption that ρ_k is small enough that, for a large share of taxpayers, $\lambda_i + \rho_k < 1$: the sum of intrinsic willingness to pay and cost of non-compliance is less than the amount of the tax.

To visualize how λ_k and ρ_k affect the potential number of taxpayers, Figure A16 illustrates distributions of λ by collector type, $f(\lambda_L)$ for Local (L) and $f(\lambda_C)$ for Central (C), for the same population of property owners, and shows values of ρ by collector type. This figure displays a case where $f(\lambda_L)$ is shifted to the right of $f(\lambda_C)$: the intrinsic willingness to pay the tax to type L is higher for most property owners than the intrinsic willingness to pay to type C. However, ρ_C is lower than ρ_L, reflecting a higher cost of punishment for non-compliance under type C than type L. Because in this instance the willingness to pay to type L outweighs type C's enforcement advantage, the proportion of property owners that will pay a collector of type L is greater than the share that will pay a collector of type C, as represented by the shaded portions underneath each curve.

The fraction of property owners who will pay the tax (conditional on being visited by a tax collector) thus depends on the intrinsic motivation λ_k and the cost of punishment for non-compliance ρ_k. This fraction will be higher for type L (and vice versa for type C) if: (*i*) the cost of punishment for non-compliance is the same, $\rho_L = \rho_C$, but the intrinsic willingness to pay the tax λ_k is higher for type L than for type C (which could be consistent with greater tax morale, trust, reciprocity); (*ii*) the intrinsic willingness to pay is the same across collector types ($\lambda_{Li} = \lambda_{Ci}$, $\forall i$), but the cost of punishment for non-compliance for type L ($a_L\theta_L$) is higher, or $\rho_L < \rho_C$ ($\uparrow a\theta \rightarrow \downarrow \rho$) (which could be consistent with greater unofficial costs to non-compliance such as shaming, withholding of services, or informal tax imposition); or

[123] Many of the ideas in this framework were inspired by Dal Bó et al. (2020).

[124] In this simple setup, we do not consider other factors that may be relevant to compliance, such as liquidity constraints, but the interpretation of λ_i could be expanded to include such factors as long as they would raise willingness to comply independent of the other inputs we specify as informing the taxpayer's compliance decision.

(*iii*) in cases such as those illustrated in Figure A16: higher intrinsic willingness to pay is larger than differences in cost of punishment for non-compliance.

Tax collectors: Tax collectors are of type Local (L) or Central (C). A collector makes visits to property owners and solicits tax payment or bribes.[125] Before outlining the collector's objective function, we first define the inputs to the collector's decision below.

Average probability of payment among visited: First, it is necessary to define the average probability of payment, which is generated by visits to property owners. A property owner i only pays — with probability $\Pr(\lambda_{ki} \geq \rho_k)$ — if visited by a tax collector of type k. Therefore, the probability of payment for a household i can be expressed as $v_i \cdot \Pr(\lambda_i \geq \rho_k)$ with $v_i \in \{0, 1\}$ being an indicator for a household receiving a visit.[126]

The average probability of payment among property owners visited by collector type k, which is a function of the total number of property owners a collector decides to visit v, is then:

$$\bar{p}_k(v) = \frac{1}{V} \sum_i v_i \cdot \Pr(\lambda_{ki} \geq \rho_k) \tag{2}$$

where $V = \sum_i v_i$, the number of households visited. $\bar{p}_k(v)$ can be decreasing, flat, increasing, or non-monotonic in v depending on the order — in terms of λ — in which collectors visit property owners.[127]

Collector targeting: Collector types employ different targeting strategies that are a function of their information about property owners' intrinsic willingness to pay and the cost of punishment for non-compliance specific to their type.

For illustration purposes, we highlight an extreme case: type L collector knows the intrinsic motivation λ_i and ρ_L, the punishment probability and cost for collector type L, so rank-orders households by λ_i as the schedule for making tax collection visits (from highest λ_i to lowest). Type C collector knows ρ_C but does not know λ_i for any households, so targets visits randomly. In this extreme case, we assume that $E[\lambda_L] = E[\lambda_C]$ and $\rho_L = \rho_C$: the willingness to pay the tax and the cost of non-compliance are the same across collector types, but collector types differ in their knowledge of property owners' λ's. However, if the number of visits v a collector makes is less than the total number of households N, then $\bar{p}_L(v) > \bar{p}_C(v)$: given type L's ability to rank-order households by willingness to pay, among the households visited by the type L collector the average probability of payment is higher than the average probability among the households visited by the type C collector.[128] This case is illustrated in Figure A17: even for different levels of visits, so long as not all households are visited, $\bar{p}_L(v) > \bar{p}_C(v)$ will hold. Even in a non-extreme case, when type C collectors possess some information about willingness to comply with the tax (curve $\bar{p}'_C(v)$ in Figure A17) — but type L are better informed — the same relationship holds. An alternative way to interpret

[125] In this section, for simplicity we refer to collectors' "tax visits" simply as "visits." In mapping this framework to the setting in Kananga, these would be visits after property registration in which the collector solicits payment of the property tax.

[126] In this simple setting we do not consider multiple (re)visits to households, but v_i could also be thought of as number of visits made to a property owner and the same relationships we identify below will hold.

[127] The average probability in Equation 2 can also be expressed using integrals as $\frac{1}{V} \sum_i v_i \cdot \left[1 - \int_0^{\rho_k} f(\rho_k) d\rho_k\right]$.

[128] If all households are visited by each collector type, then $\bar{p}_L(v) = \bar{p}_C(v)$.

this difference in strategies is that collectors observe signals about λ_i, and type L possess a stronger signal than type C that allows for more effective targeting of visits.

This inequality will also hold in the following cases where $E[\lambda_L] \neq E[\lambda_C]$ and/or $\rho_L \neq \rho_C$ (i.e., when collector types differ in information and property owners differ in intrinsic willingness to pay across collector types and/or the punishment cost of non-compliance across collector types): (*i*) $E[\lambda_L] = E[\lambda_C]$ but $\rho_L < \rho_C$, then $\bar{p}_L(v) > \bar{p}_C(v)$ even if all households are visited: this reflects a case in which the punishment power of type L is higher — the same could hold in reverse; (*ii*) $\rho_L = \rho_C$ but $E[\lambda_L] > E[\lambda_C]$, then $\bar{p}_L(v) > \bar{p}_C(v)$ even if all households are visited: this case could reflect differences in intrinsic motivation to pay that vary by collector type — e.g., type L engenders higher tax morale — the same could hold in reverse;[129] or (*iii*) in cases that depend on relative differences in $E[\lambda_k]$ and ρ_k and collection strategies when all households are not visited: one can imagine cases where, for $E[\lambda_L] = E[\lambda_C]$, $\rho_C < \rho_L$ (cost of non-compliance with type C higher), but type L can target higher-λ property owners in completed visits, while type C targets randomly — in such a case the information advantage of type L overcomes the enforcement advantage of type C.[130]

Corruption: Collectors also have the ability to solicit bribes in lieu of tax payment. This form of corruption captures the cost from the government's perspective of enlisting collectors who may have different incentives to collect bribes instead of taxes across types.[131]

A3.1.2 Collector objective functions

Collectors earn a piece-rate wage that is a portion $\delta < 1$ of the tax $T = 1$ they collect. A collector of type k chooses v and α to maximize expected utility:

$$EU = v\bar{p}_k(v)[\delta(1-\alpha) + \alpha\beta(1-r\omega) - b_k(\alpha)] - \gamma_k \frac{v^2}{2} \tag{3}$$

where v is the number of visits the collector makes according to the collection strategy described above and α is the proportion of potential collections diverted to bribes. The number of total visits conducted has payoff $v\delta\bar{p}_k(v)$ and cost $\gamma_k \frac{v^2}{2}$. The cost of visits is such that each additional visit generates a cost proportional to the visit squared (De Groot, 1988; Dal Bó et al., 2020) and γ_k is a weight that reflects differences in the costliness of visits across collector types. $\beta \leq 1$ is the proportion of the tax amount the collector is able to recover in bribes. Cost $r\omega$ is the cost for an additional unit of bribes in terms of punishment costs for the collector: r is the probability the government catches the corruption and ω is the penalty, which does not differ by collector type. $b_k(\alpha)$ is a cost of corruption that captures the social or psychological costs of corruption (such as increasing negative perceptions of oneself among property owners or guilt over diverting revenues from the state), which may

[129] In this case, type L would generate greater compliance through greater persuasion ability alone, rather than superior targeting, holding constant effort.

[130] Note that knowledge of ρ_k may also generate differences, but in this simple case we assume collectors of type k know the costs of punishment for non-compliance for their own type ρ_k.

[131] We exclude the case of bribes that could be extracted on top of tax amounts or in exchange for reductions in tax amounts as these forms of corruption appear less likely to be common in our setting. Among self-reported bribe payers, 91% did not pay the tax accordingly to our midline measure (83% according to the endline measure), indicating that bribes are most likely paid to avoid the tax.

differ by collector type based on how sensitive they are to citizen views or how aligned they are with government's objectives. We again let this cost be increasing in the square of the proportion of collections diverted to bribes such that $b_k(\alpha) = b_k\frac{\alpha^2}{2}$, where b_k is a marginal cost that can differ across collector types.

For simplicity, we do not model the bargaining process between collectors and property owners over whether to pay a bribe instead of the tax. We assume that when a collector decides to solicit a bribe instead of the tax, there is some probability — built into the portion of the tax the collector is able to recover β — that the property owner will accept, and that the bribe amount will not exceed the tax amount.[132] Collectors therefore determine whether to collect bribes in lieu of tax payment based on the benefits and costs associated with these actions.

FOC: A collector chooses v and α to maximize the objective function, giving first-order conditions:

$$v^* = \frac{\bar{p}(v)W}{\gamma - \bar{p}'(v)W} \tag{4}$$

$$\alpha^* = \frac{\beta(1-r\omega)-\delta}{b_k} \tag{5}$$

where $W = \delta(1-\alpha) + \alpha\beta(1-r\omega) - b_k\frac{\alpha^2}{2}$.

We note the following implications of these conditions.[133] First, v^* is increasing in $\bar{p}_k(v)$: a higher average probability of payment among those visited produces more visits. Because $\frac{\partial \bar{p}_k(v)}{\partial E[\lambda_k]} > 0$, v^* is also increasing in $E[\lambda_k]$, the intrinsic motivation of individuals to pay the tax, and because $\frac{\partial \bar{p}_k(v)}{\partial \rho_k} < 0$, v^* is also decreasing in ρ_k: a lower cost of punishment for non-compliance decreases the return to additional visits. Second, v^* is decreasing in $|\bar{p}'_k(v)|$: a higher marginal (negative) change in $\bar{p}'_k(v)$ for an additional visit means less visits. Third, v^* is decreasing in γ_k: a higher multiplier on cost of effort reduces visits. Finally, α^* is decreasing in $r\omega$ (or: r and ω, separately) and b_k but increasing in β: higher costs to collecting bribes reduce them, higher payoffs increase them.

A3.1.3 The Government's Decision

The government wishes to maximize value from the taxation campaign net the associated costs with employing a collector type. In deciding which collector to engage in collection, it compares:

$$V_L - V_C = (1-\delta)(q_L - q_C) - (g_L - g_C) - \Gamma(c_L - c_C) \tag{6}$$

where V_k is the value the government realizes from employing a collector type.[134] Output $q_k = v_k^*\bar{p}_k(v_k^*)(1-\alpha_k^*)$ are the revenues collected by collector type $k \in \{L, C\}$, and cost g_k represents the cost of engaging a particular collector type outside of sharing δ portion of

[132] The median bribe amount measured at midline and endline in our sample is 1000 Congolese Francs, which is 40% of the median tax rate faced by households. Moreover, 95% of reported bribe amounts are equal to or less than the tax rate assigned to a household.

[133] In order to characterize relationships between inputs into tax collection neatly, we consider only interior solutions.

[134] This is not expressed in purely revenue terms as the government places a negative value on corruption that is separate from revenues lost to bribes.

revenues, such as training, materials, and transportation, but could also represent the costs of monitoring a given collector type to limit corruption.[135]

Cost $c_k = \alpha_k^* \beta v_k^* \bar{p}_k(v_k^*)$ is the amount of revenues lost to corruption, and the difference across collector types is valued by Γ. As the revenue cost of corruption — i.e., pure leakage — is already factored into the first term, this term instead represents the cost to the government of permitting corruption, such as in undermining trust in public personnel and institutions. We express the government's decision in value terms rather than explicit revenue terms to accommodate this non-revenue cost; however, in our discussion in the paper we express Γ as a multiplier that converts these non-revenue corruption costs into revenue terms.

The difference in revenues between collector types is therefore, first, increasing (decreasing) in the square of the average probability of payment for collector type L (type C),[136] reflecting differences described above that may derive from: collector strategy (informational advantage), intrinsic willingness to pay by type (tax morale), and cost of punishment for non-compliance (sanctioning power). Second, it is increasing (decreasing) in the number of visits conducted by type L (type C). Third, it is decreasing (increasing) in the cost of effort multiplier for type L (type C). Fourth, it is decreasing (increasing) in fixed costs to employing type L (type C) g_L (g_C). Finally, it is decreasing (increasing) in the proportion of collections lost to bribes, both from lost revenues and the cost of permitting corruption separate from the impact of revenues.

Beyond the scope of this simple framework is how constraints on the government's ability to administer a taxation campaign may shape the activities of tax collectors. While the cost to effort of tax visits in the collector's objective function (Equation 3) implies that collectors' optimal number of visits will be less than the total number of taxable households, costs on the government side may additionally constrain tax visits and generate differences in the desirability of a particular collector type. In other words, the g_k term in Equation 6, representing the total cost of engaging collector type k, may in reality be subject to budget constraints. For instance, fixed allowable transportation budgets for collectors may limit the number of total possible tax visits.[137] This may in effect impose a ceiling on the number of tax visits the government can pay for, independent of the number collectors would deem optimal according to their own optimization problems.[138] For parsimony, we do not incorporate these government-side constraints in this framework. However, in Section VB, we discuss how such considerations may condition the interpretation of our findings in the context we study.

A3.1.4 Discussion

This simple framework of the government's problem captures the primary margins through which we hypothesize collector types may differ in their ability to generate value in our set-

[135] This second formulation of g_k could represent an "oversight cost" by collector type that could additionally create a wedge between the returns to types.

[136] Also decreasing (increasing) in marginal reduction in average probability of payment $\bar{p}_k'(v)$.

[137] Fixed budgets might well be optimal for the government if the return on additional tax visits becomes negative after collectors have already visited the high-propensity types in a neighborhood, as we consider in Section VB.

[138] For instance, one could conceptualize low state capacity as a ceiling on the number of tax visits, such that collectors are choosing from the set of possible visits in $\{0, \frac{N}{1/a}\}$, where $a \in (0,1)$ is increasing in state capacity.

ting. First, differences at the property-owner level affect the likelihood of payment, and these may differ by collector type: e.g., a collector type may engender stronger tax morale or trust that generates a higher willingness to pay the tax compared to another collector type. Likewise, collector types may differ in the costs they can impose for non-compliance: e.g., state agents (type C) may be better able to impose official penalties for non-compliance, while chiefs (type L) may be able to marshal other forms of punishment such as social sanctions, withholding services, or altering demands for informal taxes. These factors determine the property owner's decision to pay, highlighting that the identity of the collector may itself impact the probability of payment, holding targeting and effort fixed.

Second, differences in collector information about the probability of payment across property owners affects effort (and tax revenues). Though the strategies of collectors described above are relatively deterministic — type L rank orders by λ_{Li}, type C visits randomly, or operates with less information — this formulation captures the intuition that information differences affect the average probability of payment (among visited households) and therefore may also affect collector effort.[139]

Third, differences in the cost of effort by collector type shape effort, reflecting real-world costs associated with traveling to visit individual property owners. We hypothesize that these are mainly a function of the distance between collectors' home location and where property owners live in our setting, but may also depend on physical ability and the opportunity costs of time spent collecting taxes.

Fourth, we hypothesize that chief collectors may experience lower social or psychological costs to soliciting bribes, given they are less aligned with the government than central agents ($b_L < b_C$), and therefore will be incentivized to collect more bribes relative to type C.[140] This difference highlights a cost to employing chiefs that may affect government revenue directly: that chiefs are less aligned with the government's objectives of raising public funds, reflected in the lower social or psychological cost type L faces to collecting bribes, but which also may affect the government's choice of collector type through harms to perceptions of the state. The additional cost of corruption lies in what negative value the government places on such actions, in terms of how the level of corruption might undermine government legitimacy. If chiefs collect more taxes but also more bribes, for example, the government must decide to what extent the costs associated with permitting higher corruption when employing chiefs cancel out the higher revenues they bring in compared to central agents.

The last consideration relates to the cost of employing a particular collector type, which may differ in real-world terms, primarily in the form of compensation for transportation, if collectors differ in their location or payments to collectors entail different logistical challenges. In terms of cost-effectiveness, chief collectors may be more attractive because they live where they collect and thus do not require compensation for transportation to neighborhoods like state agents.

Mapping this framework to our setting, and focusing on the mechanisms through which

[139] The shape of $f(\lambda_k)$ also matters: if there is more (less) variation in λ_i's across citizens, then information is (more) less important. Likewise, how much effort collectors put in will matter more (less): this is reflected in the $\bar{p}'_k(v)$ term, which captures how much the average probability of payment changes with each visit. If all citizens have the same willingness to comply with the tax, then the informational advantage of collector type L is zero.

[140] This assumes that costs to bribes in terms of property owners' perceptions about the government are low.

a given collector type may possess an advantage in collection, as motivated by the literature, we hypothesize that chiefs may generate more tax revenue than state agents if: the intrinsic willingness of property owners to pay chiefs is higher (tax morale), chiefs can impose greater punishments for non-compliance (sanctioning power), or, holding the aforementioned factors constant across collector type, if chiefs are better informed about property owner willingness to comply (informational advantage) or have lower costs to effort (transaction costs of visits). Conversely, state agents may generate more revenues because the punishment costs to non-compliance are greater when state agents are collecting (sanctioning power).

A3.1.5 External Validity

The framework outlined above also allows us to consider in a simple way how contextual differences, or the impact of government interventions, could affect the relative value of employing collectors of a particular type. In this section, we discuss how (*i*) differences or changes in general tax enforcement (through increasing punishment of non-compliance), (*ii*) citizens in a context having higher tax morale or the government raising it through public good provision, (*iii*) access to richer information on citizens, (*iv*) differences or changes in administrative costs, and (*iv*) alignment of collector types with state objectives could affect the government's decision to employ collectors of a particular type.

1. *Higher enforcement*: Increasing enforcement — specifically in terms of punishing property owners for non-compliance by imposing penalties — would shift $\rho_k = 1 - a\theta$. Consider an extreme shift to perfect enforcement such that $a = 1$ ($\theta > 1$). In this case, all property owners will be willing to pay the tax, collector type is irrelevant in terms of informational advantages, and the revenue generation across types will only depend on (*i*) the relative effort and administrative costs of employing a particular type, and (*ii*) differences in bribe solicitation.

2. *Greater public good provision*: Likewise, increasing public good provision may shift λ_C such that $E[\lambda_C] \geq E[\lambda_L]$: citizens will have higher intrinsic willingness to pay the tax to type C, and this reduces the informational advantage of type L or their lower effort costs (or administrative costs).

3. *Information on citizens*: Collecting information on property owners relevant to λ_i's — through data collection or simply observing past compliance behavior — and transferring this information to type C could neutralize the informational advantage of type L. Then differences by collector again would depend more on differences in cost to effort (and administrative costs) and incentives to divert revenues.

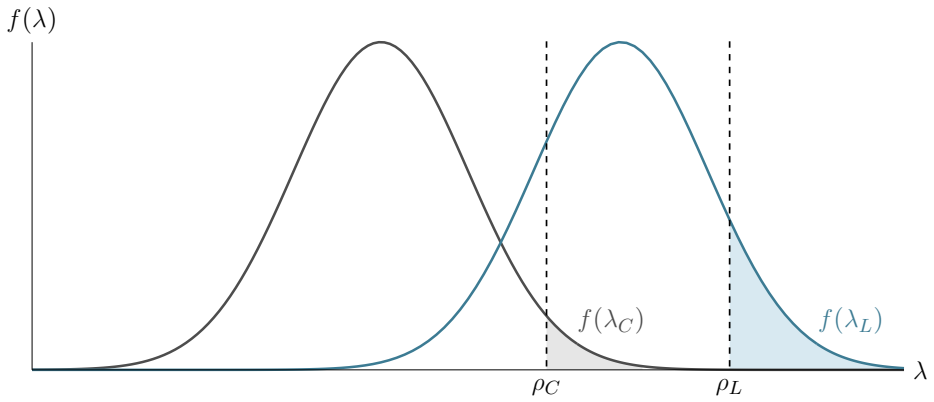
4. *Administrative costs*: Reducing administrative costs by collector type (e.g., selecting type C collectors from each neighborhood and assigning them to collect in their own neighborhood) could neutralize transport costs differences between collector types.

5. *Alignment of collectors with state*: Reducing mis-alignment of type L in terms of the social or psychological costs of soliciting bribes (through giving type L more of an incentive to care about government revenues, potentially through recognition, greater responsibility in other areas, or providing a salary) may reduce a higher hypothesized

prevalence of corruption among type L collectors and therefore make the government's decision more concentrated on the differences in revenue generation between collector types.

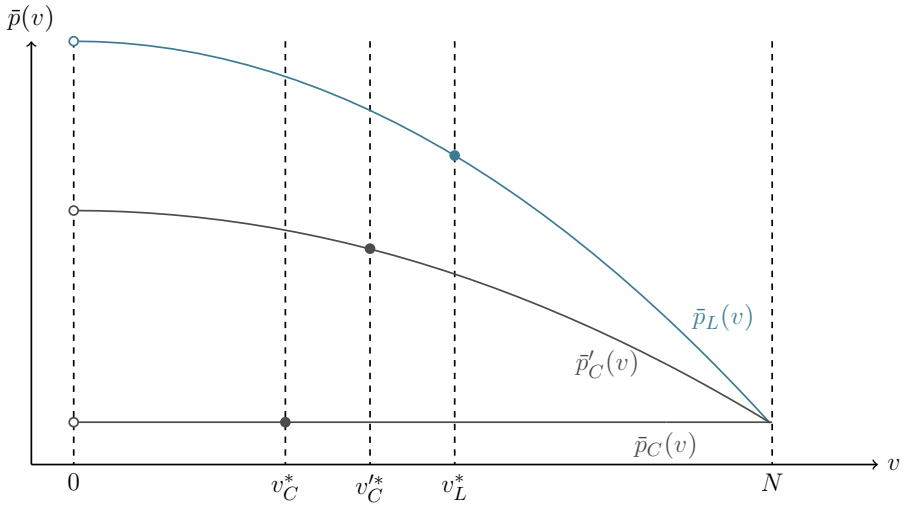
This discussion illustrates the manner in which differences across contexts may change the government's calculus in deciding between collector types or how that calculus may change if the government decides to invest in other approaches to generating tax compliance. In short, the contextual attributes or investments in raising compliance described above would all, in expectation, be positively correlated with the level of development and government resources. This suggests that the tradeoffs we identify, and the salience of the decision between collector types more generally, are intensified in low-enforcement, low-capacity settings, whereas in contexts with higher enforcement or resources for punishing non-compliance, the choice between collector types may be less crucial.

FIGURE A16: EXAMPLE OF POTENTIAL TAXPAYERS BY COLLECTOR TYPE



Notes: Curves $f(\lambda_L)$ and $f(\lambda_C)$ are the distribution of intrinsic willingness; ρ_L and ρ_C are the cost of non-compliance; and the shaded areas are the proportion of potential payers by collector type L and C. This figure is discussed in Section VC and A3.1.1.

FIGURE A17: AVERAGE PROBABILITY OF PAYMENT BY VISITS AND COLLECTOR TYPE



Notes: Curves $\bar{p}_L(v)$, $\bar{p}_C(v)$, and $\bar{p}'_C(v)$ are the average probability of payment among visited property owners by collector type and informedness. v^*_k are the optimal number of visits selected by collectors, N is the total number of property owners. This figure displays the case where $E[\lambda_L] = E[\lambda_C]$ and $\rho_L = \rho_C$: the only difference across collector types in average payment probability derives from the level of information about λ_i's of property owners and number of properties visited. We discuss this figure in Section VB and A3.1.1.

A3.2 Combined Team — Central X Local

Might combined teams — pairing chiefs and state agents together — have promise for raising revenues? This question touches on issues of team production and peer effects, which are beyond the scope of this paper. In our reading, the theoretical literature offers no clear prediction. On the one hand, free-riding issues could be severe (Alchian and Demsetz, 1972); on the other hand, peer effects and productivity spillovers could outweigh free-riding (Kandel and Lazear, 1992). However, Bandiera et al. (2010) show that heterogeneity in social ties moderate peer effects, and chiefs and provincial collectors clearly had quite disparate social networks (and other characteristics). We therefore approach this question in a reduced-form way to shed light on whether pairing one chief and one ministry agent together could provide a policy-relevant package. Reasoning that the chief would contribute local information to the team, while the ministry agent would contribute a more credible threat of enforcement, we had expected that "Central X Local" (CXL) would outperform Central and Local.[141]

However, CXL neighborhoods had tax compliance in between that of Central and Local — and overall quite similar to CLI. Figure A18 documents a compliance trend over time that approximates a linear combination of that for Central and Local. Table A38 summarizes these results. On average, CXL had higher compliance than Central, though the effect on revenues is less robust. Local still outperforms CXL.[142]

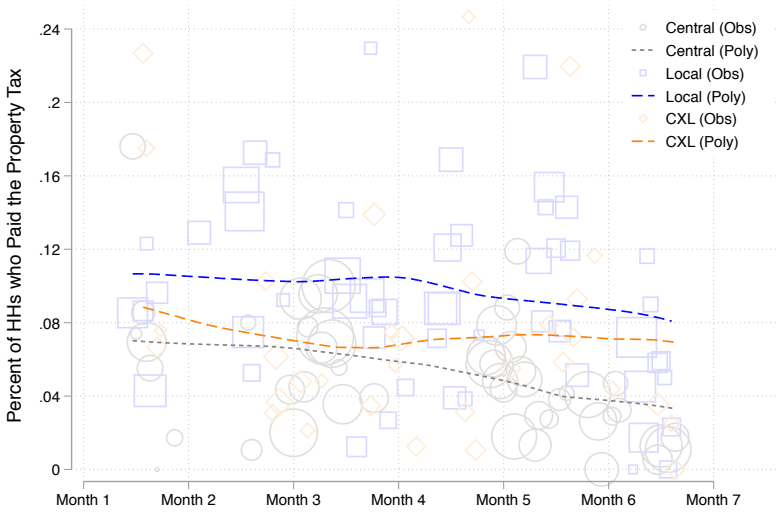
We observe no complementarities or positive peer effects between the chief and state collector. As to why the expected complementarities did not materialize, anecdotally, both types of collectors reported coordination issues in this treatment arm. For instance, chiefs and state agents complained of having problems meeting one another at the time specified, and disagreements over who should be in charge of the receipt printer and tax funds. These coordination problems are reminiscent of the challenges encountered in the hybrid subsidy targeting strategy examined in Alatas et al. (2012). However, the trends in compliance (Figure A18) provide suggestive evidence that the collectors in CXL were perhaps learning how to solve these coordination problems as the campaign went on: they in fact appear better able to counter the secular decline in compliance registered across all other arms. Toward the last period, CXL nearly rivaled Local in compliance.

In sum, on average, the CXL treatment arm achieved lower revenues than Local, yet it had higher costs (because of greater transport costs for state agents). In this setting, delegating tax collection to chiefs appears preferable on most measurable dimensions compared to a hybrid collection model involving collectors of each type.

[141] CXL is also likely easier to implement than CLI, for instance, which reinforces our interest in this arm from a policy perspective.

[142] As noted when discussing CLI and Local, the change in coefficients for CXL in Columns 1 and 6 derives from the change in the definition of the time period fixed effects described in Section IV, which are defined based on the start and end date of the treatments being compared. Thus, when Local is included in the comparison, the time period definition changes to account for trends in compliance over the full period under examination.

FIGURE A18: DECREASING COMPLIANCE OVER TIME — CENTRAL, LOCAL, CXL



Notes: This figure shows the decrease in compliance for Central, Local, and CLI over the tax campaign. Blue squares represent Local observations, gray circles represent Central observations, and orange diamonds represent CXL observations, with size indicating number of observations. Lines — dashed blue for Local, dotted gray for Central, and dashed orange for CXL — are local linear polynomials estimated separately by treatment.

TABLE A38: CENTRAL V. CENTRAL X LOCAL

	Compliance	Revenues	Visited	Visits	Compliance	Compliance
	(1)	(2)	(3)	(4)	(5)	(6)
Central X Local	0.018*	-10.647	0.019	0.065	0.029**	0.013
	(0.010)	(27.683)	(0.037)	(0.061)	(0.014)	(0.010)
Local						0.044***
						(0.007)
Visit Control	No	No	No	No	Yes	No
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	18211	18211	12476	12464	5030	32496
Clusters	142	142	141	141	140	252
Central Mean	.053	156.773	.396	.518	.102	.053
Test CXL=Local (p-value)						0.002

Notes: This table compares the Central X Local (CXL) arm to the Central arm, which is the excluded category. Columns 1, 5, and 6 report impacts on compliance. Column 2 reports impacts on revenues. Columns 3 and 4 report differences in tax visits by collectors after registration by the extensive and intensive margins, respectively. All regressions include fixed effects for house type, randomization strata, and time periods and cluster standard errors at the neighborhood level. All specifications include time fixed effects defined to maximize overlap between the treatments under comparison, as discussed in Section IV. Column 5 restricts to the subsample of properties that received any tax visits after registration. Column 6 includes a dummy for the Local treatment in the regression. The bottom row reports the p-value from a test for equality between the CXL and Local. We discuss these results in Section A3.2.

A3.3 State Collector Team Composition and Performance

One alternative explanation for the higher compliance achieved in Local is that chiefs worked with assistance and thus benefitted from a naturally hierarchical relationship. In Central, collectors were matched with peers without clear hierarchy. If hierarchy led to more efficient team production (Alchian and Demsetz, 1972), this team composition difference could account for the gap in tax outcomes between Local and Central.

To investigate this hypothesis, we exploit the two-staged random assignment of collectors (*i*) into teams, and (*ii*) to neighborhoods to examine if matches of collectors with dissimilar traits corresponds with higher levels of tax compliance and revenue. Specifically, we define a variable *Similarity* that is a dummy for the two randomly assigned collectors both lying either above or below the median for a given collector trait, such as age, education, or income. For instance, in Column 1 of Table A39, *Similarity* equals 1 for all neighborhoods in which both assigned collectors are below the median age as well as for all neighborhoods in which both assigned collectors are above the median age. A negative coefficient would indicate that teams in which the two collectors fall on either side of the median collect more tax, conditional on the average age of the assigned collectors. In fact, we observe that if anything more homogeneous teams of collectors seem to collect more tax. Specifically, collector teams with members who fall on the same side of the median age achieve about 2.9 percentage points higher property tax compliance and 41 Congolese Francs more revenue per owner compared to collector teams that straddle the median age.

Of course, some teams could straddle the median age but still be only a few years apart. Thus, we also consider the absolute value of the difference between the traits of the two collectors — in years for age and level of education (Columns 1–2, 4–5, 7–8, and 10–11) and in dollars for income (Columns 3, 6, 9, and 12). We do not observe that a larger difference between the two collectors' traits is correlated with higher performance as a tax collector team.

TABLE A39: STATE COLLECTOR PERFORMANCE BY TEAM COMPOSITION

	Outcome: Tax Compliance (Indicator)						Outcome: Tax Revenue (in CF)					
	Collector Similarity			Collector Distance			Collector Similarity			Collector Distance		
	Age	Education	Income	Age	Education	Income	Age	Education	Income	Age	Education	Income
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Similarity	0.029**	-0.004	0.014				40.800	8.346	-5.483			
	(0.011)	(0.013)	(0.012)				(33.746)	(33.983)	(37.411)			
Distance				-0.001	0.001	0.000				-0.658	1.102	-0.094
				(0.001)	(0.002)	(0.000)				(2.702)	(5.607)	(0.334)
Observations	187	187	187	187	187	187	187	187	187	187	187	187
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Avg Age Ctrl	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No	No
Avg Educ. Ctrl	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No
Avg Inc. Ctrl	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: This table examines the relationship between state collector team structure and tax compliance (Columns 1–6) or tax revenue (Columns 7–12) at the neighborhood level. The variable *Similarity* is a dummy for the two randomly assigned collectors both lying either above or below the median in the collector trait noted in the column titles. *Distance* is the absolute value of the difference between both collectors' traits, measured in years for age and level of education (Columns 1–2, 4–5, 7–8, and 10–11) and in dollars for income (Columns 3, 6, 9, and 12). All regressions include stratum fixed effects, and robust standard errors. In addition, we control for the average level of the corresponding trait for the assigned collectors in each neighborhood. The sample includes all neighborhoods assigned to Central and CLI, i.e., where state collectors were randomly assigned.

Thus, at least according to this evidence, we find evidence that heterogeneity in collector teams is not associated with higher performance. There is even some suggestive evidence that homogeneity may lead to higher ability among tax collector teams. In a companion paper (Bergeron et al., 2020c), we show that positive assortative matching among tax collectors in the Central arm would increase revenue relative to randomly assigning collectors to each other, which is perhaps consistent with the evidence in Table A39.

A3.4 Collector Exhaustion and Demoralization

Another possible mechanism behind the results is that Central collectors become exhausted or demoralized collecting month after month, while chiefs do not because they typically only collect once. Of course, one would anticipate that Central collectors would also have opportunities to learn and improve as collectors over time. Thus, the prediction is theoretically ambiguous. But it is possible that an exhaustion or demoralization effect overpowers learning and this could explain the lower performance of state collectors compared to chiefs.

To investigate this hypothesis, we examine first if state collectors did fewer tax visits over time, and if so whether this decrease is more pronounced compared to the trend in visits in Local. We find evidence that while Central collectors start doing more visits than chiefs, they end doing fewer visits, and the differential trend is statistically significant (Table A40).

TABLE A40: LOCAL V. CENTRAL: VISITS OVER TIME

	Visited	N Visits
	(1)	(2)
Local	-0.165**	-0.187**
	(0.052)	(0.093)
Local X Time Decile	0.029***	0.036**
	(0.009)	(0.015)
Time Decile	-0.031***	-0.042***
	(0.005)	(0.009)
House FE	Yes	Yes
Stratum FE	Yes	Yes
Observations	18382	18371
Clusters	212	212
Mean	.417	.552

Notes: This table examines visits from tax collector on the extensive (Column 1) and intensive (Column 2) margin across treatments and over time. Specifically, we take deciles of the time distribution of the tax campaign, and interact these with the Local treatment dummy. All regressions include stratum and house type fixed effects, and cluster standard errors at the neighborhood level.

Does this decrease in visits over time explain the higher compliance observed in Local neighborhoods? We conduct several analyses to investigate this possibility, and ultimately we find limited evidence to suggest that this mechanism explains the compliance results.

One test of this mechanism is whether controlling for visits reduces the magnitude of the treatment effect. If the decline in visits were mechanically suppressing tax payment, then controlling for this variable should fully account for the gap in tax compliance we observe between Central and Local. However, when we control for visits on the extensive

and intensive margin, the treatment effect on tax compliance stays intact. This analysis involves conditioning on an outcome of treatment, and should thus be interpreted cautiously. We therefore consider several additional tests.

First, we examine compliance comparing Central to the subset of chiefs who collected in multiple neighborhoods. If there is a kind of exhaustion effect that kicks in when collectors work in more than one neighborhood, then these chiefs would have also been affected by this exhaustion effect. However, the gap in compliance between Central and Local restricted to these repeat neighborhood chiefs remains large and statistically significant (Table A41, Column 1). Thus, it does not appear that chiefs collecting in multiple neighborhoods were less effective tax collectors, as might be predicted by a mechanism in which collecting in multiple neighborhoods leads to demoralization and exhaustion with the task.

Second, we compare Central to a different subset of chiefs who collected in multiple waves of the campaign — and we restrict Local to chiefs collecting for the second time. If the demoralization effect stems from working on the campaign in two consecutive months, rather than two separate neighborhoods, then one would not predict a difference between Central and "Repeat Collector Chiefs." However, the gap in compliance between Central and Local remains substantial even when restricting to this set of chiefs who had already collected in at least one previous wave (Table A41, Column 3). This analysis should be taken with a grain of salt due to the smaller sample size. However, those chiefs who did work month after month, like the Central collectors, still appear to have collected more tax than state agents.

Third, we subset Central to the collectors who were working for the first time. Most of these collectors were from the first wave. But there were 14 other new collectors who joined later in the campaign, too. Thus, "First Time Central Collectors" includes neighborhoods in which at least one assigned collector is working on the campaign for the first time. If demoralization kicks in after the first month — either because of natural exhaustion with the work, or because the comparison with the chiefs becomes more salient — then comparing this subset of Central collectors to chiefs should reveal no gap between collector types. However, the difference in compliance between treatment arms remains large and statistically significant (Table A41, Column 5). This may be the strongest evidence that a demoralization effect does not appear to explain the gap in compliance and revenue that we observe between chief and state collectors.

Fourth, we use the fact that the same Central collectors at times worked alone and at times consulted with chiefs in CLI. Although the program was designed to minimize contact between chiefs and state collectors — who were due to visit the tax ministry at different times of the day and of the week, for instance — the experience of working in CLI might have made salient the fact that chiefs were also working on the tax campaign. If state collectors thought of chiefs as better collectors, this comparison might lead to demoralization after CLI. To test this possibility, we can compare the compliance of Central collectors in the month before CLI and in the month after CLI too see if there is something akin to a trend break — as would be consistent with the new salience of chief collectors reducing effort or causing demoralization more generally.

This analysis is made more complicated by the secular decline in compliance across all treatments. To deal with this, we first estimate the trend in compliance in Local neighborhoods only. Then we compare Central neighborhoods before and after CLI, controlling for

TABLE A41: INVESTIGATING COLLECTOR DEMORALIZATION AND EXHAUSTION

	Tax Compliance					
	Chiefs Working Multiple Nbhds.		Chiefs Working Second Month		First-Time Central Collectors	
	(1)	(2)	(3)	(4)	(5)	(6)
Local	0.035***	0.032***	0.047**	0.051***	0.051***	0.056***
	(0.008)	(0.009)	(0.016)	(0.013)	(0.011)	(0.013)
CLI		0.024**		0.023**		0.034**
		(0.009)		(0.009)		(0.013)
Observations	16642	26064	13049	22471	16505	25927
Clusters	130	210	100	180	129	209
Central Mean	.052	.052	.052	.052	.057	.057
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
House FE	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates from Equation 1, comparing property tax compliance in Local and CLI to Central (the excluded category). All regressions include fixed effects for randomization strata, house type, and time period fixed effects and cluster standard errors at the neighborhood level. Columns 1–2 restrict the Local sample to neighborhoods where chiefs in charge of collection worked in multiple neighborhoods. Columns 3–4 restrict the Local sample to neighborhoods with chiefs who worked in multiple months (in different neighborhoods), keeping only neighborhoods in their second collection period. Columns 5–6 restrict the Central sample to neighborhoods with state agents collecting for the first time. The data include all properties registered by tax collectors merged with the government's property tax database.

the trend (in compliance or revenues) estimated in Local.[143] We summarize the results in Table A42. While the trend is statistically significant (as expected), we do not observe a systematic additional drop in compliance or revenues in months after Central collectors were exposed to the CLI arm (and thus to chiefs working on the tax campaign). Columns 1 and 3 focus on only the first exposure to chiefs in CLI, which occurred in month 2; these regressions thus compare only compliance and revenue in months 1 and 3. Columns 2 and 4 then also consider if there is an additional drop between month 3 and 5, when Central collectors were again exposed to chiefs in CLI (in month 4). Ultimately, this analysis provides little evidence in support of a demoralization effect driving lower compliance after state collectors have exposure to chiefs in CLI.

Finally, perhaps the most direct test of a pure demoralization explanation is to examine collectors' motivation in the survey we conducted with all collectors — state and chief — after the campaign had concluded. In this survey, drawing on the psychology literature (Tremblay et al., 2009) on motivation, we asked about the extent to which collectors were motivated during the campaign by (*i*) extrinsic motivation (working because of the compen-

[143] More formally, we estimate a parametric event study model around CLI exposure timing that allows for a linear Local trend in time, yielding the impact of exposure to CLI relative to the Local trend (Dobkin et al., 2018; Freyaldenhoven et al., 2019).

TABLE A42: CENTRAL: EXPOSURE TO CENTRAL + LOCAL INFORMATION

	Compliance		Revenues	
	(1)	(2)	(3)	(4)
Post CLI Exposure	-0.017	0.012	-126.423	8.685
	(0.075)	(0.018)	(168.455)	(36.417)
Local Trend (Compliance)	1.293	2.032**		
	(2.258)	(0.920)		
Local Trend (Revenues)			0.148	1.510**
			(2.029)	(0.755)
Month Periods Included	1–3	1–3, 3–5	1–3	1–3, 3–5
Time FE	No	No	No	No
House FE	Yes	Yes	Yes	Yes
Stratum FE	No	No	No	No
Period FE	No	Yes	No	Yes
Observations	6447	14164	6447	14164
Clusters	52	84	52	84
Central Mean (Pre-Exposure)	.12	.085	319.104	234.09

Notes: This table reports changes in compliance and revenues within the Central treatment arm, comparing outcomes before Central agents engaged in consultation with chiefs in the CLI arm with those after consultations took place, for the same set of Central agents. We examine two periods: changes in outcomes between months 1 and 3 (for collectors working in the CLI arm in month 2), and between months 3 and 5 (for collectors working in the CLI arm in month 4). We exclude the period straddling the final month of CLI (months 5 and 7), as there are few neighborhoods assigned to the Local treatment arm in month 7. In each period, we estimate the compliance trend in the Local treatment arm and control for it when comparing the pre- and post-periods in the Central treatment arm. All regressions include house type fixed effects. When considering multiple periods we include period fixed effects corresponding to the above-described periods. We do not include fixed effects for stratum or collectors as collectors rotate (due to random assignment to neighborhoods) to different strata and collection partners and thus including these fixed effects would result in a severely restricted sample.

sation), (*ii*) intrinsic motivation (working because they found the work intrinsically rewarding), or (*iii*) introjection (working because the job gave them a positive self-image), or (*iv*) goal orientation (working because they thought the work was socially important / their duty). We also asked a module of questions concerning "amotivation" that address demoralization concerns (Tremblay et al., 2009).

We use these questions to compute standardized indices for each motivation type and then compare the levels among chiefs and central collectors at endline. There are no statistically significant differences between collectors concerning the four aforementioned types of motivation (Table A43, Rows 1–4). However, we see that chiefs report considerably higher (by 0.42 SDs) levels of amotivation at endline (Row 5). This higher level of demoralization among chiefs is also consistent with the negative point estimates for extrinsic, intrinsic, and goal-oriented motivation (though none of these are statistically significant). Exploring the sub-components of the amotivation index, the coefficient is positive for all three survey questions — indicating that chiefs were more likely to agree with each of the statements. But the strongest association is a statement asserting that "our bosses expected too much of us." While these results must be taken with a grain of salt because they are self-reported,

nonetheless they provide further evidence that the state collectors do not appear to have been more demoralized than the chiefs — and if anything the opposite may have been true.

TABLE A43: LOCAL V. CENTRAL: ENDLINE DIFFERENCES IN COLLECTOR CHARACTERISTICS

Panel A: Motivation				
Extrinsic motivation	-0.092	0.222	0.002	111
Intrinsic motivation	-0.308	0.225	0.017	111
Introjection	0.089	0.218	0.002	111
Goal orientation	-0.235	0.212	0.011	111
Amotivation	0.486**	0.218	0.044	111
Panel B: Personality Traits				
Conscientiousness (big 5)	-0.132	0.239	0.003	111
Extroverted (big 5)	-0.384*	0.226	0.026	111
Discount factor	-0.106	0.215	0.002	111
Optimism	0.205	0.216	0.008	111
Locus of control	0.232	0.195	0.013	111
Persistence (maze)	0.727***	0.209	0.122	89
Dishonesty/cheating (RAG)	-0.222	0.213	0.010	111

Notes: This table examines endline differences in collector motivation and personality traits using data from a survey conducted with all collectors after the tax campaign. Each row summarizes a regression of the variable noted on an indicator for chiefs who worked in Local (with the omitted category of state collectors who worked in Central). All dependent variables are standardized to facilitate interpretation of magnitudes. The motivation indices in Panel A come from the psychology literature (Tremblay et al., 2009). The Big 5 indices come from Borghans et al. (2008). Locus of control questions come from the World Values Survey. The persistence measure is the total number of minutes the collector worked on an impossible maze. The dishonesty/cheating measure involves allocating money between oneself and a payoff to the government according to die rolls, as explained in detail in Lowes et al. (2017).

Ultimately, we thus find little evidence to suggest that state collector demoralization or exhaustion led to lower compliance in Central compared to Local. However, it could explain the fact that the slope of the decline in compliance is somewhat more pronounced in Central compared to Local (Figure A5).

Rather than becoming demotivated, it is also possible that state collectors increased the efficiency of their tax visits thanks to learning by doing — e.g., by becoming better at targeting high-propensity households. This explanation would be consistent with the evidence in the paper that targeting of visits to households with higher payment propensities is an important mechanism explaining the higher compliance achieved by chiefs in this context. Similarly, the fact that CLI collectors did similar (or smaller) numbers of visits than Central collectors, and yet they collected more revenue, is further evidence that the composition of visits, rather than the number of total visits, is the key driver of collector efficacy in this setting.

Further evidence that Central collectors learned and became more "efficient" over time

TABLE A44: LOCAL V. CENTRAL: ENDLINE AMOTIVATION

	Couldn't Manage Tasks	Worked Under Unrealistic Conditions	Bosses Expected Too Much	Amotivation Index
	(1)	(2)	(3)	(4)
Local	0.187	0.136	0.651**	0.486**
	(0.214)	(0.217)	(0.212)	(0.218)
Observations	111	111	111	111
Mean (Central)	.161	.094	.484	.369

Notes: This table examines endline differences in collector amotivation using data from a survey conducted with all collectors after the tax campaign. The survey questions were drawn from Tremblay et al. (2009).

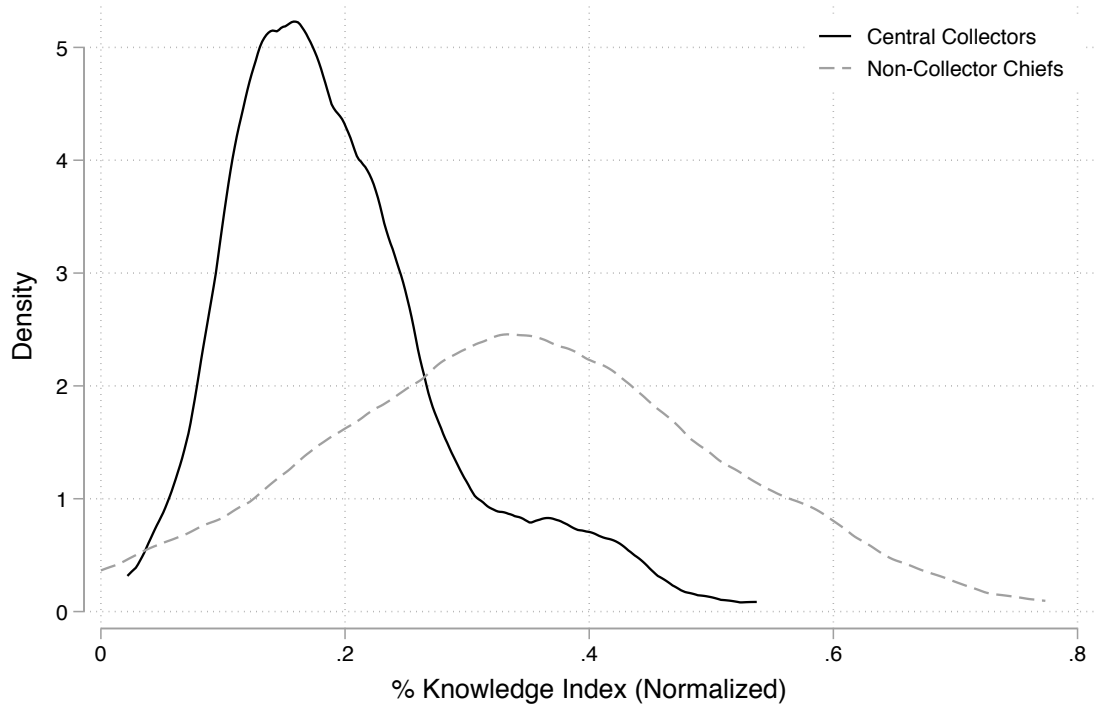
comes from a companion paper in which we examine collector peer effects in the Central arm (Bergeron et al., 2020c). In this paper, we show that being matched with a high-type collector — defined as a collector who achieves a high level of tax compliance across their set of randomly assigned neighborhoods — in time t causes their partner collector to have higher tax compliance in time $t+1$. Specifically, a 1 SD increase in cumulative exposure to high-type collectors increases tax compliance in subsequent periods by 5.1 percentage points ($p = 0.02$). However, the partner collector does not exhibit higher effort in $t+1$ in the form of more tax visits. Rather, they seem to get more efficient at collecting taxes conditional on doing a given number of visits.

A3.5 Quantifying the Knowledge Gap between Chiefs and State Agents

The targeting mechanism assumes that chiefs have access to local information that enables them to better target their tax visits to households with higher payment propensities. To illustrate the knowledge levels of both types of collectors, we administered a quiz-type survey module after the tax campaign concluded. Both types of collectors were shown photos of a set of randomly selected property owners in the chief's neighborhood and asked to provide their (*i*) names, (*ii*) jobs, and (*iii*) education levels. We know the correct answers to these questions from household surveys and can therefore estimate a knowledge index for each collector-neighborhood dyad.

Chiefs took the "quiz" for their neighborhood, while state collectors took it for neighborhoods where they had not worked to estimate the knowledge they would have had at the outset of the campaign. On average, 2.5 state collectors took the knowledge test for each neighborhood, for whom we compute the average accuracy and compare this to the local chief's score. In comparing collector types, we exclude chiefs in Local and CXL because they may have learned about their neighborhoods from collecting taxes. Thus, we restrict the sample of chiefs to all neighborhoods where chiefs did not work as tax collectors (i.e., Central, CLI, and pure control). According to this analysis, chiefs were indeed much better informed about the residents of their neighborhoods than state collectors, scoring about 70% more accurately on this quiz (Figure A19).

FIGURE A19: KNOWLEDGE QUIZ: STATE COLLECTORS V. NON-COLLECTOR CHIEFS



Notes: This figure shows the distributions of knowledge about citizens for chiefs compared to state collectors. Knowledge of the inhabitants of the neighborhood is measured by the percentage of correct answers regarding a random sample of property owners in a short quiz-type survey module conducted after tax collection. Questions included the owner's name, education level, and occupation. Chiefs took quizzes for their own neighborhoods, but we restrict the sample to chiefs who did not collect taxes (since the quiz was administered after the campaign); central agents took quizzes for randomly selected neighborhoods to simulate the knowledge they would have if assigned to a location before collecting taxes there. We discuss these results in Section VB.

A3.6 The Limits to Codifying Local Information

Information is a pillar of state capacity. States must render society "legible" in order to raise revenue and pursue other state-building projects (Scott, 1998). The paper provides direct evidence of the value of local information possessed by city chiefs in raising tax compliance. When equipped with local information, state collectors raised 30.9% more revenue.

However, the results also highlight the limits of the state's ability to codify and harness local information. Some information possessed by chiefs and useful for tax collection appears to have been simply uncodifiable. This conclusion stems from the combination of two observations: (*i*) Local realized higher tax compliance than CLI, and (*ii*) chiefs did not exhibit greater persuasive power. The remaining gap likely reflects the uncodifiable information of the chief that is relevant for tax collection, including "tacit knowledge" about payment propensities of households (Polanyi, 1958).[144]

[144] Polanyi (1958) coined the term tacit knowledge for abilities like facial recognition or language learning that cannot

What aspects of local information are uncodifiable? If such information were truly akin to tacit knowledge, then by definition we could not perfectly characterize it. However, we can compare characteristics of households who were visited after registration in Local and CLI and examine where they diverge. Overall, the characteristics of households visited in CLI are closer to those visited in Local than Central, on both visible and non-visible dimensions (Figure 1).[145] Comparing CLI to Local, the clearest difference concerns liquidity (Figure A14), with CLI collectors somewhat less likely to have visited above-median liquidity households ($p = 0.089$). The uncodifiable component of chiefs' information may thus concern household liquidity. For instance, one possibility is that chiefs received signals about the *timing* of households' liquidity constraints that enabled them to better target tax visits on the time dimension of payment propensity as well as on time-invariant dimensions (e.g., households' underlying tax morale). Such knowledge would have been difficult to convey in a one-off consultation with state collectors. We find suggestive evidence of this possibility by analyzing the time stamps on receipt data, which reveal similar distributions of tax collections occurring primarily in the morning with one crucial difference: chiefs also collected collected a small share of taxes in the evening (Figure A20). This difference in evening collection could explain 40.1% of the remaining revenue gap between Local and CLI.

An alternative interpretation is that chiefs possessed other (codifiable) information that they simply chose not to share during consultations with state collectors in CLI. Although we cannot rule it out entirely, this interpretation appears unlikely given that the households recommended by chiefs in CLI resemble closely the households that chiefs themselves targeted in Local neighborhoods.[146] Moreover, anecdotal evidence from state collectors and program supervisors confirms that chiefs were sincerely engaged during CLI consultations.[147] All told, the results suggest that, in urban settings of low state capacity, the government can achieve better outcomes — from the perspective of the state coffers as well as that of citizens — by delegating collection responsibilities to local elites rather than by trying to integrate their local information into state collection.

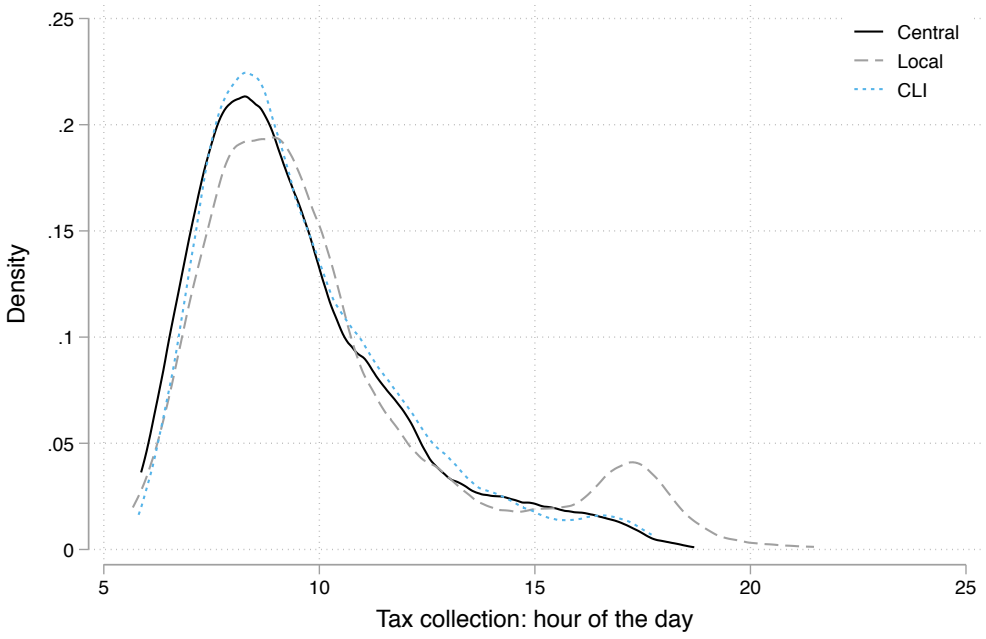
be easily expressed as the sum of explicit, codifiable facts. Williamson (1979) draws on this idea when discussing the appropriate governance structures in markets high in idiosyncratic transaction-specific human capital. Ober (2008) emphasizes the social value of political institutions capable of integrating technical and tacit knowledge.

[145] The similarity between the implied targeting functions of collectors in CLI and Local (rather than Central) provides further evidence about the compositional shift in targeting that led state collectors in CLI to achieve higher compliance than those in Central, as discussed in Section VI.

[146] The co-movement of CLI and Local in terms of tax visits and their correlations with household characteristics is evident in Figure 1 as well as Table 8.

[147] For instance, as noted above, chiefs suggested adding "willingness to pay" — in addition to "ability to pay" — as a field on the form state collectors' filled out during the consultations. They felt an important dimension about households' payment propensity was not reflected in the codification of their knowledge, and unprompted they suggested an amendment to the protocol.

FIGURE A20: TIMING OF TAX COLLECTIONS BY TREATMENT



Notes: This figure shows the distribution of tax payments according to the receipt data. We discuss these findings in Section VB.

A3.7 Cost-Effectiveness

To estimate the cost-effectiveness of state and chief tax collection, we examine campaign data on the marginal costs of tax administration, including transport costs and collector compensation.[148] State collectors were reimbursed for motorcycle taxis from the provincial tax ministry to their assigned neighborhoods. Chief collectors, by contrast, did not incur such costs because they worked near their homes. They were, however, reimbursed for weekly trips to the tax ministry to deposit their tax receipts and receive their bonus. The other key marginal cost was collectors' compensation, which was constant across treatments.

The marginal costs associated with Central and Local are summarized in Figure A21 (Panel A). Chief tax collection has roughly 30% lower administrative costs than state collection. Panel B shows back-of-the envelope estimates of the treatments' cost-effectiveness. The return on $1 is 53% higher in Local compared to Central due to the higher revenues achieved as well as the decreased administrative costs. Moreover, while Local was cost-effective, Central on average was not.[149] Further, this analysis reveals heterogeneity that could guide future policy. State collectors were similar to chiefs in cost-effectiveness when working in the city center, whereas they were much less cost-effective in the city's peripheries (Figure A22). Depending on its assessment of the social cost of bribery (cf. Section VII), governments could opt for collection strategies involving state agents in the city center

[148] Transportation costs, in particular, are emphasized in theoretical work on the tradeoffs between centralized collection and taxation by local elites (Azabou and Nugent, 1988; Levi, 1989).

[149] At the outset of the campaign, state collection was also cost-effective. But the secular decline in tax compliance over 2018 meant that over the course of the campaign, administration costs exceeded tax revenues.

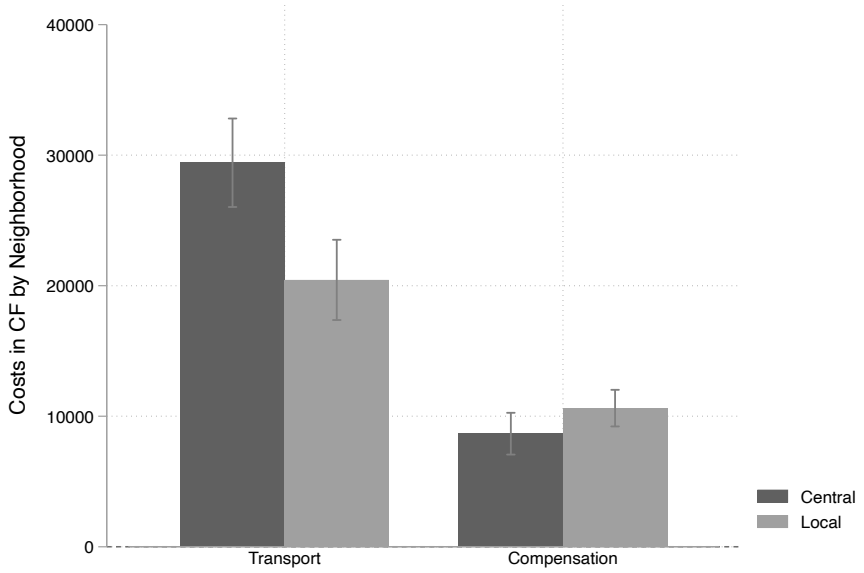
and chiefs in the periphery.

Although the revenue returns to tax administration costs were low, this is a setting of near-zero prior citizen compliance in which the government is making initial investments in fiscal capacity that it hopes will lead to higher revenues over time. Tax officials often discuss their objective of gradually inculcating a "fiscal culture" in Kananga. In other words, the government expects positive inter-temporal spillovers that make the expected future return higher than our calculations. In Section A3.1.5, we discuss how contextual differences and broader fiscal capacity investments could alter the choice of collector type. Yet even low-cost investments, such as mobile remittance of taxes by collectors (already on the tax ministry's agenda), could have large revenue impacts.[150] If chief collectors did not have to make weekly (or biweekly) trips to the government to deposit collections and receive their compensation, we estimate that $1 spent on chief collection would generate $2.1, as shown in Panel B of Figure A21.

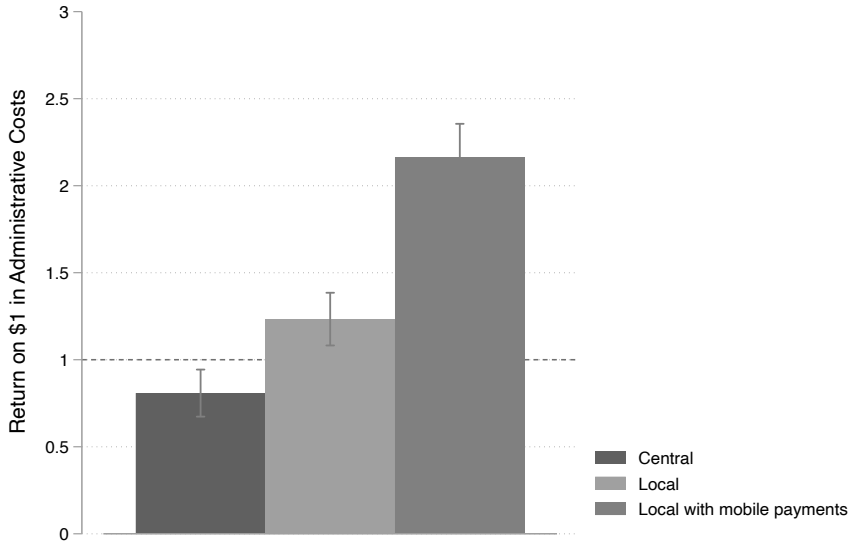
[150]Mobile banking and money transfer services are already widely used in Kananga.

FIGURE A21: COSTS AND COST-EFFECTIVENESS ACROSS TREATMENTS

A: Costs of Tax Collection Methods

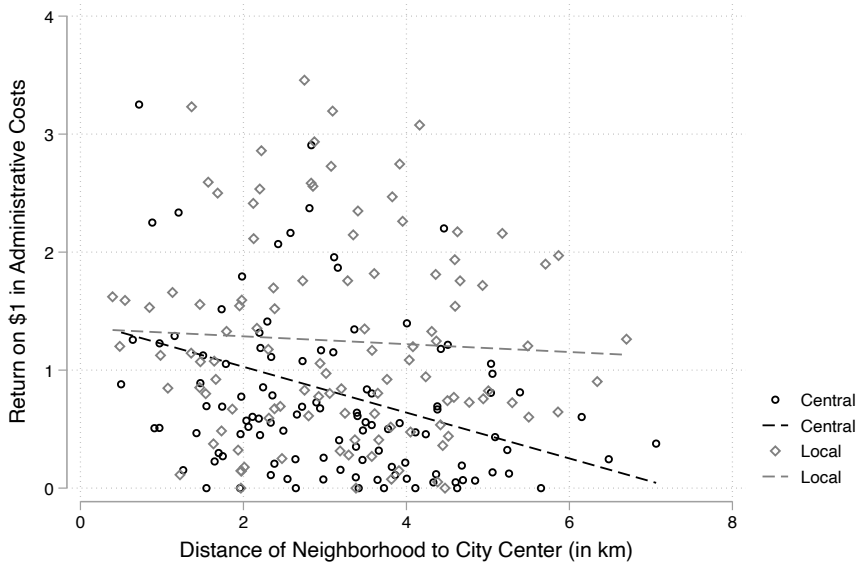


B: Cost-Effectiveness of Tax Collection Methods



Notes: This figure reports estimated costs (Panel A) and cost-effectiveness (Panel B) for the Central and Local treatments. In Panel A, costs are broken down by transport and compensation. In Panel B, cost-effectiveness is the return of an additional $1 spent on collection in particular treatment, and the hypothetical cost-effectiveness of Local with mobile payments is shown at far right. Estimates are the mean value of each measure averaging across neighborhoods. Confidence intervals are shown by the vertical bars. We discuss these results in Section A3.7.

FIGURE A22: COST-EFFECTIVENESS OF LOCAL AND CENTRAL BY REMOTENESS



Notes: This figure reports estimated cost-effectiveness for the Central and Local treatments as a function of the distance from downtown Kananga. We discuss these results in Section A3.7.

TABLE A45: LOCAL V. CENTRAL: BRIBE MULTIPLIER

	Central			Local			
	Revenues	Costs	Bribes	Revenues	Costs	Bribes	Bribe Multiplier
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Campaign Amounts (CF)	2,812,400	4,207,300	117,998	3,514,500	3,197,900	228,488	15.49
With Mobile Money Payment (CF)		4,207,300			1,086,950		34.6

Notes: This table reports measures from the tax campaign of total revenues collected and costs incurred for the Central and Local treatment arms. Columns 1 and 4 report revenues collected by treatment arm. Columns 2 and 5 report costs, which include bonuses paid to tax collectors and compensation for transportation. The second row reports costs under a hypothetical system in which chief collectors were paid (and remit tax collections) via mobile money rather than visiting the tax ministry to receive bonuses (and deposit collections). Costs for Central under this alternative system would remain the same. Columns 3 and 6 show the amounts of bribes collecting according to the measure at endline, scaled by the number of individuals surveyed at endline relative to the neighborhood population of households. All amounts are in Congolese Francs. Column 7 reports the implied multiplier on bribe payments that would be required for the government to weakly prefer employing state collectors instead of chief collectors: $\Gamma = ((R_L - R_C) - (C_L - C_C))/(B_L - B_C)$. This formula is discussed in more detail in Section A3.1.1. We discuss these results in Section A3.7.

A4 Ethical Considerations

The design of this study involved careful consideration of the potential risks to participants. In the following sections, we provide details on these risks and how we endeavored to minimize them, as well as the ethics review process we undertook.

IRB Approval. We obtained approval from Harvard University (protocol IRB17-0724) in 2017, before commencing field research. Our submission outlined the experimental design and included all survey instruments, consent forms, and other material needed to judge the potential risks and benefits to research participants. Although the D.R. Congo does not have a national ethics board, we sought out local ethical approval from the oldest and most highly regarded university in Kananga, the University of Notre-Dame du Kasaï. We submitted the same set of materials and our Harvard IRB protocol to the academic dean of the university. We received a formal approval letter in 2017.

Compensation. Randomly sampled participants in the surveys we administered received compensation to thank them for their time. They were informed of the compensation during the consent, and then received the compensation at the end of the survey. Participants received approximately USD$2 per hour of survey. Thus, the baseline survey took roughly 1 hour, and individuals received USD$2. The midline survey took 20–30 minutes, and individuals received USD$1. The endline survey took 90–120 minutes, and individuals received USD$4. We have used a similar survey respondent compensation amount in Kananga since 2013. We chose this amount based on how other international organizations had compensated survey respondents in the city in the past.

Risks and benefits. In designing the study, we judged the risks to participants to be minimal, in other words, no greater than those they would encounter in the study's absence. Concerning benefits, the data we collected from human subjects enabled us to write an evaluation that may help the government to reduce the incidence of bribe taking and to increase its revenues. We discuss each of these in turn.

The principal risk facing our participants, a random sample of the city population of Kananga, concerned potentially sensitive and identifiable data falling into the hands of other actors, such as the government. There were two primary sensitive topics broached in the surveys.

First, in our surveys, we asked questions about tax payment, bribe payment, as well as attitudes about the government. Since the topics of taxation and corruption concern behavior deemed illegal by Congolese Law, these data were highly sensitive. We were particularly concerned about the government gaining access to survey data and using these data to pursue sanctions against non-compliant (or bribe-paying) households. This was one important risk faced by survey participants.

Second, we also asked questions about the local city chief: their behavior during the tax campaign, their solicitation of bribes, their enforcement of other informal sanctions in the neighborhood among non-compliant households, as well as respondents' views of and trust in city chiefs. We were similarly concerned that these data could fall into the hands of the neighborhood chief and that there could thus be negative consequences among our survey participants.

After consulting with the Harvard IRB and the University of Notre-Dame du Kasaï academic dean, we undertook a number of steps to mitigate these risks as much as possible. We collected all data on password-protected tablets, and we wiped the memory of these tablets on a regular basis. The survey program we use (ODK) also stores responses in XML format and in a folder on the tablet that is difficult to access and interpret unless an individual has prior training. If a government official or the chief gained access to a tablet, they would have had a difficult time accessing the data. We then stored the identifiable data in our research office on password-protected computers. The office is in a walled compound that is guarded 24-7.

In light of these measures, we believe that participation in the study would not represent greater risk than respondents might encounter in their daily lives. Fortunately, there were no instances of lost or stolen tablets during the study, nor reports of theft from the research office.

The benefits of participating in this study — in a research ethics sense distinct from compensation — would primarily accrue at the societal level. Although we did not share identifiable or disaggregated survey data with the government, we did provide a report of our analysis of the impacts of the tax campaign on tax compliance, revenues, and bribe payment. The survey data was an essential component of this report, and it will help the government to improve its tax collection policies in the future.

Such improvements could lead to benefits to citizens in both direct and indirect ways. In terms of more direct social benefits, our evaluation should help the government in its efforts to reduce corruption and bribes collected by tax collectors by providing information about the level and nature of bribe-taking. To the extent that our evaluation helps the government learn how to collect more tax, this could enable the government to provide more public goods in Kananga. Indeed, revenues are sorely needed by the provincial government, which collected on average USD$0.30 per person in the province in 2015. As we note in the paper, low tax capacity is widely regarded as a key development challenge in low-income countries like the DRC (Besley and Persson, 2009).

Regarding indirect benefits, there is evidence that taxation can help promote a social contract between citizens and the government. Indeed, past evidence from the 2016 tax campaign in Kananga suggested that property tax collection raised citizen engagement with the provincial government (Weigel, 2020). We therefore view evaluations of policies used by the provincial government to expand its fiscal capacity as helping to usher in a range of governance benefits related to the tax-based social contract.

Discussion. In light of the potential risks, our measures to mitigate them, and the potential societal benefits from evaluating government tax policies, we firmly believe that this research meets widely accepted ethical standards for social science research. As indicated by the IRB approvals we received from Harvard University and the University of Notre-Dame du Kasaï, the risk-benefit ratio was also judged to be favorable by two different independent bodies with expertise in research ethics.

In addition to the specific risks and benefits to survey participants enumerated above, we discuss here several other ways in which we were involved in the taxation campaign and the possibility that by evaluating this tax campaign implemented by the government our mere presence as international researchers could influence its outcomes in more subtle ways. We

also noted these points in our IRB submissions.

First, the government had planned to collect property taxes and to involve the same types of tax collectors regardless of whether we conducted an evaluation of the campaign. However, the assignment of collectors to different neighborhoods would have not likely been randomized absent the involvement of researchers. As noted in the paper, we conducted the randomization that was ultimately used for the implementation of the tax campaign of 2018. Relatedly, we consulted with the government regarding other elements of the policy experiment design, including (*i*) the number of neighborhoods allocated to each treatment arm, (*ii*) the timing of different waves of the campaign across treatments, (*iii*) the randomization of messages on tax letters, and (*iv*) the mechanics of the Central + Local Information treatment arm.

To inform the allocation of neighborhoods to treatments, we conducted power calculations using data from the logistical pilot of the different types (and combinations) of tax collectors in early 2018. The final allocation included the largest number of neighborhoods in the Central and Local treatment arms, the primary comparison of the policy experiment. Central + Local Information (CLI) had somewhat fewer neighborhoods as a secondary comparison. During the logistics pilot, the Central X Local (one chief and one state collector) teams achieved the highest compliance, so we anticipated it would require relatively less sample to distinguish compliance in this treatment relative to the other treatments.

Given that there was considerable uncertainty ex ante about the outcomes of the different tax collection treatments examined in the context of the 2018 campaign, our position is that randomization was the most equitable way to assign tax collection responsibilities, and likewise for the use of randomization in allocating neighborhoods to different waves of the campaign and assigning message treatments on tax letters. We were pleased to assist the government to do this using our technical background in power calculations and randomized controlled trials more generally.

Regarding the design of the CLI arm, we helped the government during the logistics pilot to evaluate different approaches of transferring knowledge of the neighborhood chief to state collectors. To do this, we interviewed a number of collectors and city chiefs from the pilot neighborhoods. We then synthesized the findings from this process as well as quantitative data from the pilot for the government. As with our role in evaluating the impact of the overall campaign on government revenue, these inputs in the pilot stage of CLI were necessary to learn as much as possible from the campaign about the emergence of tax capacity in weak-state settings.

Second, we conducted technical trainings for tax ministry staff who worked on the tax campaign regarding the receipt printers used by tax collectors. Although these technologies had been purchased by the government in 2015 from an Indian company (KS Infosystems), outside of a handful of tax collectors working at the city's tolls and airport, few tax ministry staff were familiar with the receipt printers and the management of the database associated with them. We therefore helped adapt these devices for collection of the property tax and conducted a series of trainings on the use of these technologies (and the management of data).[151] None of this involvement relates to experimental variation we study in the rese-

[151] In fact, we suggested the government consider an alternative receipt printing technology, but the tax ministry leadership chose to continue using the KS machines for the 2018 campaign.

arch. We view these trainings as important investments in the technical capacity of the provincial government. The goal of the government in using the handheld receipt printers was to create a paper trail for tax collectors in order to enhance monitoring capacity and reduce the payment of bribes. We were pleased to help the government with this goal.

Third, it is possible that the very fact of our conducting an evaluation of this campaign may have changed the behavior of tax collectors or other government officials, akin to a more macro-level "Hawthorne Effect." We of course cannot rule out this possibility because we do not observe the counterfactual campaign (in which we did not conduct an evaluation). However, we suspect any such influences would likely be benign from a research ethics point of view.[152] For instance, if tax collectors learned of the surveys our enumerators were conducting in the city to evaluate the campaign, it would have most likely led them to behave in a more professional manner and to collect fewer illicit payments. We do not think there are plausible scenarios in which awareness of the evaluation could have created incentives for collectors to act in ways that would reduce the welfare of average citizens in Kananga. This is all of course quite speculative, and we do not wish to overestimate our ability to predict the direction of such big-picture "Hawthorne Effects." However, we wanted to note that these were factors we took into consideration when deciding whether and how to conduct this research.

[152] From an internal validity perspective, we took steps to ensure that any information about our evaluation was kept constant across treatment groups. For instance, all tax collector trainings were identical.

References

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A Olken, and Julia Tobias, "Targeting the poor: evidence from a field experiment in Indonesia," *American Economic Review*, 2012, *102* (4), 1206–40.

Alchian, Armen A and Harold Demsetz, "Production, information costs, and economic organization," *The American Economic Review*, 1972, *62* (5), 777–795.

Azabou, Mongi and Jeffrey B Nugent, "Contractual Choice in Tax Collection Activities: Some Implications of the Experience with Tax Farming," *Journal of Institutional and Theoretical Economics*, 1988, pp. 684–705.

Balán, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan Weigel, "The Origins of Tax Compliance and State Capacity: Tax Intervention," *AEA RCT Registry*, 2018.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social incentives in the workplace," *The Review of Economic Studies*, 2010, *77* (2), 417–458.

Bergeron, Augustin, Gabriel Tourek, and Jonathan Weigel, "The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC," *Working Paper*, 2020.

__ , Pedro Bessone, Gabriel Tourek, and Jonathan Weigel, "Bureaucrat Quality, Peer Effects and Optimal Matching: Evidence from Tax Collection," *Working Paper*, 2020.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation and Politics," *American Economic Review*, 2009, *99* (4), 1218–1244.

Blumenthal, Marsha, Charles Christian, Joel Slemrod, and Matthew G Smith, "Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota," *National Tax Journal*, 2001, pp. 125–138.

Borghans, Lex, Angela Lee Duckworth, James J Heckman, and Bas Ter Weel, "The economics and psychology of personality traits," *Journal of human Resources*, 2008, *43* (4), 972–1059.

Dal Bó, Ernesto, Frederico Finan, Nicholas Y Li, and Laura Schechter, "Government decentralization under changing state capacity: Experimental evidence from Paraguay," *Econometrica*, 2020.

Dobkin, Carlos, Amy Finkelstein, Raymond Kluender, and Matthew Notowidigdo, "The Economic Consequences of Hospital Admissions," *AEA RCT Registry*, 2018.

Freyaldenhoven, Simon, Christian Hansen, and Jesse Shapiro, "Pre-Event Trends in the Panel Event-Study Design," *American Economic Review*, 2019, *109* (9), 3307–3338.

Gibbons, Charles E, Juan Carlos Suárez Serrato, and Michael B Urbancic, "Broken or fixed effects?," *Journal of Econometric Methods*, 2018, *8* (1).

Groot, Hans De, "Decentralization decisions in bureaucracies as a principal-agent problem," *Journal of Public Economics*, 1988, *36* (3), 323–337.

Iacus, Stefano M, Gary King, and Giuseppe Porro, "Causal inference without balance checking: Coarsened exact matching," *Political analysis*, 2012, *20* (1), 1–24.

Kandel, Eugene and Edward P Lazear, "Peer pressure and partnerships," *Journal of Political Economy*, 1992, *100* (4), 801–817.

Levi, Margaret, *Of Rule and Revenue*, University of California Press, 1989.

Lowes, Sara, Nathan Nunn, James A Robinson, and Jonathan Weigel, "The evolution of culture and institutions: Evidence from the Kuba Kingdom," *Econometrica*, 2017.

Miguel, Edward and Michael Kremer, "Worms: Identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 2004, *72* (1), 159–217.

Ober, Josiah, *Democracy and knowledge: Innovation and learning in classical Athens*, Princeton University Press, 2008.

Polanyi, Michael, *Personal Knowledge: Towards a Post-Critical Philosophy* 1958.

Pomeranz, Dina, "No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax," *The American Economic Review*, 2015, *105* (8), 2539–2569.

Scartascini, Carlos and Lucio Castro, "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment," *Journal of Economic Behavior and Organization*, 2007, *115*, 65–82.

Scott, James C, *Seeing Like a State: How Certain Schemes to Improve the Human Condition have Failed*, Yale University Press, 1998.

Tremblay, Maxime A, Céline M Blanchard, Sara Taylor, Luc G Pelletier, and Martin Villeneuve, "Work Extrinsic and Intrinsic Motivation Scale: Its value for organizational psychology research.," *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 2009, *41* (4), 213.

Weigel, Jonathan, "The Participation Dividend of Taxation: How Citizens in Congo Engage More with the State when it Tries to Tax Them," *Quarterly Journal of Economics*, 2020.

Williamson, Oliver E, "Transaction-cost economics: the governance of contractual relations," *The Journal of Law and Economics*, 1979, *22* (2), 233–261.