

BUILDING STATE AND CITIZEN: How Tax Collection in Congo Engenders Citizen Engagement with the State

JONATHAN L. WEIGEL[†]

This version: March 7, 2019

First version: September 21, 2017

Abstract

I examine a field experiment randomizing door-to-door tax collection across 431 neighborhoods of a Congolese city. I test the hypothesis that citizens will demand more inclusive governance when they are taxed. As predicted, the campaign increased political participation by 5 percentage points (28%): citizens in taxed neighborhoods were more likely to attend townhall meetings hosted by the government or to submit evaluations of its performance. I argue that citizens participate more because tax collection sends a signal of state capacity, raising the expected benefits of participation. Analysis of respondents' beliefs about government capacity supports this mechanism.

*My sincere thanks to Horacio Larreguy, Nathan Nunn, Rohini Pande, and James Robinson. I also thank Jie Bai, Robert Bates, Rema Hanna, Anders Jensen, Asim Khwaja, Michael Kremer, Amy Pond, Oyebola Okunogbe, Wilson Prichard, Gautam Rao, Otis Reid, Vincent Pons as well as participants at WGAPE 2017, NEWEPS-8, the 2018 'Rookiefest' Northwestern Development Economics Conference, NHH's FAIR Seminar, and the Zurich 2018 PF-Dev Conference for invaluable suggestions. For outstanding research assistance, I thank Anissa Chabib, Morgan Yucel, and Gabrielle Corrigan as well as Elie Kabue Ngindu and a superb team of enumerators. I am grateful for the collaboration with the Provincial Government of Kasai Central and for support from the Weiss Family Program Fund, the Harvard WCFIA, and the Fulbright Scholar Program. AEA pre-registration ID: AEARCTR-0001316. The Appendix and PAP can be found at jonathanweigel.com.

[†]London School of Economics (email: j.weigel@lse.ac.uk)

1 Introduction

There is growing consensus that state capacity, and especially tax capacity, is critical for development (Besley and Persson, 2009; Acemoglu and Robinson, 2017). Recent empirical work explores why developing countries collect so little tax — 15% of GDP compared to 40% in developed countries — and how they could collect more (Pomeranz, 2015; Khan et al., 2015). However, the political economy implications of increasing tax enforcement in developing countries have received less attention, despite well-known theories of taxation and the social compact. According to classic accounts, when European rulers began systematically taxing their subjects in the early modern period, it triggered new demands for public goods and representation (Schumpeter, 1918; Tilly, 1985; Bates and Lien, 1985; North and Weingast, 1989). This bargaining process between citizens and the state is thought to underlie the co-evolution of tax compliance, participation in politics, and accountable governance. The slogan “no taxation without representation” captures the intuition. This paper examines a key proposition of this social compact theory: that tax collection increases citizen demand for political participation.

This “tax-participation hypothesis” is difficult to test because most governments do not randomly tax their citizens. On the contrary, governments may strategically target certain sectors to maximize revenues while minimizing distortions, but individuals in these sectors may be better able to participate for other reasons (Bates and Lien, 1985). Moreover, the causal arrow could go the other way if elites strategically extend the franchise to justify raising taxes to pay for needed public projects (Lizzeri and Persico, 2004). The positive relationship between tax receipts and political participation in observational data is thus difficult to interpret (Prichard, 2015). Theoretically, it is also not obvious that citizens would choose to engage *more* with a state seeking to tax them. Citizens might prefer to free ride on the advocacy of others, or to move where the state cannot tax them (Tiebout, 1956).

I test the theorized link between taxation and participation by conducting the first field experiment to randomize tax collection — across 431 neighbor-

hoods (covering roughly 33,000 properties) of Kananga, D.R. Congo (DRC), in 2016. In collaboration with the Provincial Government of Kasai Central, I randomly selected 253 neighborhoods to receive the initial phase of the first-ever property tax campaign in the city. In treated neighborhoods, tax collectors went door to door making in-person appeals for the roughly \$2 property tax, which they collected on the spot, issuing printed receipts to payers. Control neighborhoods remained in the old declarative system: citizens were supposed to pay at the tax ministry themselves, but in practice less than 1% did.

Before considering the effects of the campaign on political participation, I examine the ‘first stage’ — whether the campaign achieved the government’s goal of raising tax compliance, and thus whether it constitutes a valid test of the tax-participation hypothesis. Despite the information frictions and state-capacity constraints that inhibit tax collection in settings like Congo (Gordon and Li, 2009), the campaign raised property tax compliance from 0.1% in control to 10.3% in treated neighborhoods. Although low by developed-country standards, this 100-fold increase made property tax receipts just under 5% of the provincial government’s total revenue, on par with local governments in more prosperous African countries.¹ The provincial government evidently viewed the campaign as a success, choosing to continue field-based property tax collection after 2016.

Given that the campaign caused a substantial increase in property tax collection, I use its random assignment to test the tax-participation hypothesis: that tax collection will increase demand for citizen engagement with the provincial government. To measure such engagement, I use two real-world channels of participation that I measured by collaborating with the government.² First, the government hosted a series of townhall meetings, in which officials and citizens had a dialog about taxation and public spending in Kananga. Second, citizens could submit anonymous evaluations of the provincial government to a drop box downtown whose contents would be shared with

¹Property taxes make up 14% of local government revenues in Ghana, 10% in the Gambia, 6% in Sierra Leone, and less than 1% in Liberia and Cameroon (Fjeldstad et al., 2017). Moreover, property tax receipts are typically much lower outside of national capitals.

²This approach is similar to that of Olken (2007), Casey et al. (2012), and Paler (2013).

the governor and other top officials. Attending a townhall or submitting an evaluation exhibits willingness to exert costly effort to have a voice in the provincial government.

Tax collection increased participation according to both measures: residents of treated neighborhoods were nearly 5 percentage points more likely to attend a townhall meeting or to submit an evaluation form, a 28% increase compared to control. Consistent with historical accounts of taxation stimulating citizen-state bargaining, townhall participants demanded better public infrastructure and a more responsive government in exchange for taxes.³ Submitted evaluation forms were also highly critical — over 90% expressed disapproval of the government — with demands for greater transparency, inclusiveness, and public goods spending.

I rule out several alternative explanations of the observed increase in participation. First, I show that familiarity with and trust in the research team is balanced across treatment and control, making it unlikely that the results are explained by an artifact of survey enumeration. Second, I demonstrate that the treatment effect is not caused by a decline in participation in the control group, rather than an increase in participation the treatment group. Third, I present evidence that this result does not reflect a sense of unfairness stemming from awareness of the control group, which had not received tax collectors when outcomes were measured. Finally, given the 6-8 month time gap between tax collection and participation, it is unlikely that the main result could be explained by a salience effect of taxation in treatment.

I consider several extensions of the analysis. I first provide suggestive evidence that the tax campaign and the increase in participation with the provincial government crowded out engagement with local city chiefs, consistent with the idea that the formal state and local forms of governance are substitutes (Cheema et al., 2006). Then, I find little evidence of spillovers on compliance or participation, or of decay of the treatment effect over time.

³“Erosion threatens our neighborhoods, and the government does nothing,” asked one individual, “so why should we pay?” 71% of citizen comments were similar complaints about the inadequate level of public goods to justify taxation, or general demands for transparency and less corruption.

I next turn to the mechanism linking tax collection and participation. It is often assumed that tax-*payers* participate more because they expect a quid pro quo (Prichard, 2015) or exhibit an endowment effect (Martin, 2014). However, in this case, the increase in participation was not driven by the 10% of treated individuals who paid. Rather, *everyone* in treatment neighborhoods — payers and nonpayers — participated at higher levels, compared to people in control neighborhoods. According to instrumental variables estimation as well as simple correlations, the tax collection campaign does not appear to have increased participation through its effect on tax payment. The campaign also does not appear to have lowered coordination costs associated with participation, another possible mechanism.

Instead, I argue that the tax campaign sent a signal of state capacity that raised the expected benefits of participation. In settings like Kananga, the formal state is effectively absent from most citizens' lives. Then, observing a systematic door-to-door tax campaign in their neighborhood leads citizens to believe that the government is more capable of meaningfully impacting their future well being than they had previously thought.⁴ They therefore anticipate greater returns to engaging with the government to try to influence future public goods spending and tax policy. I outline a simple decision-theoretical framework in Section 6.2 to make this argument concrete.

Consistent with this mechanism, I show that the treatment effect is more pronounced in areas with less past exposure to the formal state — areas in which the signal sent by the campaign would have been stronger. I also examine the effects of the tax campaign on individuals' stated beliefs about the provincial government's 'extractive' and 'productive' capacity (Besley and Persson, 2009). Citizens in treated neighborhoods updated their beliefs in two main ways: (i) they realized the government had more revenue due to greater extractive capacity; and (ii) they believed the government would spend more of that revenue productively. Interestingly, however, treated citizens who chose to participate were less certain that these new revenues would be spent well in the absence of citizen monitoring. Thus, treated citizens appear to have

⁴Coate and Morris (1995) similarly model the informational aspects of public projects.

participated more because they believed the government capable of higher spending, but they also perceived a need to monitor that spending and sought to influence it toward their preferred policy. That awareness of new tax revenues would stimulate participation is consistent with evidence from Brazil showing that municipalities were less corrupt and spent more on public goods when revenues came from taxes not transfers (Brollo et al., 2013; Gadenne, 2017).

This paper provides the first field-experimental test of the tax-participation hypothesis, a key proposition of theories of the social compact and the origins of inclusive governance (Schumpeter, 1918; Tilly, 1985; Bates and Lien, 1985; North and Weingast, 1989) as well as the political resource curse literature (Brollo et al., 2013; Gadenne, 2017). The closest past studies are lab experiments that simulate taxation and participation (Martin, 2014) and survey experiments that prime citizens about the share of taxes in government revenues (Paler, 2013; de la Cuesta et al., 2015). There is also evidence from developed countries of electoral payoffs from technologies that reduce tax evasion (Casaburi and Troiano, 2015). In addition to providing evidence on the tax-participation link, the paper outlines a novel mechanism whereby tax collection signals state capacity and raises the expected benefits of participation. This informational component of tax enforcement is relevant for understanding historical state building (Levi, 1989; Brewer, 1990) as well as the political economy implications of present-day tax reforms.

The paper also contributes to the empirical literature on tax and development, which has focused less on the political economy effects of increasing tax enforcement (Besley and Persson, 2013). Past work examines how governments can raise compliance through third-party reporting (Kleven et al., 2011; Pomeranz, 2015; Naritomi, 2015; Carrillo et al., 2017; Jensen, 2018), tax collector incentives (Khan et al., 2015), providing information about enforcement or peer behavior (Del Carpio, 2013; Pomeranz, 2015), and reducing bureaucratic barriers to compliance (Kleven and Waseem, 2013; Best et al., 2015). In addition to the main result for participation, this paper shows that governments in settings of near-zero compliance can raise property tax receipts

through in-person tax appeals that reduce the transaction costs of payment.

Finally, to my knowledge, the paper examines the first field experiment randomizing tax collection. Closest in this regard are Dunning et al. (2015), who randomize tax holidays in Uruguay, and Khan et al. (2015), who randomize tax collector incentives.

The paper reviews the setting (Section 2), experimental design (Section 3), and data, estimation, and balance (Section 4), before turning to the main results (Section 5) and mechanisms (Section 6).

2 Setting

The DRC is the fourth most populous country in Africa, and one of the five poorest in the world.⁵ Median monthly household income in the study site is roughly \$70, PPP \$111, (Lowes et al., 2017). The country is often termed a ‘kleptocracy,’ due to the corrupt rule of long-time president Mobutu Sese Seko, or a ‘failed state,’ due to its history of civil conflict (Sanchez de la Sierra, 2019). It has low state capacity across all dimensions, and especially in terms of tax capacity. In tax revenue as a percent of GDP, the DRC ranks 188 out of 200 countries for the period 2000 to 2017.⁶

Kananga, a city of roughly 1 million, is the seat of the Provincial Government of Kasai Central. With nearly 6 million people in the province, total provincial tax receipts from 2010-2015 were around \$2 million per year. These receipts chiefly came from trade and rental taxes levied on a handful of firms in downtown Kananga, such as mining and mobile-phone companies. Although there are many taxes on the books, few are enforced among private citizens in Kananga. Before the 2016 property tax campaign, only 40% of individuals knew the name of the provincial tax ministry, and 5.6% of individuals in the sample knew of the property tax (see Appendix Table 1). The most common taxes that residents of Kananga paid were market fees and a vehicle tax for owners of cars and motorcycles. But less than 10% of individuals reported

⁵See, e.g., <http://wdi.worldbank.org/table/1.2>.

⁶Data available: <https://data.worldbank.org/indicator/gc.tax.tot1.gd.zs>.

paying any taxes in 2015.⁷ The lack of a broad tax base is a challenge to governments across the developing world (Gordon and Li, 2009).

Property taxes are thought to be efficient and progressive, and urbanization in Africa is fueling rapid growth in real estate values, making a strong case for property taxation (Fjeldstad et al., 2017). Because valuations can be difficult for low-capacity governments, many African municipalities have simplified property valuations to size-based assessments or fixed-amount levies on properties under a certain threshold (Fjeldstad et al., 2017). The Provincial Government of Kasai Central has followed suit. Roughly 90% of property owners in Kananga must pay a fixed annual property tax of 2,000 Congolese Francs (CF), about \$2, which is the median household's total daily income.⁸ Larger 'midrange' houses built of modern materials (i.e. not mudbricks), about 9% of total property owners, pay 6,600 CF. Finally, 'villas', Belgian-built compounds with a garage (1% of property owners), are measured; their owners face a rate increasing in size. Prior to the 2016 tax campaign, property owners were supposed to visit the tax ministry themselves to pay. But except for a handful of firms, compliance remained near zero.⁹

Why did the provincial government begin enforcing the property tax in 2016? According to the former finance minister, an unanticipated national policy triggered a 40% reduction in provincial tax receipts, leading the provincial government to increase property tax enforcement to recoup these losses. Specifically, the 2015 *découpage* (administrative splitting) of the 11 old provinces into 26 new provinces meant that the diamond-rich region around Tshikapa, a large source of revenue for the Kananga-based tax ministry, was no longer part of the province.¹⁰ Facing shortfalls, the governor turned to property taxes.

⁷This low figure is partially offset by contributions in informal taxes (Olken and Singhal, 2011), the most notable of which is *salongo*, an activity organized by local notables (avenue chiefs) in which citizens sweep the streets and clean up after storms. About 30% of respondents reported that *salongo* occurs at least once per month in their neighborhood, though only 16% of households reported regularly contributing.

⁸Properties owned by state employees, churches, and the elderly are exempted.

⁹Of the <300 property tax payments recorded in 2015, 86% were made by firms.

¹⁰Although decentralization was noted in the 2007 constitution, its sudden implementation in 2015 was a surprise, as evidenced by the chaos it engendered in provincial-level politics (Wille, 2015). The *découpage* is widely thought to be a tactic of incumbent Joseph Kabila

The government, though on paper a democracy, is authoritarian, and citizens have few formal avenues of participation in politics. Elections were canceled in 2016 and again in 2017. Nonetheless, individuals in Kananga voice grievances to their political leaders in two main ways. First, they hold local meetings about public-good failures and other political demands and then nominate a representative to bring the case before a provincial deputy. Second, individuals, or groups of individuals, author formal letters of complaint to the provincial government. The measures of participation used in this study are versions of these forms of political engagement.

In sum, Kananga is a good setting in which to test the tax-participation hypothesis because it shares key features with the states in early modern Europe examined in historical accounts of the emergence of inclusive institutions and the social compact (Schumpeter, 1918; Tilly, 1985). These theories discuss low-capacity autocratic states, struggling to cope with fiscal crises by building a tax bureaucracy. In early 2016, the Provincial Government of Kasai Central similarly sought to systematize property tax collection in response to a sudden drop in revenues brought on by an external shock.

3 Experimental design

The treatment, randomly assigned on the neighborhood level, is the door-to-door property tax collection campaign, which ran from April to December in 2016. I defined the unit of randomization, the neighborhood, by dividing a satellite map of the city into 431 polygons that approximate *localités*, the lowest administrative unit in the city.¹¹ Neighborhood borders are typically natural boundaries like roads, ravines, or other features easily identifiable from the ground. Among the 431 polygons, 253 were selected randomly to receive the tax campaign in its first phase. The 178 control polygons were scheduled

to sew bureaucratic confusion and justify postponing the 2016 elections, which he did.

¹¹The government did not have maps of *localité* borders, hence the need to define these on a satellite map. See Appendix Figures 1 and 7 for examples.

to receive the tax campaign in mid 2017.¹²

For the randomization, I constructed 33 strata defined by (i) satellite grid cells of Kananga, and (ii) the estimated population of the neighborhood.¹³ Stratifying in this way addresses a potential inference problem that the experiment was designed to solve: the targeting of certain households or neighborhoods when states extend the tax net. For instance, the state might differentially tax wealthier areas, whose inhabitants may be more likely to participate for other reasons independent of taxation. Because wealth and other characteristics revealed to be in the tax collectors' selection function cluster spatially in downtown urban areas, stratifying on geographic location and population helps improve balance along these key dimensions that may be particularly vulnerable to selection.

Before the tax campaign, households in *all* neighborhoods received informational fliers in French or Tshiluba, the most widely spoken local language, announcing that (i) the provincial government would be collecting property taxes in the months ahead, and (ii) money collected would be used to “promote the economic development of the province.”¹⁴ Distributing fliers in treatment and control helps ensure that estimated treatment effects reflect the impact of tax collection rather than simply information about the tax or the campaign.

The 51 government tax collectors working on the property tax campaign were randomly assigned to new teams of three every twelve work days.¹⁵ Teams

¹²The government ultimately decided to suspend all tax collection in 2017 after violence broke out in the province early that year. It recommenced property tax collection in 2018. For information about the conflict, see Appendix Section 1.3.

¹³I used 11 satellite map grid cells that fully partition the city. Population in each neighborhood was estimated by counting houses visible from satellite images.

¹⁴See Appendix Section 2.1 for more information about these fliers.

¹⁵The collectors were 78% male with an average age of 33 years. All of them were from Kananga and fluent in Tshiluba, the local language. Roughly half were full-time employees of the tax ministry, and half were interns. In keeping with standard policy at the tax ministry, a small performance-based bonus was paid out to those working on the campaign: 18% of the total deposited. This size bonus is analogous to the incentive pay for Pakistani property tax collectors in Khan et al. (2015). Additionally, 40% of property owners in each treated neighborhood were randomly sampled before tax collection for a double bonus: collectors received 36% of the money they collected from these households. This randomized double bonus is examined in a separate project on the effects of collector characteristics on tax compliance. The average weekly bonus was about \$4, though more

were then randomly assigned to treated neighborhoods. The order of neighborhoods was also random. Collectors completed two tasks in each neighborhood.

1. **Property register.** First, collectors completed a brief property register to identify all liable property owners in the neighborhood. Collectors assigned a unique code to each house, written in chalk on the wall or door. These codes appear on tax receipts to identify compliant households in the administrative data. The property register was verified by members of the research team with GPS devices to ensure the collectors respected neighborhood boundaries. Collectors received a printed copy of the register before tax collection.
2. **Tax collection.** After completing the register, collectors began door-to-door tax collection. When an individual paid the tax, collectors used a tablet application to print a receipt (Appendix Figure 2). Collectors left the receipt with the taxpayer, with an electronic record saved in the tablet’s memory. When collectors deposited the money, tablet data were automatically downloaded, enabling program supervisors to check that the amount deposited equaled the amounts on all receipts issued.

Collectors memorized the following message during training to solicit the tax from households: “This compound has a legal obligation to pay the property tax for the year of 2016. The provincial government will use the money to promote the economic development of the province. If you do not pay today, please indicate a date and time when you will pay and I will return then.” Collectors kept track of appointments and were told to revisit households until they paid. According to household surveys, the modal number of collector visits in treatment neighborhoods was 2, though 21% of the sample report 3 or more visits.

The treatment is the combination of the property register and tax collection (Table 1). Control neighborhoods experienced neither component. As in the past, citizens in these neighborhoods were expected to pay at the tax ministry themselves. The main analysis considers the reduced-form impact of the tax campaign as a whole. This is a theory- and policy-relevant estimand given

productive collectors earned more than \$10.

that states invariably need information about citizens before they can collect taxes from them (Kleven et al., 2011).

4 Data, estimation, and balance

4.1 Data

Data come from four sources: (1) administrative data on property tax payment, (2) a baseline survey before the campaign, (3) a midline survey during the campaign, and (4) an endline survey after the campaign.

Administrative data come from the government’s official tax database. This database was managed by a private company, Hologram Identification Systems, which integrated raw data from collectors’ tablets with existing bank data. I link official tax records to survey data using the unique household tax identification numbers assigned during property registration.

Baseline survey enumeration occurred just before the property tax campaign. Independent enumerators randomly sampled households following skip patterns while walking down each avenue in a neighborhood: e.g. visit every X^{th} compound, where X is determined by the estimated number of compounds and a target of 5 per neighborhood. Enumerators then conducted midline surveys, on average 2-4 weeks after collectors had completed a neighborhood. For control neighborhoods, enumerators similarly waited at least two weeks after an adjacent neighborhood had received tax collectors. Enumerators conducted a short survey in all compounds, asking whether households were visited by tax collectors and whether they paid the property tax. Finally, enumerators administered the endline survey in 2017, after the tax campaign. In each neighborhood, enumerators first conducted a screening survey of roughly 20 households, randomly sampling again with a skip pattern. I then randomly selected a subsample of screening survey participants for the full interview, choosing higher-quality houses with slightly higher probability to focus on the population most affected by the campaign.¹⁶

¹⁶Appendix Section 2.2 describes this sampling strategy. I also construct weights and re-

Because of insecurity in Kananga in early 2017, enumerators were unable to conduct the endline survey in the commune of Nganza, representing about 15% of the city’s population. All 71 neighborhoods from this commune were dropped before respondents could be sampled and invited to participate. Because of the spatial stratification used for randomization, the number of neighborhoods ineligible for endline enumeration is balanced (Appendix Table 6). During endline, 453 of the 3,421 (13.21%) sampled households could not be surveyed. Common causes included (1) being too busy, (2) being on a trip, and (3) declining participation without a reason. These forms of attrition are also balanced.

Table 1 summarizes the activities of the collectors and the enumerators. All research components of the study — baseline, midline, and endline surveys — were held constant across treatment and control. Sampling and enumeration procedures of these surveys were identical, as indicated by the balanced length of surveys (Appendix Table 6). What varied across treatment groups was assignment to the tax campaign.

4.2 Outcome measurement

The paper examines two sets of outcomes. First, it examines whether the tax campaign increased visits from collectors and raised tax compliance. For this analysis, I use two variables.

1. *Visited by tax collector*: an indicator that the household received visits from provincial tax collectors in 2016, self-reported at midline.
2. *Paid property tax*: an indicator that the household paid the property tax in 2016, measured by linking administrative compliance data to household surveys by unique tax ID numbers. In control, I use fuzzy name matching within neighborhoods to match administrative records with household surveys.¹⁷

Second, testing the tax-participation hypothesis requires measures of polit-

estimate all results to be representative of the population (Appendix Section 4).

¹⁷This method is reliable because there are <30 records of non-campaign property-tax payments in 2016 made by individuals.

ical engagement. I cannot use voting data because the DRC is not a liberal democracy: elections were canceled in 2016 and 2017. Moreover, it is more in line with the underlying theory to test this hypothesis in a nondemocracy with a broken social contract. The theory is that taxation stimulates political participation and ultimately the emergence of inclusive and accountable political institutions (which may include elections). To test this theory, one needs to measure how citizens exert voice in politics *in the absence of such institutions*. Self-reported political participation is often subject to measurement error caused by social desirability, time inconsistency, and/or anonymity concerns in repressive settings. Thus, I worked with the provincial government to embed measurement strategies in two forms of political engagement that come at a cost to individuals: attendance at townhall meetings, and submission of government evaluations.

Specifically, in early 2017, the provincial government held five townhall meetings, chaired by the finance minister and the director general of the tax ministry, to provide a venue for dialog with citizens. Endline participants in treatment and control received official invitations to one of these meetings (Appendix Figure 6).¹⁸ The meetings were advertised as a chance to obtain information about taxation and public spending in Kananga and to ask questions of government officials. The actual proceedings were formal and highly consistent with this description, as I describe further below and in Appendix Section 2.4. Townhall meeting attendance indicates a willingness to exert costly effort to have a voice in the government. Citizens had to remember the date and time of the meeting and pay for their transport to the provincial assembly building, located up to 13km (on average 5km), from participants' homes (Appendix Figure 7). Motorcycle taxis ask up to \$2 for a oneway trip from the outskirts of Kananga to the city center. Nonetheless, 483 individuals (24.9% of those who received invitations) participated in a townhall meeting.

The second measure is the submission of anonymous evaluations of the

¹⁸Militia-related insecurity in Kananga increased in early April, and the government discontinued the meetings, urging all citizens to stay in their homes. Thus, participants sampled after April 1 never received an invitation.

provincial government to a locked drop box across from the vice governor’s official residence in downtown Kananga.¹⁹ All endline participants in treatment and control received evaluation forms after the endline survey. Participants then chose whether or not to fill out and drop off their evaluation. The form had one question about the respondent’s overall level of satisfaction with the government, followed by four agree-disagree statements concerning (i) opportunities for participation, (ii) access to information, (iii) spending on public goods, and (iv) citizen reporting of problems.²⁰ Citizens could also write additional suggestions in a text box at the bottom. They were informed that the governor and other top officials would receive the evaluations plus a summary of their contents. Filling out the form and paying the transport to the drop box downtown again demonstrates willingness to engage in costly participation with the provincial government. In total, 396 individuals (13.6% of total respondents) submitted their evaluations.

For completeness, I consider five dependent variables.

1. *Townhall meeting attendance*: an indicator for individuals who attended a townhall meeting.
2. *Evaluation form submission*: an indicator for individuals who submitted an evaluation.
3. *Townhall or evaluation*: an indicator for individuals who either attended a townhall or submitted an evaluation.
4. *Townhall and evaluation*: an indicator for individuals who both attended a townhall and submitted an evaluation
5. *Costly participation index*: a standardized index composed of indicator variables for townhall attendance and evaluation submission.²¹

Finally, I consider survey evidence about citizen demand for the provincial government to provide public goods. First, respondents answered questions

¹⁹This is similar to the comment forms in Olken (2007) and the postcards in Paler (2013).

²⁰See Appendix Section 2.4 for further details.

²¹I use standardized indices throughout the paper to facilitate interpretation of coefficient magnitude (in terms of standard deviations). I construct these indices by first standardizing each component variable, summing over all questions, and standardizing the new synthetic variable again. I use this indexing procedure whenever there are multiple measures of the same underlying variable.

about whose responsibility it is to provide public goods across six different sectors (such as education and infrastructure), choosing for each among the provincial government and other possible providers (the national government, NGOs, churches, etc). From these data, I use the standardized sum of sector-specific indicators for choosing the provincial government. This variable is thus increasing in the amount of public goods provision demanded from the provincial government relative to other providers. Second, enumerators posed three sets of opposing viewpoints concerning the optimal level of public service provision by the provincial government. These hypothetical questions are combined into an index that is also increasing in the extent to which participants envision a large role for the provincial government in public goods provision. Both indices are examined individually and in an aggregate index.²²

4.3 Estimation

I primarily use OLS to estimate the following equation:

$$y_{ijk} = \beta_1 I_{jk}^{Campaign} + \alpha_k + \mathbf{X}_{ijk}\boldsymbol{\Gamma} + \mathbf{X}_{jk}\boldsymbol{\Phi} + \varepsilon_{ijk} \quad (1)$$

where i indexes individuals, j neighborhoods, and k the strata used during randomization. $I_{jk}^{Campaign}$ is an indicator for neighborhoods that receive the door-to-door tax campaign, meaning that β_1 estimates the average causal effect of the tax campaign on the outcome of interest (y_{ijk}), i.e. political participation. Standard errors are clustered at the neighborhood level (356 in total). In addition, α_k are strata fixed effects, and \mathbf{X}_{ijk} and \mathbf{X}_{jk} are individual- and neighborhood-level covariates. All regressions control for gender, age, and age squared, with additional covariates at times also included, as noted below.

4.4 Balance

To check the randomization, I estimate Equation 1 with thirteen individual-level variables from the endline survey, thirteen neighborhood-level variables

²²See Appendix Section 6 for the exact text of the underlying questions.

from the baseline survey,²³ and six variables about survey enumeration itself (Appendix Table 6). In total, two individual-level covariates, household wealth index and business owner status, are imbalanced at the 10% level, and one neighborhood-level covariate, quality of public lighting, is imbalanced at the 10% level. Thus, as expected, 9.3% of variables are found to be significant at the 10% level. An omnibus test of joint orthogonality fails to reject the null for the individual variables ($F = 1.32, p = 0.20$) and the neighborhood-level variables ($F = 0.71, p = 0.75$). To be conservative, the three imbalanced covariates are included in \mathbf{X}_{ijk} and \mathbf{X}_{jk} , respectively for the main specifications, with robustness checks showing invariance to the specific covariates used in the Appendix.

5 Results

5.1 Effects on tax compliance

This section considers the ‘first stage’ — whether the campaign raised tax compliance through household visits by collectors. It is not obvious that a tax campaign in the DRC would succeed. Bureaucrats are underpaid and have low morale, while citizens have little exposure to formal tax collection. Will collectors undertake this work as planned, and will citizens pay when collectors arrive at their doorstep for the first time?

Table 2 summarizes OLS estimations of Equation 1. The campaign caused an 81.5 percentage-point increase in reported visits from tax collectors.²⁴ It also caused on average an 10.3 percentage-point increase in property tax payment, a 100-fold rise relative to control. For this and subsequent estimations, Appendix Section 4 contains a series of robustness checks, including specifications with (1) only gender, age, and age squared as covariates, (2) all possible

²³Two exceptions are road quality and public lighting, which were measured at endline.

²⁴In control neighborhoods, 5% of individuals reported visits from tax collectors. This likely reflects noncompliance among collectors, who at times crossed into to the wrong (control) neighborhoods. Such noncompliance was expected given that the borders between neighborhoods are not always clearly delimited and must be checked using GPS. This noncompliance would, if anything, bias treatment effects toward zero.

covariates listed in the pre-analysis plan, (3) enumerator fixed effects, (4) sampling weights, and (5) heterogeneous treatment effects by wealth.

Although a 10 percentage-point increase in tax compliance is substantial, the majority of individuals still *evaded* paying the tax, despite visits from collectors. Why did the campaign cause some, but far from all, individuals to pay the tax? A companion paper provides a full treatment of this question (Weigel, 2018). Briefly, tax compliers tended to have more education, income, wealth, and formal employment (Appendix Table 3). Tax collectors also underscored the role of cash on hand in explaining compliance, noting that more people paid at the beginning of the month (just after salaries are paid). In addition, individuals who at baseline perceived a higher probability of punishment for evasion were more likely to pay — as were individuals who *ex ante* professed more positive attitudes toward the provincial government. These results are consistent with classic cost-benefit models of tax compliance (Allingham and Sandmo, 1972) as well as models emphasizing “tax morale” (Luttmer and Singhal, 2014).

It is also worth noting that the tax campaign does not appear to have increased bribes, according to multiple measures (Appendix Section 1.2).²⁵ This result is unsurprising for two reasons. First, because this was the first-ever citizen tax campaign, collectors faced high uncertainty about the government’s plans to audit their work and sanction bribe takers. Second, collusive bribery is more likely when collectors and citizens have repeated interactions (Khan et al., 2015). In contrast, this first year of tax enforcement involved, in most cases, a single-shot interaction between collector and citizen. The negligible impact on bribe payment means that the campaign could only affect political engagement through collector visits and tax payment.

5.2 Effects on political participation

Given that the campaign increased collector visits and tax compliance, I use its random assignment to test the hypothesis that taxation raises demand for

²⁵Reporting bribes is not taboo in Kananga: in a study of tolls in Kananga, nearly half of motorcycle taxi drivers openly admitted to paying bribes (Reid and Weigel, 2017).

political participation. Estimations of Equation 1, summarized in Table 3, support this hypothesis. The campaign triggered a 4.4 percentage-point increase in townhall attendance (Column 1) and a 2.6 percentage-point increase in evaluation form submission (Column 2). To capture the intensive margin, Columns 3 and 4 show that the tax campaign stimulated participation in either outcome by 4.9 percentage points and in both outcomes by 2.8 percentage points. These treatment effects amount to a 0.14 standard-deviation increase in participation (Column 5).

The results are robust to the checks described above, as well as estimating average effect size (AES) coefficients.²⁶ Controlling for the distance between participants' houses and the location of the townhall meeting and the evaluation drop box does not affect the results (Appendix Table 8). As noted in Table 3, constructing p -values using randomization inference or Bonferroni adjustments does not meaningfully affect the statistical significance of the estimates.²⁷

Of the 600 individuals who participated in a townhall meeting or submitted their evaluation, 179 (30%) did both; 128 of these 179 (72%) hailed from treated neighborhoods. The provincial assembly building, where the townhall meetings occurred, and the evaluation drop box were about 1km apart in downtown Kananga (Appendix Figure 7). However, evaluation form submission did not increase on the days of townhall meetings. Most double participants appear to have made independent trips to attend the townhall and to submit their evaluations.

Did individuals participate to make demands of the provincial government, or did they simply have more factual questions about the 2016 tax campaign? Examining the statements made by participants during townhall meetings suggests that citizens used them as an opportunity to bargain with government

²⁶See Appendix Section 4 for robustness checks and Appendix Table 9 for AES coefficients.

²⁷The Bonferroni-adjusted p -value is calculated following Aker et al. (2011) to adjust for correlation between *Townhall meeting attendance* and *Evaluation form submission*. If m is the number of correlated outcome variables and ρ is the average correlation coefficient among the other outcome variables, the Bonferroni p -value with a correlation adjustment equals $1 - (1 - p)^g$, where $g = m^{(1-\rho)}$.

officials over the quality of governance in Kananga. Over 70% of these statements were (i) demands for better governance in exchange for compliance with provincial taxes, or (ii) related questions about public spending, public goods in Kananga, or provincial corruption (Appendix Figure 10). “Why should the inhabitants of Lukonga [a commune of Kananga] pay taxes,” one participant asked, “when the roads are in such disastrous condition?”²⁸ Such complaints evoke a bargaining process in which citizens demand better governance in exchange for tax compliance (Bates and Lien, 1985).

Additional suggestive evidence about the motive behind participation comes from examining whether, conditional on attending a meeting, citizens from treated neighborhoods were more likely to ask about factual details of provincial taxation (21% of total townhall statements). This comparison is difficult to interpret because speaking at the meetings is endogenous to participation. Nonetheless, it reinforces the descriptive evidence above that, according to simple difference-in-means tests, treated individuals were no more likely to ask factual questions about the tax campaign, but they were roughly twice as likely to ask about provincial spending and public goods ($p=0.050$).

Further evidence comes from submitted evaluations, which did not mention taxation but asked about the inclusiveness and transparency of the government. Submitted evaluations were highly critical: over 90% expressed overall disapproval of the provincial government. Similarly, respondents overwhelmingly demanded more avenues of participation, access to information, and public goods spending (Appendix Figure 13). In addition, 39% of individuals wrote in additional suggestions, of which the most frequent topics include: general demands for better governance, demands for specific public goods projects, and demands for greater monitoring of the provincial government and improved transparency.²⁹ “We ask our government to draw its attention especially to Quartier Kapanda, Avenue Lubanza,” wrote one participant, “where we are threatened by erosion, and we note that our government has never built anything to counter erosion in this quarter.”

²⁸Participant question from January 30 townhall meeting (author’s translation).

²⁹Appendix Section 2.4 provides more details about submitted evaluation card contents.

Participants in treatment neighborhoods were particularly likely to use the evaluations to demand better governance. If we re-estimate the specification in Column 2 of Table 3 using an outcome variable indicating submission of (i) evaluations that express *disapproval* of the government, and (ii) evaluations that contain critical *written-in suggestions* or demands,³⁰ individuals in treatment were still more likely to participate compared to control (Appendix Table 7). This evidence reinforces that the treatment effect reflects an increase in demand for inclusive, high-quality governance.

Finally, if the increases in participation reflect greater demand for good governance, individuals in treated neighborhoods should also hold stronger views about the obligation of the provincial government to provide public goods. Regression results examining survey-based indices described on p. 14 confirm this intuition (Table 4). Individuals in treated neighborhoods demand a larger role (by 0.112 standard deviations) for the provincial government in public goods provision relative to other possible providers, such as the national government or NGOs.³¹ Importantly, this result does not just reflect changes in beliefs about the *current* levels of public goods provision. An analogous set of questions asked how much citizens perceive the provincial government to be currently providing in the same sectors. No systematic differences appear across treatment and control (Appendix Figure 27). The evidence in Table 4 therefore suggests that the tax campaign expanded the extent of public goods provision that citizens *demand* of the provincial government.

5.3 Alternative explanations

Rather than demand for better governance, do higher rates of participation in treatment reflect (1) experimenter demand effects, (2) a decline in participation

³⁰This latter variable equals 1 only if the written-in comment was critical or made a demand of the provincial government. Comments that were complimentary of the government (5.5% of total comments) and comments about the Harvard research team (3.3%), militia-related violence (3.9%), or some other topic (4.4%) are coded as 0.

³¹The standard errors are larger when considering sector-based questions rather than hypothetical questions, but the magnitude of the coefficients is nearly identical. Appendix Figure 23 shows results for the constituent survey questions of these indices.

in control rather than an increase in treatment, (3) a sense of unfairness due to awareness of untaxed control neighborhoods, or (4) the short-run salience of taxation in treated neighborhoods? This section explores these possibilities.

5.3.1 Experimenter demand effects

One concern is whether the observed increase in participation is an artifact of the research components of the experiment. Treated citizens might have been more likely to participate if they had more contact with or were treated differently by enumerators, became more trusting of the research team, and thus felt more emboldened to participate as a result.

To preclude such issues, all research procedures were held constant across treatment and control, as evidenced by the balance in measurable characteristics of survey enumeration (Appendix Table 6). Moreover, enumerators administered surveys in a random order, frequently alternating between control and treatment neighborhoods. Individuals in treatment and control received the same information about the townhall meetings and government evaluations, and participation always occurred *after* endline survey enumeration to minimize potential demand effects.

To test formally for different levels of trust or familiarity with the research team, we consider survey questions asking respondents (1) how much they trust foreign research organizations, (2) whether they know the employer of the enumerator, (3) whether they participated in surveys in the past, (4) whether they did not provide a phone number to the enumerator (indicating potential mistrust of the researchers), and (5) whether they provided an incorrect or fake phone number to the enumerator (also indicative of mistrust). No systematic differences appear across treatment and control (Appendix Table 10).

An indirect demand effect could arise if tax collectors encouraged citizens to participate. However, this is implausible because the townhalls and evaluations had not yet been scheduled or announced at the time of tax collection.

5.3.2 Declining participation in control

A second alternative explanation is that the treatment effects result not from higher participation in treatment but from lower participation in control. It is possible that control individuals expected visits from tax collectors, and when they never received them, they concluded that the government was less capable than they previously thought — and hence decided to participate less.

I investigate this hypothesis using a sample of 630 baseline participants whom enumerators re-surveyed after the tax campaign.³² Although I cannot measure changes in participation, I examine changes in beliefs about the provincial government within individuals over time, specifically: (1) the responsibility of the provincial government in public goods provision (the same sector-based question examined in Table 4); (2) trust in the provincial government and tax ministry; and (3) the share of taxes that respondents perceive to be spent well and not wasted or stolen. Appendix Table 11 summarizes fixed-effects regressions with an indicator (*Post*) for measurement after the tax campaign, interacted with the treatment indicator. If attitudes towards the government deteriorated in control, there would be negative point estimates on *Post*. For none of these measures is the coefficient negative and statistically different from zero.³³ At least for this set of individuals tracked from baseline to endline, those in control not seem to have updated negatively about the government.

5.3.3 Awareness of the untaxed control

Treated individuals might have participated more because they were aware that control neighborhoods had not yet been taxed, and they thought this was unfair. The main result could thus be an experimental artifact, a function of having measured outcomes before the control group received the tax campaign.

³²I collected these data for a companion paper on the determinants of compliance (Weigel, 2018). This repeated baseline sample is not part of the endline sample for this paper, but it is helpful here to examine changes in beliefs within individuals.

³³The increase in the perceived responsibility of the government to provide public goods in the treatment group (Column 1) corroborates the results in Table 4.

At first glance, this explanation appears implausible because households were informed that the campaign would eventually reach all neighborhoods. Still, treated individuals could have thought it unfair that they were taxed first.

To explore this possibility, I examine whether treated households near the border with control neighborhoods were more likely to participate compared to households farther from control. If awareness of untaxed control fueled participation in treatment, then presumably individuals living near a border with control (and thus more aware of neighborhood-level differences in tax collection) would have been more likely to participate compared to those further from the border. However, plotting the participation rate in treatment as a function of minimum distance to control reveals no such relationship (Appendix Figure 15). Moreover, complaints about the fact that some neighborhoods had been taxed while others had not did not arise during townhall meetings or on government evaluations. Awareness of the randomized rollout of the campaign appears to have been low. This is not surprising because the unit of randomization, the neighborhood, was quite fine, averaging only 131 plots. If larger regions of the city had been taxed before others, citizens might have been more likely to notice the phased rollout.

5.3.4 Short-term salience of taxation

Another interpretation of the results is that the increase in participation is driven by the short-term salience of taxation in treated neighborhoods. In this interpretation, the treatment is akin to a prime, and individuals are thought to participate more in townhall meetings simply because taxation is top of mind, not because they have higher demand for public goods or good governance.

Although many public programs may function in part through salience effects, this interpretation is difficult to sustain in the present context. First, there was on average a 6-8 month gap between tax collection and the forms of participation measured as outcomes. Salience and priming effects are unlikely to persist this long. Second, although townhall invitations did mention taxation as a subject of the meetings, the evaluation forms made no mention of taxation. The treatment effects on this outcome — and the survey outcomes in

Table 4 — thus could not plausibly stem from the salience of taxation. Finally, the idea that citizens respond to taxation by demanding better governance *in exchange for taxation* is in fact the theory this paper seeks to test. It is this process of “tax bargaining” that is thought to trigger the evolution of more inclusive and responsive governance (Bates and Lien, 1985). The exchanges between citizens and government during townhall meetings are thus consistent with the tax-participation hypothesis.

5.4 Further analyses

5.4.1 Crowd out of local participation

This paper tests the hypothesis that citizens respond to taxation by trying to hold accountable the government that taxes them, i.e. the provincial government in this case. But could an increase in participation with the provincial government crowd out participation at other levels of government?

Although I lack measures of costly participation at other levels of government, I provide suggestive evidence from survey questions about engagement at the national and local level. Specifically, respondents indicated their current and future participation in national elections, parties, marches, protests, and rallies, which I combine in the index *Engagement with national politics*. A separate index, *Interest in politics*, combines questions about political news consumption and knowledge, a chance to learn information about the national government, and hypothetical questions about the role of citizens in politics. To measure local engagement, the survey asked about views of and engagement with city chiefs, local notables with two main responsibilities: (1) organizing weekly *salongo*, an informal tax in which citizens contribute labor toward local public goods, such as maintaining neighborhood roads (Olken and Singhal, 2011); (2) mediating local disputes to avoid escalation to the formal court system. I combine all questions about city chiefs in the index *Engagement with local city chiefs*.

There are no detectable differences in national political engagement or interest in politics across treatment and control (Appendix Table 12). Column

3, however, suggests that the tax campaign crowded out participation at the local level. Treated individuals report fewer consultations with city chiefs as well as diminished views of their quality (Appendix Figure 26). The effect is more pronounced among relatively poorer individuals. Although only suggestive, this result has an intuitive interpretation. City chiefs are more active in poorer, peripheral neighborhoods, where the formal state is essentially absent. As the state expands its presence by collecting taxes, citizens may substitute engagement with the provincial government for engagement with local chiefs. Formal taxation may crowd out informal labor taxation like *salongo*. This result supports the view that building the state can undermine local, informal forms of governance (Cheema et al., 2006).

5.4.2 Spillovers

Because of the cluster-level randomization, I can estimate spillover effects of the tax campaign following Miguel and Kremer (2004). Specifically, I exploit random variation in (a) the number of treated neighborhoods adjacent to control neighborhoods (controlling for the number of total adjacent neighborhoods), and (b) the length of control neighborhoods' borders shared with treatment neighborhoods (controlling for the total length of each control neighborhood's borders).

There is evidence of spillovers in reported visits from tax collectors in control (Appendix Table 13). This is not surprising due to the lack of clear on-the-ground markers between some neighborhoods. However, there is not an accompanying increase in tax compliance in control, perhaps because the spillover effect on visits is small in magnitude. There is also no statistically significant spillover on participation. However, the estimates are consistently positive and of non-trivial magnitude, so it is possible I am underpowered to detect an effect. In fact, the presence of externalities on participation but not payment would be consistent with the mechanism proposed in Section 6.2, through which the tax campaign sent a signal of state capacity that raised the expected benefits of participation.³⁴

³⁴Such spillovers would bias the treatment effect toward zero.

5.4.3 Persistence of the treatment effect

Will the increase in participation persist over time? As noted, the average time gap between tax collection and participation is 6-8 months, so the main estimates already demonstrate persistence. Moreover, I can exploit the variation in this time gap to estimate decay formally. This variation is random because the order in which both collectors and enumerators worked in neighborhoods were random.

Appendix Figure 16 shows the estimated treatment effect after taking quartiles of the data according to the lag between tax collection and participation. The treatment effect becomes smaller over time, but this decay is only marginally significant. Specifically, the difference in the treatment effect between periods 1 and 2, and between periods 1 and 3, is not significant. But the change in magnitude from period 1 to period 4 is marginally significant ($t = 1.79$). Thus, although there appears to be a slight decline in the impact of the tax campaign over time, the degree of persistence is perhaps the more surprising implication given that a tumultuous period of political uncertainty (including the cancellation of the 2016 national election) and civil conflict occurred between tax collection and participation.

6 Mechanisms

This section examines three possible mechanisms behind the increase in participation caused by the tax campaign: (1) individual *payers* in treated neighborhoods participated at higher rates because they expected reciprocal benefits or derived greater expressive utility from voicing their grievances; (2) the tax campaign sent a signal of state capacity that raised the expected benefits of participation; (3) the tax campaign lowered the coordination costs of participation by stimulating common grievances and communication. Although the evidence in this section is more suggestive, it supports the second mechanism.

6.1 Tax payment as the cause of participation

Many accounts of the political economy effects of tax collection assume that *payers* are those who participate more and demand better governance. Tax payment could stimulate a sense of ownership over public revenues, leading taxpayers to expect public goods and better governance as a *quid pro quo* (Prichard, 2015). Alternatively, tax payment could trigger participation through a behavioral response akin to an endowment effect or a version of the sunk cost fallacy (Martin, 2014).

A naive test of these mechanisms is to examine whether payers participated more than nonpayers in treated neighborhoods.³⁵ Although payment is an endogenous outcome of treatment, this correlation can still be informative, especially in the case that payment and participation are uncorrelated. Given that the likely unobserved sources of bias (income, education, views of the government, etc) in a regression of participation on payment would bias the coefficient on payment away from zero, estimating a zero correlation coefficient would be difficult to reconcile with a payment-based mechanism. Interestingly, payers were no more likely to participate compared to nonpayers in treatment neighborhoods (Appendix Table 14).

Similarly, one can compare participation among individuals in treated neighborhoods who did and did not receive visits from tax collectors. Although collectors were supposed to visit all households in a neighborhood, they sometimes skipped households likely due to idiosyncratic human error.³⁶ A mech-

³⁵Comparing payers to non-payers in the full sample would be harder to interpret because it would compare compliers in treatment to a mix of never-takers in treatment plus compliers and never-takers in control. (I assume away the existence of always-takers since payment in control is effectively zero.) A less complicated comparison is compliers to never-takers in treatment neighborhoods only, shown in Appendix Table 14.

³⁶I suspect that collectors typically skipped households by accident, due to the fact that neighborhoods bear little resemblance to a grid, and it is easy to lose track of one's position in the neighborhood, even when guided by a GPS device. Moreover, collectors received a piece-rate wage for documenting each house in the property register, so they had little incentive to skip houses. To reinforce this interpretation, I compare the coordinates taken during the collectors' property registration survey to coordinates from the midline survey, conducted by enumerators with greater experience using GPS devices, to identify unvisited properties. This analysis does not reveal a pattern consistent with collectors deliberately skipping certain houses, i.e. those that are larger and might pay bigger bribes (Appendix

anism operating through payment would imply no difference in participation between nonpayers who were and were not visited. On the other hand, a signaling or coordination mechanism would predict differences (as discussed in the next section). As shown in Appendix Table 14, there is a significant positive association between participation and tax collector visits, but not between participation and tax payment. These correlations suggest that the mechanism operates through the experience of collector visits rather than tax payment per se.

A more rigorous test of payment-based mechanisms requires an instrument for endogenous tax payment. Assignment to treatment is an obvious candidate, but the exclusion restriction would be violated given that a first-ever door-to-door tax campaign likely has other direct effects on participation since it conveys information about the government. I therefore need instruments for two endogenous regressors — collector visits only ($I^{Visited\ only}$) and collector visits plus payment ($I^{Visited\ and\ paid}$) — to identify the causal effect of paying taxes on participation separate from other informational effects of the campaign captured by $I^{Visited\ only}$:

$$y_{ijk} = \beta_1 I_{ijk}^{Visited\ only} + \beta_2 I_{ijk}^{Visited\ and\ paid} + \alpha_k + \mathbf{X}_{ijk}\mathbf{\Gamma} + \mathbf{X}_{jk}\mathbf{\Phi} + \varepsilon_{ijk} \quad (2)$$

A common pitfall of IV analysis with multiple endogenous variables is reliance on instruments that identify the same endogenous regressor, leaving the other regressor unidentified (even if the joint first-stage F -stat is large). Fortunately, one can construct F -stats for each endogenous variable independently, thereby verifying that both regressors are separately identified by the instruments (Angrist and Pischke, 2008, pp. 217-218). In the tables that follow, these F -stats will be reported as “AP F -stat.”

I thus construct leave-one-out jackknife IV (JIVE) instruments for $I^{Visited\ only}$ and $I^{Visited\ and\ paid}$, respectively. These JIVE instruments exploit the random

Figure 19). Instead, unvisited houses appear idiosyncratically distributed, indicative of human error. To be conservative, I also instrument for visits.

assignment of tax collectors to neighborhoods. The intuition behind these instruments is that a collector’s effort in a given neighborhood can be predicted by his or her observed effort in all other assigned neighborhoods. The instruments are constructed as follows.

1. Predict a fixed effect, $\hat{\lambda}_{i,-j}$, for collector i in neighborhood j by estimating Equation 1 with tax collector dummies and the endogenous variable as the outcome in all assigned neighborhoods other than j .
2. Take a linear combination of the collector-specific fixed effects to construct a neighborhood-level instrument, i.e.

$$Payment\ propensity = \sum_{i=1}^3 \delta_i * \hat{\lambda}_{i,-j}$$

where δ_i weights the collector-specific fixed effects.³⁷

I construct JIVE instruments for both endogenous variables: *Visit propensity* for $I^{Visited\ only}$, and *Payment propensity* for $I^{Visited\ and\ paid}$. The logic of this strategy is that collectors vary in their effort and effectiveness, and the two traits are not perfectly correlated. Some collectors make many visits (high effort) but collect few taxes (low effectiveness). Others make fewer visits but are more skilled at convincing citizens to pay taxes. Appendix Figure 17 plots, for each tax collector who worked on the campaign, the correlation between visits and payment, conditional on household covariates and stratum fixed effects. Despite the fact that collecting tax payments was impossible without visiting households (and thus the correlation cannot be negative), for only 38% of collectors is the correlation coefficient statistically different from zero. This considerable variation in the observed effort and effectiveness of tax collectors is reassuring for this estimation approach.

The JIVE instruments can be thought of as a continuous predictor of treatment intensity along these two dimensions (effort and effectiveness): they equal 0 for control neighborhoods, and then vary between 0 and 1 for treated neigh-

³⁷For simplicity, collectors are weighted evenly, though due to sick days and other factors some worked for more days than others.

neighborhoods depending on the predicted effort (or effectiveness) of the assigned collectors. Some neighborhoods are randomly assigned to a set of collectors likely to exert high effort; others are assigned to collectors likely to demonstrate high effectiveness. If these qualities are sufficiently uncorrelated, there should be a first stage for $I^{Visited\ only}$ and for $I^{Visited\ and\ paid}$.

Table 5 reveals that both instruments predict $I^{Visited\ only}$, and *Payment propensity* instrument strongly predicts $I^{Visited\ and\ paid}$. Although the endogenous regressors will be jointly identified by the full set of instruments in two-stage least squares, the fact that only *Payment propensity* predicts $I^{Visited\ and\ paid}$ is reassuring that there are indeed valid instruments in both of the first stage equations. The F -stat reported here is the standard joint test of the exogenous instruments; the individual Angrist-Pischke (AP) F -stats for 2SLS with multiple endogenous variables are reported in Table 6 showing the second-stage results. Including enumerator fixed effects, as in all robustness checks (Appendix Section 4), further strengthens the first stage.

This estimation generates little evidence that the increase in participation goes through tax payment. Although standard errors are large, the estimated coefficient on $I^{Visited\ and\ paid}$ is always negative and statistically indistinguishable from zero. On the other hand, the coefficient on $I^{Visited\ only}$ is consistently positive and often statistically significant (especially when I include enumerator fixed effects). Its magnitude is 2-3 times as large as the reduced-form effect of the campaign on participation. Ultimately, the large standard errors make this analysis only suggestive. But the most natural interpretation is that tax payment does not appear to have an effect on participation separate from the effect of being visited by tax collectors. Indeed, we can never reject equivalence of the coefficients on $I^{Visited\ and\ paid}$ and $I^{Visited\ only}$. In sum, the available evidence is consistent with a mechanism in which the campaign increased participation through collector visits rather than through tax payment.

6.2 Tax collection as a signal of state capacity

One such mechanism is as follows: a first-ever tax campaign stimulates participation by sending a signal of state capacity that raises the expected benefits of participation. The intuition is that citizens who observe the campaign update that the government has greater means and is more capable than previously thought. They therefore anticipate greater returns to engaging with members of the government, for example by advocating for more public spending in their neighborhood. This mechanism predicts higher participation among everyone in treatment neighborhoods — payers and nonpayers.

6.2.1 Decision-theoretical framework

Imagine there is a government and one citizen who is uncertain about the capacity of the government. The government sets a policy $g(\theta, \lambda)$, where $\theta \in \{H, L\}$ indicates whether the government is high or low capacity, and $\lambda \in \{1, 0\}$ indicates the citizen's decision to monitor the government. The citizen incurs a cost c to participate, and receives utility $u(g(\theta, \lambda))$ from the policy.

Government capacity (θ) is meant generally. It could be 'extractive capacity,' i.e. ability to collect taxes, or 'productive capacity,' i.e. ability to provide public goods and enforce contracts (Besley and Persson, 2009). A signal of either type of capacity triggers participation because the citizen believes the government will be more likely to affect his future well being — through tax collection or public goods provision — and thus he has an incentive to try to influence public policy to be as favorable as possible.

Concretely, the government can provide public goods, which increase the citizen's utility, and extract taxes, which decrease the citizen's utility. The citizen's preferred policy (high public goods, low taxes) results when the government is high capacity and when the citizen participates. To simplify notation, call this policy g^+ . When the government is low capacity, the government always provides the same policy (low public goods, low taxes) regardless of citizen participation: $g(L, 1) = g(L, 0)$. In this case, the citizen has no incentive to participate. Call this policy g^0 . When the government is high capacity and

the citizen does not participate, however, the policy is worse for the citizen than g^0 because the government collects taxes without providing public goods. Call this least-preferred (by the citizen) policy g^- . To summarize:

$$u(g^+) \geq u(g^0) \geq u(g^-) \quad (3)$$

Before the tax campaign, the citizen believes the government is high capacity with probability $p \sim F(\cdot)$. If he participates, his expected utility is:

$$EU_1 = p(u(g^+) - c) + (1 - p)(u(g^0) - c) \quad (4)$$

If he doesn't participate, his expected utility is:

$$EU_0 = p(u(g^-)) + (1 - p)(u(g^0)) \quad (5)$$

The citizen chooses the action that maximizes expected utility. There is a threshold point p^* of indifference between participating and not participating:

$$p^* = \frac{c}{u(g^+) - u(g^-)} \quad (6)$$

In this expression, the quantity $(u(g^+) - u(g^-))$ is the participation dividend, which we might term d . The derivative with respect to d is negative:

$$\frac{\partial p^*}{\partial d} = -\frac{c}{d^2} < 0 \quad (7)$$

Thus, as the participation dividend increases, citizens can be less confident that the government is high capacity but still choose to participate.

Now assume that the government launches a tax campaign, which sends a signal about its capacity (θ). The citizen knows that a high-capacity government administers a tax campaign with probability α , and a low-capacity government administers a tax campaign with probability β . Then as long as $\alpha \geq \beta$, by Bayes' Theorem, the posterior probability (q) that the government is high capacity conditional on administering a tax campaign is:

$$\frac{\alpha p}{\alpha p + \beta(1 - p)} = q \geq p \tag{8}$$

Let $F(\cdot)$ be a uniform distribution, i.e. $p \sim U(0, 1)$, and $\alpha = 0.8$ and $\beta = 0.4$. We can then simulate the distribution of q , as shown in Figure 1. A threshold (p^*) is shown in red at a value of 0.7. Individuals with values of p to the right of this threshold participate; those to the left do not. There is more mass to the right of the threshold in the posterior distribution, indicating that individuals with priors to the left of the threshold have shifted in their beliefs to the right, choosing to participate only after receiving the signal sent by the tax campaign. Thus, the tax campaign catalyzes citizen engagement with the state by conveying information about the capacity of the state.

This very simple framework suits weak-state settings, such as the DRC, in which the government is effectively absent *ex ante* and thus a citywide tax campaign plausibly shocks citizens' beliefs about its capacity. The framework does not likely suit high-capacity states in which citizens are habituated to taxation. In such settings, an increase in tax enforcement may have an ambiguous effect on participation: some citizens might choose to protest new taxes, while others might invest in strategies for evasion. That said, a low-tax, low-capacity equilibrium characterizes the settings in early modern Europe that scholars draw on to develop tax-participation hypothesis (Schumpeter, 1918; Tilly, 1985; North and Weingast, 1989).³⁸

6.2.2 Evidence

One implication of this mechanism is that the treatment effect should be larger in neighborhoods with less past exposure to the state. Where the state has been effectively absent, receiving a visit from government agents conducting a property register and collecting taxes should send a stronger signal of capacity compared to neighborhoods habituated to the state. Thus, in neighborhoods unaccustomed to the state, more individuals should update their beliefs beyond

³⁸Tilly (1985) notes that European monarchies had low capacity prior to expanding taxation. He argues that state capacity was a byproduct of the quest to raise tax revenue.

the threshold and choose to participate.

To test this hypothesis, I measure past state exposure on the neighborhood level in two ways: (1) the number of past visits to the neighborhood from state agents reported at baseline; (2) the number of individuals who report ever having participated in a political protest at baseline. The former measure captures state activity in the neighborhood, while the latter captures respondents' exposure to the state outside of the neighborhood. I use neighborhood estimates of these variables and split the sample at the median. The treatment effect is indeed larger in neighborhoods with *less* past state exposure. An F -test rejects the equivalence between the effects in low- and high-exposure neighborhoods. These results are consistent with the idea that citizens update their beliefs more — and thus participate more — when they are less accustomed to the presence of the state.

As a second test, I show that citizens' self-reported beliefs about the capacity of the government shift in response to the tax campaign. I estimate q using survey data. Following Besley and Persson (2009), I examine both extractive capacity, the government's ability to raise tax revenue, and productive capacity, its ability to enforce contracts and provide public goods. As Besley (2018) notes, a government can raise revenues through coercion or by fostering voluntary compliance. I thus split extractive capacity into a coercive and a voluntary component.

1. **Extractive capacity - coercive compliance.** Coercion requires (i) information about taxpayers, and (ii) a credible threat of punishment for evasion. As measures, I thus use two survey-based indices.³⁹
 - (a) *Information about citizens*: increasing in how much information the government is perceived to possess about citizens (e.g. household location, compliance status, occupation, income).
 - (b) *Ability to punish evaders*: increasing in the perceived likelihood of punishment against households that refuse to pay the property tax or pay a bribe instead.
2. **Extractive capacity - voluntary compliance.** Voluntary compliance

³⁹See Appendix Section 6 for details on all variables.

requires citizen approval of the tax ministry and confidence that its collectors will not simply pocket taxpayer money. I measure this with the following indices.

- (a) *Performance of tax ministry*: increasing in citizens' overall trust in and approval of the provincial tax ministry.
- (b) *Integrity of tax collectors*: increasing in the perceived amount of money collected in property taxes that will reach state coffers.
- (c) *Perceived citizen compliance*: increasing in the share of other households whom respondents think paid the property tax in 2016. This approximates the state's revealed extractive capacity.

3. **Productive capacity.** Once the state has resources, it needs capacity to deploy those resources productively rather than wasting or stealing them. Productive capacity is thus a function of the technology of public goods provision as well as the ability to control high-level corruption and spend prudently. I examine the following survey-based measures.

- (a) *Ability to provide public goods*: increasing in the perceived ability (i.e. technology) of the provincial government to provide public goods (electricity, paved roads, security) efficiently and effectively, assuming it has the will to do so.
- (b) *Performance of government*: increasing in citizens' trust in and approval of the provincial government in general.
- (c) *Share of taxes spent well*: increasing in the perceived share of tax revenues that will be spent on public services or other 'good uses' and not lost to high-level corruption or misallocation.

Table 8 summarizes estimations of Equation 1 using each of these variables as the outcome. Concerning the coercive component of extractive capacity (Panel I), the tax campaign increased citizens' perceptions about how much information the government possesses about citizens, especially the locations of their properties and their tax compliance status. But it did not substantially impact beliefs about the credibility of punishment for evasion (though I may be underpowered to detect a small effect). These inferences drawn by treated citizens are essentially correct. Thanks to the campaign, the government does

have a new database with detailed information about potential taxpayers that it can use to collect more tax in the future. Moreover, to my knowledge, the government did not pursue sanctions against many noncompliant households. It seems to have used its enforcement capacity to pursue arrears for a handful of firms accused of evading other taxes and fees.

Regarding the voluntary component of extractive capacity, treated citizens updated their beliefs considerably. At endline, they viewed the tax ministry more positively and had more confidence in its collectors, thinking a greater share of taxpayer money would be deposited in the state account (rather than staying in collectors' pockets). It might seem odd that citizens updated *positively* about the government tax apparatus after it started taxing them. However, in a 'failed state' in which government agents are seldom observed doing meaningful work, the fact that a citywide tax campaign involving the use of tablets and receipt printers would send a positive signal is not surprising.

Importantly, citizens in treated neighborhoods also perceived much higher levels of citizen compliance with the property tax. In other words, they updated about the de facto extractive capacity of the government. The mixed results on coercive compliance paired with consistently positive updating about voluntary compliance suggests that citizens expected others to pay taxes chiefly due to the enhanced legitimacy of tax collectors (i.e. the belief that more of their tax money would reach state coffers). Again, citizens were correct in believing that quasi-voluntary factors like legitimacy and trust motivated payment in this setting (Weigel, 2018).

It is possible that this pattern of belief changes simply reflects the fact that payers convinced themselves that tax collectors were trustworthy after they paid, an example of ex post motivated reasoning. Because there were more payers in treatment, such motivated reasoning could explain the average effect. However, re-estimating Table 8 with only nonpayers returns similar results (Appendix Table 15), albeit with slightly smaller coefficients. Nonpayers clearly also drew the same inferences as a result of the tax campaign, making a motivated reasoning interpretation unlikely. It might at first appear counter-intuitive that households that *evaded* the tax would update positively — not

negatively — about the state’s capacity. But, again, this must be put in the context of a nearly absent state. In such a setting, receiving home visits from tax collectors, being assigned a tax ID number and entered into a government database, and being asked to contribute to formal taxation is a stronger signal about the government than is the fact that this year they managed not to pay.

Panel III explores whether treated citizens think the greater extractive capacity of the government will lead to higher public goods provision. There is little evidence that the tax campaign increased perceptions of the *technology* of public goods provision. Not surprisingly, observing tax collectors did not lead citizens to think the government could now build a road more efficiently. Treated citizens also did not evaluate the government as a whole more positively, as they did the tax ministry. However, they did update about the *share* of tax revenues that would go to public goods spending or other good uses. This result mirrors the higher confidence among treated citizens that collectors would deposit tax money to the state. Updating about the share of revenues that would be well spent implies that, conditional on the same public goods provision technology, treated citizens did perceive the government to have greater productive capacity after the tax campaign.

In sum, the tax campaign caused citizens to believe that (i) the government had more revenue due to greater extractive capacity, and (ii) that it would spend more of that revenue productively. These results are consistent with a mechanism by which the tax campaign signals capacity and raises citizens’ expectations about the benefits of participation. They are also consistent with evidence from Brazil showing that citizens are more successful in holding the government accountable when revenues come from taxes rather than (unobserved) transfers (Brollo et al., 2013; Gadenne, 2017).

To provide further suggestive evidence about the importance of beliefs about government capacity in citizens’ decisions to participate, I examine the beliefs of participators and non-participators in treatment neighborhoods (Appendix Table 16). Although these comparisons are not identified, they can nonetheless help to interpret the average effects on beliefs in Table 8. Participators appear to have updated their beliefs in line with the average with three exceptions.

First, they were more likely to believe the state would punish tax evasion compared to non-participants. Second, participants in treatment were less confident than non-participants that tax collectors would not simply pocket the money they collected. Similarly, they were also less confident that tax revenues would fund public goods and not be wasted or embezzled.

Thus, while participants in treatment also updated about state's extractive capacity and total revenues, they were more concerned about the *uses* of tax money compared to non-participants. This pattern of correlations has an intuitive interpretation. Observing the tax campaign caused citizens to update about the size of the budget and the potential for public spending, creating an incentive for participation. This incentive was offset for some by confidence that the government would spend the money in a productive manner (i.e. provide g^+) even without citizen monitoring. Only those who were less confident about government spending of tax revenue chose to participate in order to monitor and influence spending toward their preferred policy.

This interpretation is reinforced by the fact that participants made specific demands that the government repair roads and counter erosion (p. 19) in their townhall and evaluation comments. Moreover, a number of citizens explicitly demanded transparency and accountability regarding the new revenues raised during the property tax campaign. "The provincial government should do more," wrote one individual, "and inform us how *this money* will be spent on public infrastructure and not wasted on other things." Another individual wrote: "I ask that the government show the population what it achieves with *this money*" (emphasis added). These individuals sought to monitor spending of the tax campaign revenues, consistent with the proposed mechanism.

A different interpretation is that citizens participated to try to access patronage goods rather than to try to influence public goods provision. Although g encompasses any public and private goods that the government distributes, several pieces of evidence make a patronage story less likely. First, although citizens might have expected handouts at townhall meetings (there were none), it is hard to imagine they would have expected patronage goods to result from submitting evaluations because the forms were anonymous and deposited in a

drop box that did not involve interaction with government officials. Second, although citizens made neighborhood-specific demands during townhall meetings, they did not make overt requests for fully individualistic benefits. Third, when asked how the money would be spent, most people guessed roads (49%) or education (19%), while only 11% said waste/leakage (Appendix Figure 18). That said, this interpretation of greater *patronage capacity* is consistent with the theoretical framework, and I cannot rule it out entirely.

6.3 Taxation lowers the cost of coordination

Individuals might have anticipated being more effective in lobbying for public spending in their specific neighborhood if multiple residents attended townhall meetings, or submitted evaluations making similar demands. If the tax campaign lowered the costs of coordination by stimulating common grievances and communication about the government, taxation, and public services, a coordination mechanism could explain the increase in participation.

In the Appendix, I consider several tests of this mechanism (Section 5). First, I examine if treated townhall participants were more likely to arrive at townhall meetings with other members of the neighborhood. There is marginally significant evidence that treated individuals may have coordinated more with others in the neighborhood to travel to townhalls together. Second, I use the GPS coordinates of participants' households to measure if individuals who attended a townhall or submitted an evaluation are more clustered geographically within treatment neighborhoods relative to control, as one would expect if lower coordination costs were the key mechanism. I find no evidence of greater clustering in treatment. Finally, I use baseline data to test if the campaign had larger effects in neighborhoods with higher collective action potential, proxied by the baseline level of ethnic homogeneity, population density, and city chief activity. There is no evidence of such heterogeneous effects.

In sum, although there is some suggestive evidence that the tax campaign could have stimulated coordination among citizens to attend townhall meetings, it is unlikely that lowering the cost of collective action is the principal

mechanism explaining the reduced-form increase in participation.

7 Conclusion

This paper analyzed the first door-to-door property tax collection campaign in the city of Kananga, D.R. Congo, which increased tax compliance by 10 percentage points. It used the random assignment of the campaign to test the tax-participation hypothesis, finding that citizens in taxed neighborhoods were nearly 5 percentage points more likely to attend a townhall meeting or to submit a government evaluation. Participating individuals demanded more public goods and more accountability from the government. The available evidence supports a mechanism through which tax collection sent a signal of state capacity that raised the expected benefits of participation.

One implication of this study is that the political response to tax collection may be disproportionately large relative to the increase in tax revenues. Even if the state only succeeds in marginally raising revenues, its efforts to do so may trigger a large increase in citizen participation due to the signal of state capacity sent by an increase in tax collection/enforcement. However, the political response may also be larger for the property tax than for indirect and less visible taxes, such as consumption taxes that pass through to consumers in the form of higher prices. These observations could thus explain why developing countries collect so little revenue in property taxes, despite the theoretical advantages of property taxation: forward-looking governments may anticipate a large political response from taxing property and choose other tax instruments instead.

References

Acemoglu, Daron and James A Robinson, “The Emergence of Weak, Despotic and Inclusive States,” *Working Paper*, 2017.

- Aker, Jenny C, Rachid Boumnijel, Amanda McClelland, and Niall Tierney**, “Zap it to me: The short-term impacts of a mobile cash transfer program,” *Working Paper*, 2011.
- Allingham, Michael G and Agnar Sandmo**, “Income tax evasion: A theoretical analysis,” *Journal of Public Economics*, 1972, 1 (3-4), 323–338.
- Angrist, Joshua D and Jörn-Steffen Pischke**, *Mostly harmless econometrics: An empiricist’s companion*, Princeton university press, 2008.
- Bates, Robert H and Da-Hsiang Donald Lien**, “A note on taxation, development, and representative government,” *Politics & Society*, 1985, 14 (1), 53–70.
- Besley, Timothy**, “State Capacity, Reciprocity and the Social Contract,” *Working Paper*, 2018.
- **and Torsten Persson**, “The origins of state capacity: Property rights, taxation and policy,” *The American Economic Review*, 2009, 99 (4), 1218–1244.
- **and –**, “Taxation and development,” *Working Paper*, 2013.
- **and –**, “Why do developing countries tax so little?,” *The Journal of Economic Perspectives*, 2014, 28 (4), 99–120.
- Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem**, “Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan,” *Journal of Political Economy*, 2015, 123 (6), 1311–1355.
- Brewer, John**, *The sinews of power: War, money, and the English state, 1688-1783*, Harvard University Press, 1990.
- Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini**, “The political resource curse,” *The American Economic Review*, 2013, 103 (5), 1759–1796.
- Carpio, Lucia Del**, “Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru,” *Working Paper*, 2013.
- Carrillo, Paul, Dina Pomeranz, and Monica Singhal**, “Dodging the taxman: Firm misreporting and limits to tax enforcement,” *American Economic Journal: Applied Economics*, 2017, 9 (2), 144–164.

- Casaburi, Lorenzo and Ugo Troiano**, “Ghost-house busters: The electoral response to a large anti-tax evasion program,” *The Quarterly Journal of Economics*, 2015, 131 (1), 273–314.
- Casey, Katherine, Rachel Glennerster, and Edward Miguel**, “Reshaping institutions: Evidence on aid impacts using a preanalysis plan,” *The Quarterly Journal of Economics*, 2012, 127 (4), 1755–1812.
- Cheema, Ali, Asim Ijaz Khwaja, and Adnan Qadir**, “Local government reforms in Pakistan: context, content and causes,” *Decentralization and local governance in developing countries*, 2006, pp. 257–84.
- Coate, Stephen and Stephen Morris**, “On the form of transfers to special interests,” *Journal of Political Economy*, 1995, 103 (6), 1210–1235.
- de la Cuesta, Brandon, Helen V Milner, Daniel Nielson, and Steve Knack**, “No greater representation with taxation: Experimental evidence from Ghana and Uganda,” *Working Paper*, 2015, pp. 23–24.
- Dunning, Thad, Felipe Monestier, Rafael Piñeiro, Fernando Rosenblatt, and Guadalupe Tuñón**, “Is Paying Taxes Habit Forming? Theory and Evidence from Uruguay,” *Working Paper*, 2015.
- Fjeldstad, Odd-Helge, Merima Ali, and Tom Goodfellow**, “Taxing the urban boom: property taxation in Africa,” *CMI Policy Paper*, 2017.
- Gadenne, Lucie**, “Tax me, but spend wisely? Sources of public finance and government accountability,” *American Economic Journal: Applied Economics*, 2017, 9 (1), 274–314.
- Gordon, Roger and Wei Li**, “Tax structures in developing countries: Many puzzles and a possible explanation,” *Journal of Public Economics*, 2009, 93 (7), 855–866.
- Jensen, Anders**, “Employment Structure and the Rise of the Modern Tax system,” *Working Paper*, 2018.
- Khan, Adnan Q, Asim I Khwaja, and Benjamin A Olken**, “Tax farming redux: Experimental evidence on performance pay for tax collectors,” *The Quarterly Journal of Economics*, 2015, 131 (1), 219–271.
- Kleven, Henrik J and Mazhar Waseem**, “Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan,” *The Quarterly Journal of Economics*, 2013, 128 (2), 669–723.

- Kleven, Henrik Jacobsen, Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez**, “Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark,” *Econometrica*, 2011, 79 (3), 651–692.
- Levi, Margaret**, *Of rule and revenue*, University of California Press, 1989.
- Lizzeri, Alessandro and Nicola Persico**, “Why did the elites extend the suffrage? Democracy and the scope of government,” *The Quarterly Journal of Economics*, 2004, 119 (2), 707–765.
- Lowes, Sara, Nathan Nunn, James A Robinson, and Jonathan L Weigel**, “The evolution of culture and institutions: Evidence from the Kuba Kingdom,” *Econometrica*, 2017, 85 (4), 1065–1091.
- Luttmer, Erzo F P and Monica Singhal**, “Tax Morale,” *Journal of Economic Perspectives*, 2014, 28 (4), 149–168.
- Martin, Lucy**, “Taxation, loss aversion, and accountability: theory and experimental evidence for taxation’s effect on citizen behavior,” *Working Paper*, 2014.
- Miguel, Edward and Michael Kremer**, “Worms: identifying impacts on education and health in the presence of treatment externalities,” *Econometrica*, 2004, 72 (1), 159–217.
- Naritomi, Joana**, “Consumers as Tax Auditors,” *Working Paper*, 2015.
- North, Douglass C and Barry R Weingast**, “Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England,” *The Journal of Economic History*, 1989, 49 (4), 803–832.
- Olken, Benjamin A.**, “Monitoring Corruption: Evidence from a Field Experiment in Indonesia,” *Journal of Political Economy*, 2007, 115 (2), 200–249.
- Olken, Benjamin A and Monica Singhal**, “Informal taxation,” *American Economic Journal: Applied Economics*, 2011, 3 (4), 1–28.
- Paler, Laura**, “Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government,” *American Political Science Review*, 2013, 107 (4), 706–725.
- Pomeranz, Dina**, “No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax,” *American Economic Review*, 2015, 105 (8), 2539–2569.

- Prichard, Wilson**, *Taxation, responsiveness and accountability in Sub-Saharan Africa: the dynamics of tax bargaining*, Cambridge University Press, 2015.
- Reid, Otis and Jonathan Weigel**, “Citizen Participation in Corruption: Evidence from Roadway Tolls in the Democratic Republic of the Congo,” *Working Paper*, 2017.
- Sanchez de la Sierra, Raul**, “On the origin of states: Stationary bandits and taxation in Eastern Congo,” *Journal of Political Economy*, 2019.
- Schumpeter, Joseph A**, “The crisis of the tax state,” *International Economic Papers*, 1918, 4, 5–38.
- Tiebout, Charles M**, “A pure theory of local expenditures,” *Journal of Political Economy*, 1956, 64 (5), 416–424.
- Tilly, Charles**, “War making and state making as organized crime,” *Violence: A reader*, 1985, pp. 35–60.
- Weigel, Jonathan L**, “The taxman cometh: A virtuous cycle of compliance and state legitimacy in the D.R. Congo,” *Working Paper*, 2018.
- Wille, Christophe**, “Carving up Congo: President Kabila’s Latest Ploy to Stay in Power,” *Foreign Affairs*, 2015.

8 Tables and figures

Table 1: Activities of collectors and enumerators

Activity	Treated	Control	Timing	N	J
Tax collectors					
Property register	Yes	No	Apr-Dec 2016	20,902	253
Tax collection	Yes	No	Apr-Dec 2016	20,902	253
Enumerators					
Baseline survey	Yes	Yes	Mar-Apr 2016	2,384	431
Midline survey	Yes	Yes	Apr-Dec 2016	33,019	431
Endline survey	Yes	Yes	Jan-May 2017	2,913	356

N = sample size, J = number of clusters.

Table 2: Effects of the campaign on collector visits and compliance

	Visited by tax collector (1)	Paid property tax (2)
Campaign	0.815*** (0.013)	0.103*** (0.007)
Stratum FE	Yes	Yes
R^2	0.640	0.054
Observations	27,443	27,443
Clusters	356	356
Control Mean	0.050	0.001

Visited by tax collectors is an indicator for households reporting at least one visit by tax collectors in 2016. *Paid property tax* is an indicator for individuals' who paid the property tax in 2016 according to the administrative data. See p. 12 for details on these variables. Data: midline survey merged with government tax database.

Table 3: Effects of the campaign on participation

	Townhall meeting attendance (1)	Evaluation form submission (2)	Townhall or evaluation (3)	Townhall and evaluation (4)	Costly participation index (5)
Campaign	0.044** (0.020)	0.026** (0.012)	0.049*** (0.016)	0.028*** (0.009)	0.144*** (0.043)
Covariates	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes
R^2	0.062	0.055	0.067	0.038	0.070
Observations	1934	2912	2913	2913	2913
Clusters	252	356	356	356	356
Control Mean	.18	.1	.18	.035	-.057
Dep. Var.	0-1	0-1	0-1	0-1	Standardized
Rand. Inf. p	0.034	0.045	0.0050	0.0040	0.0012
Bonferroni p	0.042	0.052			

Townhall attendance is an indicator variable that equals 1 if a participant attended a townhall meeting. *Evaluation form submission* is an indicator variable that equals 1 if a participant submitted his or her evaluation. *Townhall or evaluation* indicates that a participant attended either a townhall meeting or submitted an evaluation form. *Townhall and evaluation* indicates that a participant attended a townhall meeting and submitted an evaluation form. *Costly participation index* is a standardized index of *Townhall attendance* and *Evaluation form submission*. See p. 14 for details on these variables. Data: endline survey merged with townhall attendance and submitted evaluation records. The sample size is smaller in Column 1 because the government discontinued townhalls after April 1 due to insecurity in Kananga. Endline respondents sampled after this date never had a chance to attend a meeting.

Table 4: Effects of the campaign on the perceived responsibility of the provincial government to provide public goods in Kananga

	Responsibility of the provincial government in public goods provision		
	(full index)	(sector-based)	(hypotheticals)
	(1)	(2)	(3)
Campaign	0.112** (0.052)	0.088 (0.053)	0.088** (0.041)
Covariates	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes
R^2	0.041	0.043	0.030
Observations	2913	2813	2900
Clusters	356	356	356
Control Mean	-0.066	-0.051	-0.053

All outcomes are standardized indices increasing in the perception that the provincial government should be the primary provider of public goods in Kananga. The outcome in Column 1 is an aggregate index. The outcome in Column 2 is based on sector-specific questions about the government's responsibility relative to other possible providers (national government, NGOs, etc). The outcome in Column 3 is based on hypothetical survey questions about the role of the provincial government in service provision. See p. 14 for details on these variables. Data: endline survey.

Table 5: IV - First stage: Predicting visits and tax payment using randomly assigned tax collector effort and effectiveness

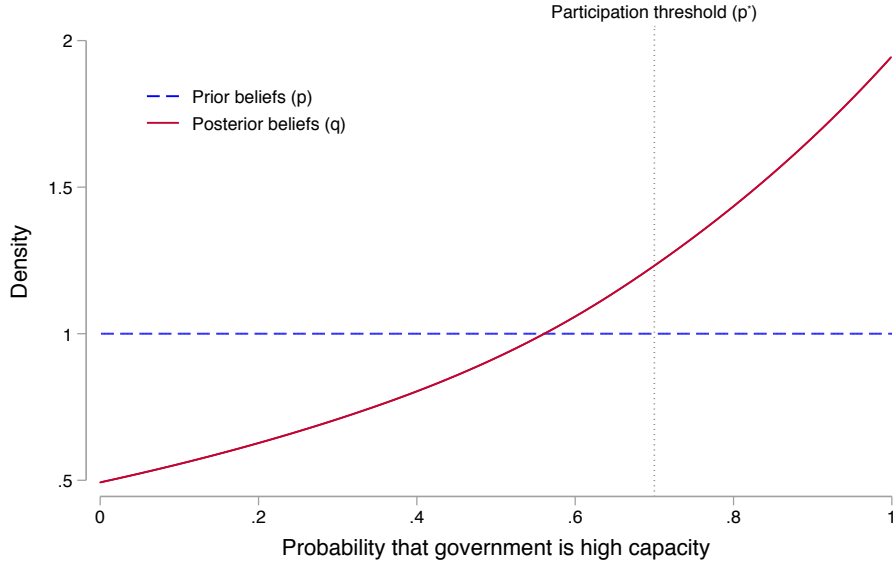
	Visited only		Visited and paid	
	(1)	(2)	(3)	(4)
Payment propensity (JIVE)	0.767*** (0.076)	0.709*** 0.073	0.267*** (0.041)	0.287*** (0.040)
Visit propensity (JIVE)	0.361*** (0.127)	0.443*** (0.119)	-0.116* (0.064)	-0.155** (0.061)
Covariates	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes
Enum FE	No	Yes	No	Yes
R^2	0.214	0.239	0.083	0.099
Observations	2913	2913	2913	2913
Clusters	356	356	356	356
Dep. Var. Mean	0.487	0.487	0.065	0.065
F -stat	147.861	144.754	34.337	41.278

Visited only is an indicator for household visited by tax collectors that did not pay the property tax. *Visited and paid* is an indicator for households who were visited and paid the property tax. *Payment propensity (JIVE)* is a leave-one-out estimator that uses randomly assigned tax collectors' observed payment rates in other neighborhoods to predict the payment rate in a given neighborhood. *Visit propensity (JIVE)* is a leave-one-out estimator that uses randomly assigned tax collectors' observed visit rates in other neighborhoods to predict the visit rate in a given neighborhood. See p. 29 for details about these instruments. Data: endline survey merged with government tax database.

Table 6: IV - Second stage: Distinguishing the effects of collector visits and tax payment on participation

	Townhall meeting attendance		Evaluation form submission		Townhall or evaluation		Townhall and evaluation		Costly participation index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Visited only	0.171*	0.173**	0.038	0.055	0.136	0.158**	0.106*	0.101**	0.454*	0.486**
	(0.094)	(0.078)	(0.075)	(0.057)	(0.099)	(0.076)	(0.060)	(0.045)	(0.262)	(0.191)
Visited and paid	-0.368	-0.264	-0.007	-0.059	-0.209	-0.236	-0.265	-0.189	-0.889	-0.800
	(0.360)	(0.268)	(0.317)	(0.230)	(0.421)	(0.300)	(0.269)	(0.191)	(1.155)	(0.803)
Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Enum FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1934	1934	2912	2912	2913	2913	2913	2913	2913	2913
Clusters	252	252	356	356	356	356	356	356	356	356
AP F -stat (Visited only)	40.978	63.355	42.721	73.352	42.743	73.402	42.743	73.402	42.743	73.402
AP F -stat (Visited and paid)	11.824	18.060	9.164	16.479	9.163	16.484	9.163	16.484	9.163	16.484
F -test p (equivalence)	0.225	0.190	0.907	0.686	0.504	0.286	0.254	0.212	0.338	0.187

The outcomes are identical to those in Table 3. As in Table 5, *Visited only* and *Visited and paid* indicate households that received visits from tax collectors but did not and did pay, respectively. AP F -stats report the endogenous regressor-specific Angrist-Pischke F -statistic for 2SLS with multiple endogenous regressors (see p. 28). F -test (equivalence) reports the p value for tests for equivalence of the coefficients on *Visited only* and *Visited and paid*. Data: endline survey merged with government tax database as well as townhall attendance and submitted evaluation records. The sample size is smaller in Columns 1-2 for the same reason noted in Table 3: endline respondents sampled after April 1 did not have an opportunity to attend a townhall meeting because government discontinued these meetings due to insecurity in Kananga.



Simulations assuming a uniform distribution of prior beliefs, $\alpha = 0.8$, and $\beta = 0.4$.

Figure 1: Simulated distributions of prior and posterior beliefs about government capacity.

Table 7: Heterogeneous effects of the campaign on participation by past exposure to the formal state

	Townhall or evaluation		
	(1)	(2)	(3)
Campaign	0.049*** (0.016)	0.062*** (0.023)	0.082*** (0.021)
Campaign X Past visits (high)		-0.030 (0.033)	
Past visits (high)		0.038 (0.027)	
Campaign X Past protest (high)			-0.073** (0.034)
Past protest (high)			0.035 (0.022)
Covariates	Yes	Yes	Yes
Stratum FE	Yes	Yes	Yes
R^2	0.067	0.068	0.069
Observations	2913	2913	2913
Clusters	356	356	356
Control Mean	.18	.18	.18
F-test p-value		.01	.00049

The outcome is the same as Column 3 in Table 3. *Past visits (high)* indicates neighborhoods above the median level of past visits from government agents reported during baseline. *Past protest (high)* indicates neighborhoods above the median level of past citizen participation in protests reported during baseline. See p. 34 for further details about these variables. Data: endline survey merged with participation records and neighborhood-level measures from baseline survey.

Table 8: Effects of the campaign on citizens' beliefs about the extractive and productive capacity of the provincial government

<i>Dependent variable</i>	β	SE	R^2	N	μ_c
<i>Panel I: Extractive capacity - coercive compliance</i>					
Information about citizens	0.152***	0.044	0.085	2910	-0.080
Ability to punish evaders	0.048	0.048	0.044	2883	-0.017
<i>Panel II: Extractive capacity - voluntary compliance</i>					
Performance of tax ministry	0.122***	0.047	0.065	2791	-0.076
Integrity of tax collectors	0.188***	0.044	0.043	2732	-0.119
Perceived citizen compliance	0.348***	0.052	0.102	1954	-0.179
<i>Panel III: Productive capacity</i>					
Ability to provide public goods	-0.012	0.053	0.038	2484	0.009
Performance of government	0.045	0.049	0.042	2795	-0.030
Share of taxes spent well	0.108**	0.050	0.054	2766	-0.062

Each row summarizes an OLS estimation of Equation 1, with the dependent variable noted in the first column. β is the coefficient on the treatment indicator, followed by the cluster-robust standard error, R^2 , number of observations, and control group mean. There are 356 clusters. Each dependent variable, described briefly on p. 34 and in detail in Appendix Section 6, is standardized to facilitate interpretation of coefficient magnitude. Data: endline survey. The number of observations varies across regressions due to non-response for specific survey questions.