Balancing Purse and Peace:

Tax Collection, Public Goods and Protests

Benjamin Krause*

Download Latest Version

November 19, 2020

Abstract

Strengthening state capacity in low income countries requires raising tax revenue while maintaining political stability. The risk of inciting political unrest when attempting to increase taxes may trap governments in a low-tax equilibrium, but public goods provision may improve both tax compliance and political stability. To test these questions empirically, I partner with the national tax authority and a local mayor's office in Haiti to cross-randomize both tax collection and public goods across one of the country's largest cities. Effects are measured both via administrative data on tax revenue as well as through novel measures of political unrest. In the paper's main result, I show that hand-delivering property tax invoices reduces individual tax compliance by 48%, and increases independently observed measures of localized political violence by 192%. In contrast, providing a valuable and visible public good (namely municipal garbage removal) increases tax compliance by 27%, and reduces localized political violence by 85%. Importantly, public goods provision significantly mitigates the adverse effects of tax collection in neighborhoods receiving both treatments. A cost accounting exercise suggests that providing the public good in this setting could pay for itself within the first year. These findings suggest that it may be possible to peacefully shift to a new equilibrium of higher tax compliance with a sufficient initial investment perhaps financed through foreign aid or other transfers.

*University of California, Berkeley Agricultural and Resource Economics; benjaminkrause@berkeley.edu; see here for my UC Berkeley Job Market Candidate Page as well as related work at benjaminkrause.com.
Previous versions of the paper have been presented under the title, "Taxation Toward Representation".

Acknowledgments

This work has been done in partnership with the Mayor's Office of Carrefour under the leadership of Jude Edward Pierre, the national tax authority of Haiti (*Direction Générale des Impôts* (DGI)) with the collaboration of Serge Rosenthal, Jr., as well as SOLUTIONS, SA led by Kurt Jean Charles. They and their respective teams work each day in service to their fellow citizens, and it is to them and their efforts that this research is dedicated.

Principal funding for implementation for this research was provided by the Mayor of Carrefour and SOLUTIONS, SA and funding for research has been provided by the Weiss Fund for Research in Development Economics, J-PAL Governance Initiative, CEGA Development Economics Challenge, and the UC Berkeley ARE Department.

The implementation and research has been made possible by the tireless field management of Etienne Frederic with assistance from Jean Dana Darius and "Jah" Laurore Adam with oversight from Gary Philoctete and financial accountability from Marie Nadege Josaphat. I also want to thank all of the people who worked to help collect data during this project including: Berlineda Râteau, Bien-Aime Steeve Kenley, Blanc Darline, Blonde Daphta Guerrier, Casimir Mazeau, Dorvil Ronelle Laurenzia, Etienne Jimmytri, Gedeon Bijou, Guerrier Woodberline, Herolna Pierre, Inoncent Guirlene, James Germain, Jefferson Frédéric, Judith Laguerre, Junior Lundi, Lafleur Gaspard-Flis, Larue Arold, Larue Ronise, Laurenville Panel, Laurore Adam Gandhi, Marlene Frédéric, Simon Patiouko, and Tham Ludger.

I am grateful to my dissertation co-chairs, Jeremy Magruder and Edward Miguel, my orals chair, Aprajit Mahajan, and the rest of my orals committee: Ben Faber, Maximilian Auffhammer, and David Zilberman. Furthermore, I would like to recognize Fred Finan, Paul Gertler, Alain de Janvry, Supreet Kaur, Raul Sanchez de la Sierra, Elisabeth Sadoulet, and Dmitry Taubinsky for their feedback and guidance. Finally I thank participants in the UC Berkeley Development Lunch Seminar, UC Berkeley ARE Development Workshop, 2018 GARESC Conference, 2019 WEAI Conference, 2019-2020 BELS Fellows, and UC Berkeley ARE 2015 Cohort Seminar as well as all of the members of my cohort in the ARE department for their collaboration and support.

All errors are my own.

1 Introduction

> *Little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism, but* ***peace****, easy* ***taxes****, and a tolerable administration of justice.* —Adam Smith

> *It's* ***not in our habit to pay taxes****. Haitians haven't been educated to pay taxes ...* ***Who here can throw the first stone****?* —(now former) Prime Minister Jean-Henry Céant

State capacity has been a central focus of the literature on economic development (Fukuyama, 2004; Janvry and Dethier, 2012; Thomas, 2015), and strengthening it requires effective taxation (Pomeranz and Vila-Belda, 2019). Two levers supported by the literature and available to policy makers to increase tax compliance are tax collection effort itself and public goods provision (Besley and Persson, 2009). Theory suggests that political instability and the threat of violence in response to efforts to increase taxes may trap governments in a low-tax equilibrium (Besley and Persson, 2013). Other work has predicted that public goods provision may improve political stability (Acemoglu, 2005; Besley and Persson, 2011). This paper tests the effects of both tax collection and public goods, as well as their interaction, on both tax compliance and political unrest by cross-randomizing these interventions at-scale and for multiple years across one of the largest cities in Haiti.

Though the empirical literature has explored some aspects of these relationships separately (Khan et al., 2015, 2019; Pomeranz and Vila-Belda, 2019), challenges of implementation, setting and measurement have limited their full experimental examination. This paper overcomes implementation and setting issues by combining international assistance and multiple overlapping partnerships with government offices in one of the world's lowest-capacity states[1] to coordinate interventions over several years. The interventions were implemented by the mayor of this city and the national tax authority with additional financial support from the national government and in-kind donations from a sister city in the Caribbean. The design is a stratified cluster cross–randomization of the interventions over all census blocs of the urban city. This includes a population of more than half–a–million people though interventions primarily focused on the owners of the approximately 50,000 buildings in the tax registry. This paper overcomes the measurement challenges by first identifying, accessing and matching previously siloed and uncoordinated data sets to measure tax compliance in this low-capacity setting. Then to independently measure political unrest and violence, this paper leverages findings from qualitative data collection to introduce two sets of novel metrics. To measure political speech, I census and geo-locate all graffiti in the city and estimate effects on localized political graffiti. To measure the more destructive and violent forms of unrest common in Haiti, I track and estimate effects on the localized construction of protest barricades and roadblocks which are often also lit on fire to magnify the signal of political mobilization in this context.

To motivate and rationalize the empirical analysis, I provide a simple model of the

[1] Consistently near the bottom, in the years leading up to this study, 2014-2017, Haiti ranked in the bottom 1% of the World Bank Governance Indicators for Government Effectiveness https://databank.worldbank.org/source/worldwide-governance-indicators/preview/on

optimal tax evasion decision in a setting of low-compliance. The standard assumption in the theory literature is that increasing tax collection effort with its associated threat of financial penalties unambiguously lowers evasion and increases tax payments (Allingham and Sandmo, 1972). Empirical evidence has consistently supported this assumption (Slemrod, 2019). Instead, building off of theory developed for low-compliance settings (Frey, 1997) and informed by my qualitative and baseline data, in my model tax collection may potentially crowd-out non-pecuniary incentives which are determined in part by the level of public goods (Besley, 2020). As a result, if public goods are absent and the credibility of the state's threats are low, my model's first prediction is that increasing tax collection effort reduces tax payments. As a result of my modeling the non-pecuniary incentive as dependent on the level of public goods, my model's second prediction is that increasing public goods mitigates the crowding-out effect and increases tax payments.

To test the effects of tax collection, I randomize efforts to increase compliance for the principal source of funding for local governments in this setting, property taxes. The mayor's agents hand-deliver invoices for property taxes to those randomly assigned to receive such visits. In settings with much higher levels of baseline tax compliance, interventions like this one have been shown to have statistically and economically significant positive impacts on tax payments (Pomeranz, 2015; Castro and Scartascini, 2015; Holz et al., 2020). In contrast, I find taxpayers reduce compliance by 48%. This negative effect can similarly be shown across nearly all metrics and is consistent with my model's predictions of enforcement crowding-out non-pecuniary incentives in the absence of public goods or credible penalties.

The literature has shown tax collection increases political engagement. There is experimental evidence of tax collection increasing attendance in local government meetings and submissions to city comment boxes (Weigel, 2020) as well as suggestive evidence of it increasing demands for representation (Ross, 2004) and willingness to punish political officials (Martin, 2013). Building on these findings, I find tax collection significantly increases political unrest and destructive or violent action. Treated households report nearly tripling time spent protesting. I further find increases in independently observed measures of unrest: I find political graffiti doubles in treated areas and becomes more negative in tone. I also find that the localized construction of protest barricades increases by 192%.

To test the effects of public goods provision, I work with the mayor's office to randomize a highly salient service which is also the only public good broadly provided by local government in this setting, door-to-door garbage removal. Suggestive evidence from heterogeneity analysis has indicated that public goods provision may increase tax payments (Castro and Scartascini, 2015). At baseline in my setting, respondents noted both that the lack of public services was their primary justification for not paying property taxes and that increasing public services would be the best way to motivate their compliance. When asked which service should be the government's top priority, a plurality of respondents cited garbage removal. Consistent with these findings and my model's predictions, I provide experimental evidence of public goods increasing tax compliance by 27%. I also find evidence of the effect increasing with the intensity of treatment. Using a reduced form estimate leveraging the quasi-random variation in number of visits by a garbage removal team generated by mainte-

nance needs, gang hijackings and staffing conflicts leading to missed routes on an otherwise randomized schedule, as well as an IV estimate, I find an additional visit by the garbage removal vehicle causes a 1% and 1.3% increase in tax payments respectively.

Experiments randomizing relatively small increases in public goods have found suggestive evidence of increased support for public officials (Blattman et al., 2017). As demonstrated by an incentive-compatible willingness-to-pay experiment at baseline as well as additional survey responses, my public goods intervention is highly valuable and visible. I find evidence that it reduces self-reported protesting by 50% and localized political unrest and violence in the form of barricade construction by 85%. I find evidence that for barricade construction as well as for political graffiti, public goods mitigate the tax collection effects in areas receiving both. I find that the remaining political graffiti in these areas is more positive in tone. Both political graffiti and barricade construction are also found to be affected by increases in intensity of treatment with the former increasing by 2% over the mean and the latter by 3.5% over the mean for an additional visit by a garbage removal vehicle.

I test for spillover effects that might influence the interpretation of my principal findings on tax compliance. Leveraging both randomization at the census bloc-level and as well as at the property-level, I compare the responses of control properties co-located in blocs with other properties receiving the tax collection treatment to those in blocs randomly assigned to not have any properties visited by tax agents. I find evidence supporting the sign, magnitude, and significance of my estimated tax collection effects. Using a novel placebo treatment through which instead of delivering and discussing a tax invoice, government agents deliver a letter with the same personal identification and property information on it but no mention of taxes, I find evidence that visits by government officials alone do not decrease tax payments. This suggests that the negative tax collection effect is the results of the explicit discussion of taxes during the government agent's visit. I leverage geographic variation to test for spillovers from properties not receiving public goods by first limiting analysis to properties farther away from bloc boundaries and find evidence that my preferred specification may be a underestimate. I then analyze the responses only among blocs not assigned to receive the public goods intervention and compare the responses of properties in blocs adjacent to public goods treatment blocs to those that do not border public goods treatment blocs. The results provide evidence against large negative responses from properties farther removed from the public goods treatment and are consistent with my preferred specification being an underestimate.

To further investigate mechanisms, I use a separate novel treatment aimed at increasing the social exposure of tax compliance. The effect of one's tax compliance decision becoming known to one's neighbors has been shown to increase tax payments through a pro-social signaling mechanism (Slemrod et al., 2019). However, this work has been done in settings with relatively high levels of initial compliance. In contrast, psychology research of settings with low rates of pro-social coordination has identified a mechanism known as antisocial punishment which contributes to the stability of these low-compliance equilibria (Herrmann et al., 2008). I incorporate social incentives into my model (Shimeles et al., 2017). However, informed by focus group discussions and baseline survey responses, unlike previous work, I

instead assume these incentives encourage non-compliance. The resulting third prediction of my model is that increasing the probability of social exposure of tax compliance reduces tax payments. I then test and find evidence supporting this prediction by randomly assigning half of all tax invoices and resulting tax invoice deliveries to inform the taxpayer that if they pay their taxes, the mayor will place a sign on the exterior of their house acknowledging tax compliance. I find that taxpayers who receive this message respond by further reducing tax payments. This effect is concentrated in blocs randomly assigned to only receive the tax collection treatment. In blocs randomly assigned to also receive public goods, the effect is mitigated, and the evidence suggests it may even be reversed. These findings suggest that in the absence of public goods, antisocial punishment contributes to the stability of the low-tax equilibrium, and that the pro-social signaling mechanism previously identified in the literature may be motivated by receiving public goods.

I indirectly test two additional mechanisms identified in my model by randomizing the framing of a subset of the tax collection treatment with two additional "nudges". Following previous experiments (Hernandez et al., 2017), I first make salient financial penalties of non-compliance. I find these appeals can mitigate the adverse tax collection effect in the absence of public goods. Second, I test culturally-specific appeals to obligation and shared responsibility communicated through local proverbs. I find making obligations salient in this way has a net positive effect, increasing tax payments by more than 40% in areas without public goods and suggestive evidence that effects may be even larger when combined with public goods.

Before concluding, I review an accounting exercise to assess the costs and benefits for policy. My findings suggest that the tax collection intervention making obligations salient as well as the public goods intervention could both plausibly generate positive net revenue for the government. Even the significant initial capital costs for garbage removal could be repaid from revenue generated - possibly even in the first year. Nevertheless, these setting are by definition poorly resourced, and thus international institutions or better-funded central governments may be required to make these initial investments possible. Lastly, evidence from this paper indicates that public goods reduce destructive unrest and may also induce pro-social behavior. This suggests that even if not resulting in positive net revenue at first, if these effects over time decrease losses or increase the rate of return, investments in public goods may still yield a positive return for the government. In short, this paper presents experimental evidence that public goods provision provides a possible pathway for governments to maintain political stability while transitioning to a higher-tax equilibrium – thus balancing purse and peace.

In what follows, I first discuss the most relevant literature and my contributions in 2. Then in 3, I present the model motivating this research. Section 4 explores the context of Haiti and introduces my study city, Carrefour. In 5, I describe my methodology and research design. Section 6 presents results followed by robustness checks in 7 and the exploration of mechanisms in 8. I discuss an accounting exercise of the costs and benefits for the government in 9, and I conclude in 10.

2 Literature Review

Understanding how to increase tax revenue, especially in low-tax equilibrium settings, has been identified in the literature as both key to state capacity building and an open area of empirical inquiry (Besley and Persson, 2013). Standard models of tax compliance based on expected utility maximization of evasion given a perceived probability of detection (Becker, 1974; Allingham and Sandmo, 1972; Srinivasan, 1973) predict much lower levels of tax compliance than observed in the empirical data – especially in developed countries. More recent models incorporating public good provision, social norms, moral sentiments, or other 'tax morale' motivations have demonstrated greater ability to more accurately predict compliance rates (Erard and Feinstein, 1994; Torgler, 2005; Kirchler, 2007; Luttmer and Singhal, 2014; Bernheim and Taubinsky, 2018).

Empirical work has successfully increased tax compliance primarily by focusing on tax collection directly. A recent review article concluded that the "incontrovertible" finding of the empirical public literature is that "raising the chance of getting caught high enough deters evasion" (Slemrod, 2019). Field experiments implementing behavioral nudges to induce higher tax participation have shown significant returns (Hallsworth, 2014; Pomeranz, 2015; Pomeranz and Vila-Belda, 2019) even in middle and lower-income countries for both firms (Shimeles et al., 2017; Carrillo et al., 2017; Holz et al., 2020) as well as for individuals (Castro and Scartascini, 2015; Kettle et al., 2016; Hernandez et al., 2017). These effects have been shown to hold even in the few low-compliance settings where work has been done to date (Dwenger et al., 2016; Weigel, 2020; Bergeron et al., 2020). There is also work examining the relative efficiency of different tax policies which are more accessible to lower capacity states which have similarly been successful in increasing compliance through tax collection effort (Khan et al., 2015; Best et al., 2016; Khan et al., 2019).

The modern democratic state arose – according to this branch of research – from a negotiated bargain of formal representation in exchange for taxes (Levi, 1988; North and Weingast, 1989; Tilly, 1990). Some have even made explicit causal claims that either the level of taxation (Buchanan, 1980; Huntington, 1991) or the ratio of taxation to services provided (Bates and Lien, 1985) leads to an increased demand for representation. Cross-country analyses (Besley and Persson, 2013; Ross, 2004), experiences in weak states with extractive resources (Morrison, 2005), and results from a lab–in–the–field (Martin, 2013) have provided some empirical support to these theories.

However very little has been done to experimentally test these relationships. Related work has shown significant political and electoral effects resulting from cash transfers (Manacorda et al., 2011), windfalls from the central government (Brollo et al., 2013), and interventions to curb tax evasion (Casaburi and Troiano, 2016). More recent experiments have found evidence that increased public service provision may improve public opinion of local leaders (Blattman et al., 2017; Khan et al., 2020), and increased tax collection efforts can increase interactions with the local government (Weigel, 2020). How interventions impact democratic accountability, and in particular the engagement of the poor, remain open empirical questions (Pande, 2020).

Existing experimental work on these topics have generally been limited to settings of relatively high pre–existing compliance within relatively established states. The papers that have attempted to randomize tax collection effort in low-compliance settings have focused primarily on relatively small taxes as a percentage of total income or wealth. In addition, the papers that have randomized public goods have focused on interventions that have gone largely unnoticed by the treated population. I am not aware of any previous work that has randomized both types of interventions in the same study. In this paper, I randomize both collection effort for a financially significant tax that is the principal source of funding for the local government, hand-delivering invoices for personal property taxes, as well as public goods provision of a highly salient service that is effectively the only public good broadly provided by the local government, door-to-door garbage removal. I do this in a setting of extremely low tax compliance and limited state capacity.

This is significant because, with regards to tax compliance, theory and even some empirical work in other disciplines suggest that the incentives faced in low-compliance, low-capacity states may be quite different than in higher-compliance settings. For instance, theory by Frey (1997) cautions that increasing enforcement of government policies in settings that initially have few external incentives may crowd-out the 'civic virtue' motivating baseline levels of pro-social behavior and result in a net decrease in the desired compliance. I am able to test for this by randomizing tax collection effort for a financially significant amount of money in just such a setting. Besley (2020) separately models the related concept of 'tax morale' that is generally regarded as incentivizing greater compliance to instead be dependent upon the level of public goods provided so that in low-capacity states with little-to-no public goods, the incentive further depresses tax compliance. By randomizing public goods starting from a baseline of effectively zero, I am able to test this theory as well. Finally, recent work in economics has highlighted the important role of perceived social norms in decision making (Bursztyn et al., 2017). Empirical work following a natural experiment in a high-compliance setting has shown that appealing to social norms by increasing social exposure of tax compliance – that is, making one's tax compliance decision known to one's neighbors – increases tax payments (Slemrod et al., 2019). Informed by focus group and baseline survey responses, I build on the work of Herrmann et al. (2008) in psychology, and introduce the mechanism of antisocial punishment to this setting. The intuition is that when a group of people perceive that the social norm is one of non-compliance, then a desire to avoid being perceived by the others as contravening that norm further erodes existing levels of compliance. I model, test and find evidence of this mechanism by introducing a novel treatment that experimentally increases social exposure of tax compliance.

Finally, with regards to effects on political unrest, though of significant interest, this sort of behavior has been very difficult to independently measure (Cantoni et al., 2017). To overcome this challenge, I introduce two novel metrics for independently measuring political unrest. First, to measure political speech, I conduct a census of and geo-tag the graffiti across the city. I then use the presence, prevalence, and tone of political graffiti specifically as outcomes of interest. Second, to measure the most violent or destructive political unrest, I track the construction of barricades in neighborhoods which are built, and often lit on fire, as a form of protest in this setting. Tracking both where these are constructed and which

areas are affected provide additional outcomes of interest. As a result, I am able to provide novel experimental evidence of the effects of both tax collection and public goods on political unrest – and on violent or destructive unrest in particular.

3 Model

Background and intuition for the set-up are provided in 3.1. The model is formally presented with Equation 1, the First Order Conditions in Equation 2 and the comparative statics and resulting predictions for my primary research questions regarding tax compliance in 3.2. Additional predictions for proposed mechanisms are presented in A.

3.1 Model Motivation and Set-up

To motivate this research, I model the taxpayer's decision to evade taxes. I enrich the microfoundations developed for the dynamic model of state capacity building in Besley (2020) by modifying the 'civic-minded' incentive so that under the appropriate circumstances the net effects can be crowded-out by enforcement as in Frey (1997). I also incorporate a sense of obligation like that of Reckers et al. (1994) which helps to motivate compliance in a low-capacity state wherein the threat of punishment is not credible even when caught evading. Finally, to account for the prominent role of social norms in such settings, I incorporate a social incentive building off of Shimeles et al. (2017).

In my model, the taxpayer chooses a level of evasion, $e \in [0, 1]$ to maximize the monetary value of the sum of public goods received plus net income:

$$\max_{e} \quad g \quad + \quad w\Big(1 \; - \; (1 \; - \; e)\Big(t \; - \; \Big(e\rho \; + \; \tau(g \; - \; wt)\Big) \; - \; q\Sigma\Big) \; - \tau f e\Big) \qquad (1)$$

where $g \geq 0$ are the public goods received and w is the taxpayer's given level of private income. The actual tax rate set by the government is $t \in [0, 1]$. However the weight of the tax rate, in other words its perceived cost for the taxpayer, is adjusted based on three non-pecuniary incentives.

First, a sense of obligation, $\rho \in [0, 1]$, effectively decreases the weight of the tax rate as the taxpayer feels compelled to make this payment out of duty. This sentiment increases in the level of evasion chosen, e, as she feels increasingly worse for contravening her sense of duty.

Second, the perceived fairness of the tax, $(g - wt)$, is determined by subtracting the total tax bill, wt, from the amount of public goods received, g. When the taxpayer receives in public goods a monetary value equivalent for her of exactly what she is being asked to pay,

then $g - wt = 0$ and the motivation has no effect on the perceived tax as she is indifferent. However, when she receives more than she is being asked to pay, then $g - wt > 0$ and this motivation decreases the weight of the tax rate as she perceives it to be generous - acting as complement to tax obligation in this way. Conversely, when the taxpayer receives less in public goods than she is asked to pay in taxes, $g - wt < 0$, then this motivation increases the weight of the tax as she perceives this tax to be unfair - acting against the tax obligation effect. This incentive is multiplied by the level of effort the government expends on tax collection, $\tau \in [0, 1]$, as it has the effect of increasing the salience of the perceived fairness for the taxpayer.

Third, the social incentive, Σ, captures the taxpayer's desire to be seen as socially conforming with the behavior of her neighbors. It is multiplied by the probability that her evasion decision will become known to her neighbors, q, which I refer to in this paper as social exposure.

Finally, the model includes a financial penalty when caught evading composed of the perceived fine rate, f, the level of evasion chosen, e, and the taxpayer's given private income, w, the amount of effort the government spends on tax collection, τ.

The resulting first order condition (FOC) maximizing with respect to evasion, e, is:

$$FOC_e: \quad e^* = \frac{t - \tau(g - wt) - q\Sigma + \tau f}{2\rho} \tag{2}$$

3.2 Model Predictions

The comparative statics from totally differentiating the FOC reveal the following relationships and corresponding predictions regarding effects on tax evasion, e.

I begin with the effect of increasing tax collection effort, τ. Here, note that the first two non-pecuniary incentives discussed above, the sense of obligation and perceived fairness, when taken together provide a model of 'civic virtue' consistent with Frey (1997). Compliance is motivated through the sense of obligation in the absence of other external incentives or tax collection effort ($\tau = 0$). This is reinforced if the state attempts to increase its efforts to collect taxes ($\tau > 0$), when the tax is perceived to be generous ($wt - g < 0$). However, when the level of taxation is perceived to be unfair ($wt - g > 0$), the effect on tax evasion motivated by 'civic virtue' will be positive, encouraging more evasion. Increasing tax collection effort also increases the perceived penalty faced when evading, which would have the opposite effect of discouraging evasion. The ambiguity of the resulting net effect can be seen in the comparative static:

$$\frac{\partial e}{\partial \tau} = \frac{(wt - g) - f}{2\rho} \tag{3}$$

The sign on this relationship is dependent on the relative magnitudes of the total tax bill, wt, public goods received, g, and fine rate, f. Specifically, for tax collection effort to increase compliance, the combined level of public goods and financial penalties must be sufficiently high so to offset the tax bill. However, in low-tax compliance states, the level of public goods received is quite low or even zero. Similarly, the perceived fine faced may also be approaching zero if the state lacks credibility or capacity to enforce punishment when an evader is identified. With these simplifying assumptions, the comparative static becomes:

$$\frac{\partial e}{\partial \tau} = \frac{wt}{2\rho} > 0 \tag{4}$$

and the resulting prediction can be summarized:

Prediction 1: In the absence of public goods or credible fines, increasing tax collection effort *increases* evasion.

This leads directly into my second research question, the effect of increasing public goods, g, on tax evasion, e. Building off of the intuition discussed above regarding the perceived fairness of the tax, the resulting comparative static is:

$$\frac{\partial e}{\partial g} = -\frac{\tau}{2\rho} < 0 \tag{5}$$

and the resulting prediction can be summarized:

Prediction 2: Increasing public good provision *decreases* evasion.

Additional predictions for exploring mechanisms are discussed in A. Evidence from both qualitative data collection and my baseline survey supporting the modeling assumptions made here are presented in C.

4 Context

Haiti's history and institutions are briefly summarized in 4.1. The study city is introduced and baseline summary statistics are discussed in 4.2.

4.1 History and Institutions

Haiti's political history has been characterized by violence, instability, poverty and an antagonistic relationship between the state and its citizens (Girard, 2010; Dubois, 2012). Famously established via a slave rebellion, the country has since struggled to consolidate democratic institutions due to international exclusion, interference and occupation as well as domestic in-fighting, corruption and, for much of the mid-twentieth century, autocratic rule. Since emerging from dictatorship in 1988 with a new constitution, the country has suffered a series of coups, prolonged periods of political unrest and failed democratic transitions (Farmer, 2005). In the intervening three decades, there has been only one peaceful handover of power between political parties via election and the average Prime Minister has served less than 18 months before being ousted - often in the face of unrelenting protesting and political unrest (Maguire and Freeman, eds, 2017). Today, Haiti is one of the poorest countries in the world[2]. It also suffers among the lowest levels of government effectiveness ranked by the World Bank as being in the bottom 1%. Among countries in the Western Hemisphere, Haiti has the lowest levels of trust in state institutions, participation in elections and levels of taxation (Cohen et al., 2017). Non-state institutions in Haiti have proven more resilient. Haiti has the highest level of civic engagement in the Western Hemisphere (Cohen et al., 2017). Survey responses indicate that households contribute significant portions of their resources in terms of money, materials and time to informal provision of public goods (Figures 10 and 11).[3]

4.2 Study City

My study city is Carrefour, Haiti (Figure 1). It is one of the six (6) independently governed municipalities that make up the greater Port-au-Prince metropolitan area. It has an estimated population of between 500,000-1,000,000 people making it one of the largest cities in a country of 10 million total inhabitants. There are approximately 50,000 properties in the tax registry, though the mayor's office estimates that this potentially represents less than half of the total built property in the city. Though globally poor with median monthly per capita household expenditures of about $65 USD or just over $2 USD/day (Figure 12), the city is considered middle class by Haitian standards with many professionals and trades persons who commute into the capital each day.

In 2011, with funding from the United States Agency for International Development (USAID) the Government of Haiti launched a digital system for municipal finance oversight. The focus has been on individual property taxes which represent the primary source of municipal tax revenue and effectively the only progressive tax in Haiti. The current payment

[2] All reports of Gross Domestic Product (GDP) Purchasing Power Parity (PPP) in this paper are in 2018 US Dollars as calculated by the World Bank. Accessed 29 April 2018: https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.CD?locations=ET-US-CL-AR-HT-CD-IT.

[3] Indeed, my qualitative work identified more than a dozen different ways that people talk about or refer to collective action activities in my research city. The most common being: "kombit", "kove", "douvanjou", "kolon", "tet ansanm" and "team". The concept is roughly analogous to the Kenyan "harambee" more commonly referenced development economics.

Figure 1: Map of Study City and Units of Randomization

Red lines indicate census bloc boundaries which are the first level of randomization in the study. DGI is the national tax authority local office where taxes are paid. Mairie Carrefour is the Mayor's Office. The bold yellow line traces the national highway running through downtown as well as the ring road that bypasses it. Prior to the start of the intervention, these were the only roads receiving regular garbage removal services and residents were responsible to get their trash to the road.

rate in the study city is officially about 10% though estimated to be about half that when unregistered properties are also considered, and the new mayor, eager to increase tax compliance, outsourced promotion of tax payment to the private Haitian data science firm that USAID originally contracted to build and implement the system, SOLUTIONS, SA.

Table 1: Property Tax Brackets

Tax Bracket	Min Value USD	Max Value USD	Tax Rate	# Properties	% Properties
1	$ 257	$ 769	6%	2,403	6%
2	$ 770	$ 1,538	7%	33,650	84%
3	$ 1,539	$ 2,307	8%	1,602	4%
4	$ 2,307	$ 3,076	9%	1,402	3%
5	$ 3,077	NA	10%	1,103	3%

All property values are in whole USD (calcuated using 65HTG : 1USD). Property values are estimated market return for one year of renting the property as determined by a uniform calculation performed by the municipality each year which takes as input data collected on the housing market and a set of observable criteria for each property including square meters of space, number of rooms, location in city, access to street, water, power, construction materials, and roofing materials.

Property taxes are the principal source of funding for local governments and the only effective progressive tax in the country. As can be seen in Table 1, properties pay an increasing percentage of that estimated annual market rental rate for the property starting with 6% in the lowest tax bracket (6% of properties) up to 10% for properties who estimated annual rental value is greater than 3,077 USD. Note that a considerable majority of properties, 84%,

Table 2: Administrative Tax Data Summary Statistics

Statistic	Mean	Median	N
Water Access	0.60	1.00	33,763
Electricity	0.71	1.00	33,763
Road Access	0.70	1.00	33,763
Located on Bloc Boundary	0.31	0	40,060
Improved Roof	0.57	1	40,060
Number of Floors	0.18	0.00	33,763
Paid FY17	0.11	0	40,060
Paid FY18	0.12	0	40,060
Rental Value/Yr (USD)	1,246.87	769.23	40,060
Tax Rate	0.07	0.07	38,601
Tax Bracket 1	0.06	0	40,060
Tax Bracket 2	0.84	1	40,060
Tax Bracket 3	0.04	0	40,060
Tax Bracket 4	0.03	0	40,060
Tax Bracket 5	0.03	0	40,060
Total Tax Bill (USD)	89.46	53.85	40,060

N indicates the number of properties for which there was data. Analysis is limited to those blocs eligible for garbage removal services. Some household characteristics are only available for a subset of the city.

fall into the second step and owe 7% of their annual rental value or between $54-108 USD annually. The top 3 tax brackets combined only account for 10% of all properties.

Summary statistics[4] for my study city are provided in Table 2. Table 3 compares those who had paid at least once in the three years prior to the start of interventions (Paid) to those who had not paid during that period. Properties that have paid are more likely to have access to electricity, be located on a boundary road between blocs, have improved roofs, and are more likely to be multi-story. These are all easily observed characteristics which suggest that the visibility of the property may play a role in compliance. Of note, there is

[4] Though data for tax billing in the mayor's office and tax payment in the national tax authority were both digitized as part of the 2011 USAID project mentioned, these two data sets were not linked to each other. Therefore, one of the obstacles this project had to overcome was joining the two data sets based only on the variables contained within each. This was a non-trivial task for several reasons. First, there are multiple types of identification number issued by different government offices, none have wide-spread adoption in the population and neither data set was consistent in collecting or labeling these identifiers. Second, even for common terms like popular names, titles and addresses the data sets have a variety of spellings because Haiti officially has two distinct orthographies (one based on French and the other on Haitian Kreyol), mixing the two even within the same word is not uncommon and only a little more than half of the population can read and write in either language. Furthermore, there is no convention as to the order of writing one's name as one or more family names may alternatively be written before or after one or more given names, and many family names are also common given names. Third, properties can be identified by either their traditional address or by a census code comprised of three distinct three-digit numbers, but rarely were both collected and often what was collected was a nonstandard mix of the two. There is not even a convention for the order of writing the three different numbers in the census code, indeed, it was relatively rare to find all three in any one entry. To overcome these challenges, I employed a combination of data science and field verification to map more than 90% of payments from the 5 years of data made available so far for this study back to tax billing. This created what I understand to be the first city-wide matched set of administrative data in Haiti.

Table 3: Who Had Paid?

Covariates	Not Paid	Paid	p-value	N
Water Access	0.59	0.61	0.08	33763
Electricity	0.71	0.75	0.00	33763
Road Access	0.70	0.71	0.51	33763
Located on Bloc Boundary	0.30	0.33	0.00	40060
Improved Roof	0.56	0.60	0.00	40060
Number of Floors	0.17	0.23	0.00	33763
Paid FY17	0.05	0.61	0.00	40060
Rental Value/Yr (USD)	1255.81	1180.22	0.36	40060
Tax Rate	0.07	0.07	0.00	38601
Tax Bracket 1	0.04	0.17	0.00	40060
Tax Bracket 2	0.85	0.71	0.00	40060
Tax Bracket 3	0.04	0.05	0.17	40060
Tax Bracket 4	0.03	0.03	0.78	40060
Tax Bracket 5	0.03	0.04	0.00	40060
Total Tax Bill (USD)	89.23	91.13	0.65	40060

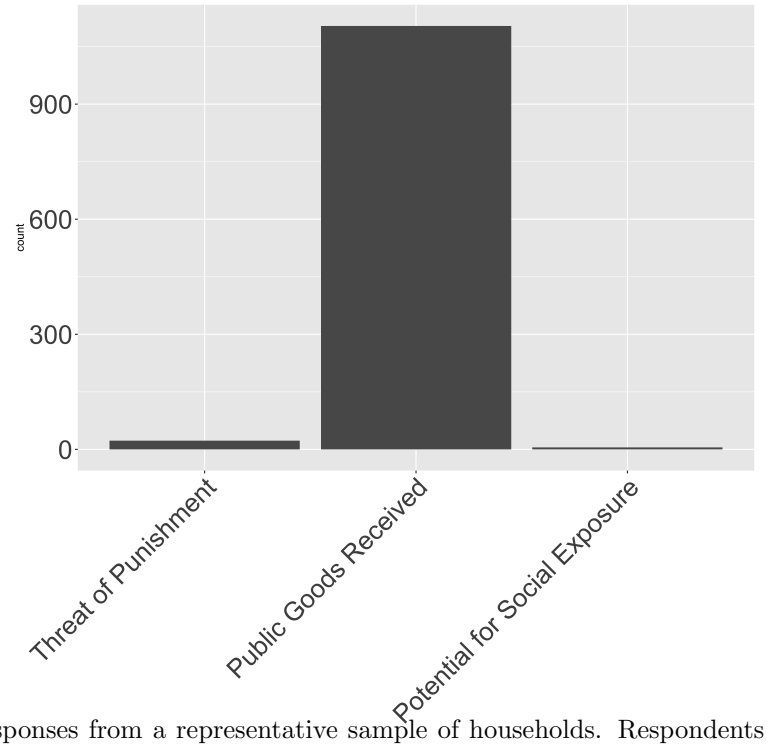
Comparing the mean values of those properties in the registry but had not paid in the previous 3 years (Not Paid) to those in the registry who had paid at least once in the previous 3 years prior to the start of interventions (Paid). Statistical significance of the difference in means is indicated by the p-value. N indicates the number of properties for which there was data. Analysis is limited to those blocs eligible for garbage removal services.

no statistically significant difference in total tax bill which suggests that these differences above are not driven by more valuable properties systematically being more likely to pay. Indeed, as is indicated in the differences in representation across tax brackets, there is greater representation of both less expensive and more expensive properties among those who paid.

Beyond characteristics gleaned from these Summary Statistics, I was also able to collect both qualitative data and a quantitative baseline survey to generate hypotheses regarding motivations for tax compliance. Discussed in greater detail in C, responses indicated incentives consistent with my modeling assumptions presented in 3. As an example, in response to the open ended question of "What would motivate you to pay property taxes?", nearly all referenced public goods (Figure 2).

Figure 2

"What would motivate you to pay property taxes?"



Baseline Survey Responses from a representative sample of households. Respondents were asked the open ended question, "What would motivate you to pay property taxes?" Only the most common responses are presented.

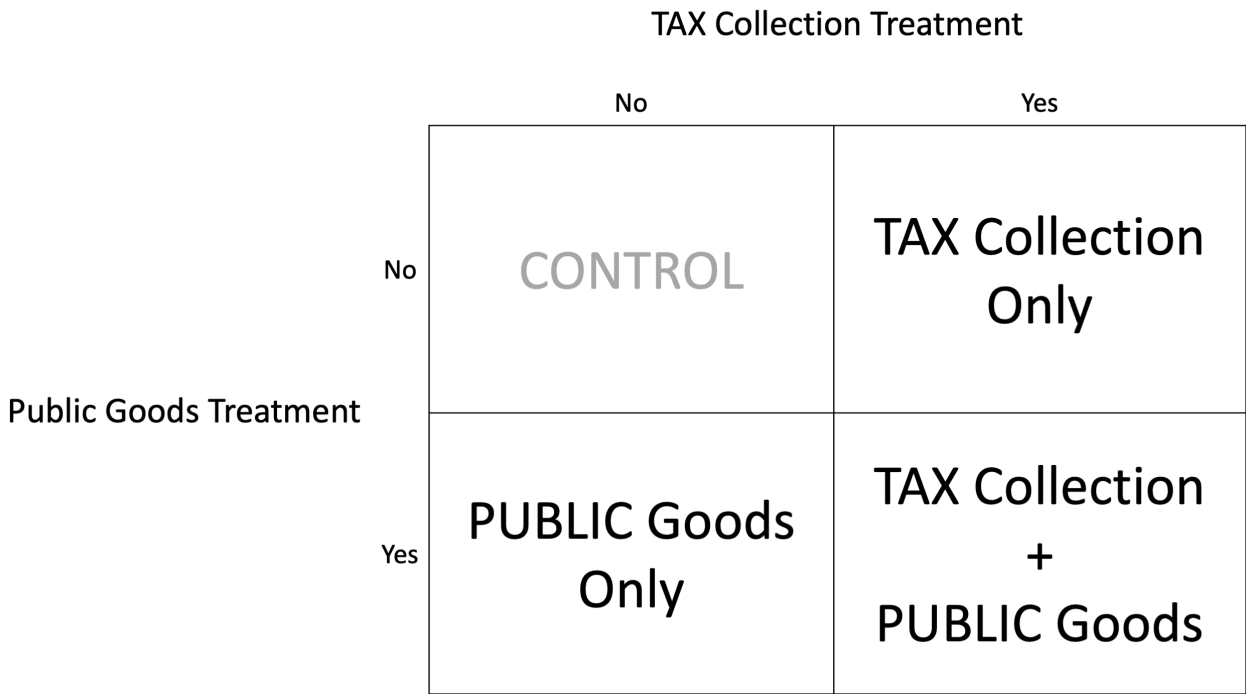
5 Methodology

I first summarize the Experimental Design in 5.1 before proceeding to describe the interventions in greater detail in 5.2.

5.1 Experimental Design

In partnership with the mayor's office and national tax authority, I cross–randomized both tax collection and public goods at–scale and over two fiscal years and the 331 census blocs that comprise the urban territory of my study city. Interventions began prior to the start of the 2019 fiscal year in August of 2018 and ran through the end of the 2020 fiscal year in September 2020. The focus is primarily on the owners of the approximately 50,000 buildings for which personal property taxes are owed (Table 2 for summary statistics and Figure 13 for geographic density of properties by bloc). Effects on independently observed measures of political unrest are collected at the census bloc-level. Effects on tax compliance are measured at the property-level from the administrative data from the central tax authority. Furthermore, a representative sample of properties were randomly selected for a panel of survey data.

Figure 3: 2x2 Factorial Design for Treatments



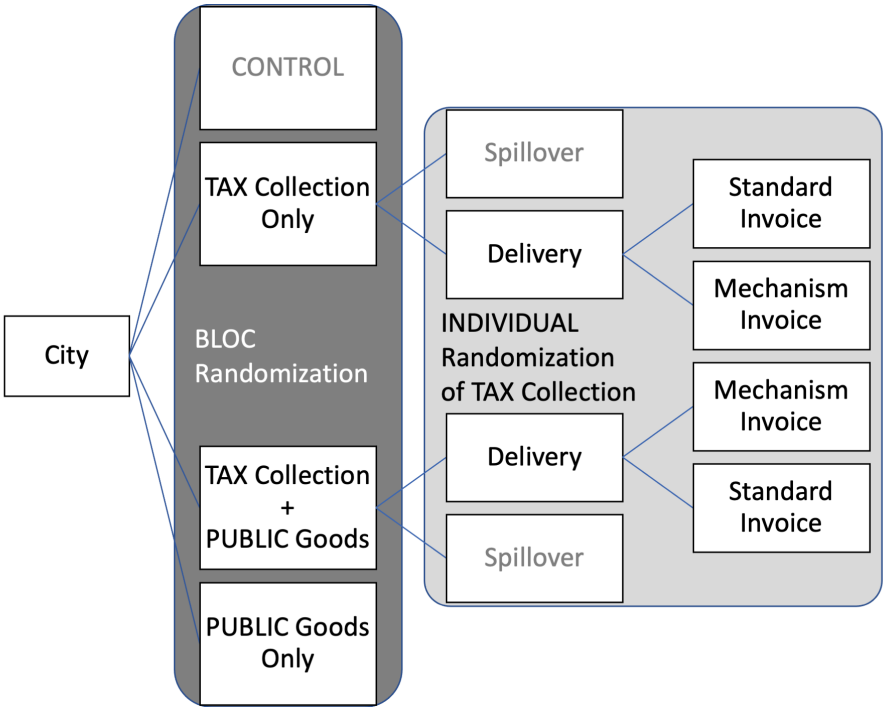
Stratified randomization performed at the bloc-level.

The experimental design is a clustered randomization with each census bloc representing one cluster. Blocs were stratified by crossing bloc–level mean history of past payment rate {above median, below median} with three divisions based on property value: {above median, below median, no property value data} and initial public goods provision {bordering baseline garbage services , not}. Blocs in each stratified grouping were then randomly assigned to one of the treatment arms show in Figure 3. This provides four different experimental conditions for analysis: 1) a set of pure control blocs wherein no new activities take place, 2) a set of blocs with only increased public goods provision, 3) a set of blocs with only increased tax collection effort and finally 4) a set of blocs where both public goods and tax collection effort are increased to identify any interaction effects.

5.2 Interventions

Conditional on being randomly assigned to receive the tax collection treatment at the bloc-level, properties were further randomized at the property-level to either receive a visit from a tax agent who hand-delivered a personalized tax invoice for the property taxes due for the fiscal year 2019 ("Delivery") or not ("Spillover") as illustrated in Figure 4. The standard invoice included a general reminder of property tax obligations, the deadline and information for how to proceed with payment as well as one's name, address, estimated property value, and total tax bill (Figure F). To further identify mechanisms, the type of invoice delivered was also randomly assigned as being either of the standard form ('Standard Invoice') or

Figure 4: Experimental Design Tree



Bloc randomization in dark gray depicts the same design information as presented in the "2x2 Factorial Design for Treatments" in Figure 3. The light gray area depicts the additional individual-level randomization conditional on being assigned tax collection treatment (either alone or with public goods) at the bloc-level. "Spillover" indicates a property is assigned to not receive an invoice from a tax agent while "Delivery" indicates that a property is assigned to receive an invoice. Then conditional on being assigned to receive an invoice, properties are either randomly assigned to receive a "Standard Invoice" which is used as the basic tax collection treatment for analysis or to receive a "Mechanism Invoice" which includes additional framing or 'nudges' both printed on the invoice and discussed during the delivery visit.

incorporating some additional framing or 'nudge' both printed at the top of the invoice itself as well as discussed during the visit ('Mechanism Invoice'). Of particular note among these frames is the social exposure treatment discussed previously wherein taxpayers were informed that if they paid their taxes that the mayor would acknowledge compliance by placing a sign on the exterior door of their house. This and the other frames will be discussed at greater length in Section 8 when I discuss mechanisms.

With respect to public goods, in blocs randomly assigned to this treatment, the mayor increased the primary public good provided by the municipal government: garbage removal services and solid waste management. This is the only city–wide service provided by mayors in Haiti to the general public (other services are provided via explicit fee–for–service mechanisms). Furthermore, though not considered the biggest problem in my study city (Figure 5), my qualitative research and subsequent baseline survey indicated that this is the service most associated with the government and which the most people would like to see the government prioritize (Figure 6). In contrast, water and electricity are almost exclusively provided by private companies, public lighting infrastructure reaches only a small percentage

Figure 5

"What is the biggest problem in your city?"

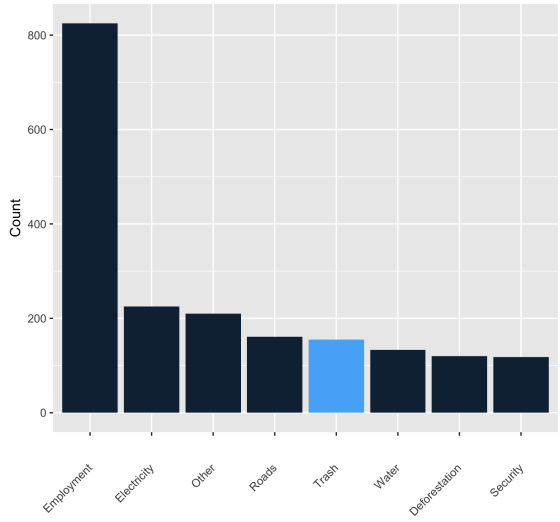
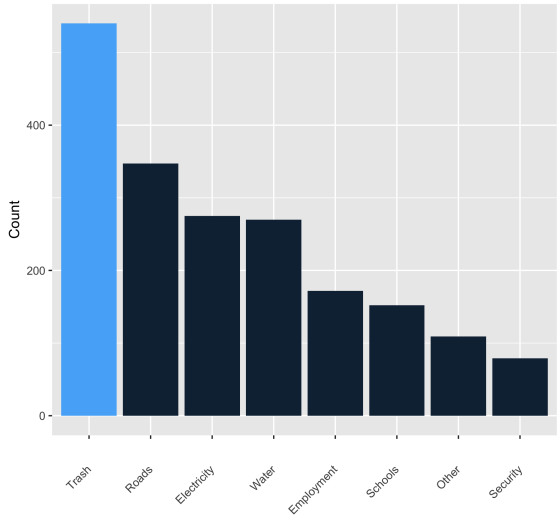


Figure 6

"What should be the government's top priority?"



Baseline Survey Responses from a representative sample of households. Respondents were asked the open ended questions indicated above and responses were categorized by the enumerators.

of streets within most cities and is unreliable on all but the most essential thoroughfares. 80% of all schools are private, and 95% of all students who pass the 4th grade national exam attend private schools. Other services, like the maintenance of public spaces, ravines, and canals, are geographically–specific.

To further test the potential of this particular public good prior to the start of interventions, I included an incentive compatible Becker–DeGroot–Marschak willingness to pay (WTP) solicitation for garbage removal services in my baseline survey. I find support across the full tax schedule suggesting that even if people were only to receive garbage removal services, WTP for those services would be sufficient for many to pay the equivalent of their full property tax bill (Figure 14).

Treatment was defined by first driving all streets in the study city to determine which blocs were accessible for the garbage removal vehicles and which streets within each bloc were accessible for door-to-door services. From the total set of 331 census blocs in the urban area of my study city, 241 were deemed accessible and thus eligible for public goods treatment. These were then randomized per Figure 3 and the stratification discussed in 5.1. I then designed a series of 24 routes incorporating only those blocs randomly assigned to receive public goods (either only public goods or public goods and tax collection together). Assignment to public goods treatment meant that all drivable roads within the bloc - including boundary roads - were incorporated into the garbage routes. The order of the routes was then randomized into a schedule designed to visit half of the routes each week - or every route on a bi-weekly basis. I then digitized the routes and loaded them onto a smartphone application that provided turn-by-turn directions even when offline so long as a GPS connection was maintained. I then oversaw the training of the drivers for the routes and the use of the smartphone application.

To track service delivery, I worked with the mayor to install GPS tracking devices in all thirteen (13) public vehicles which remotely monitor location, velocity, acceleration, and whether or not the engine is on. I further had data collectors with a second GPS tracker ride along with removal routes periodically to verify compliance. I also collected baseline trash delivery data at the local dump on origin, quantity, type of vehicle, and source of funding for each delivery received.

No blocs experienced a reduction in services as a result of the randomization introduced in this study. Rather, the mayor directed additional resources, most notably six (6) new trash compacting trucks donated from a sister city in Martinique, to increase the frequency and intensity of service. Indeed, most treatment blocs started receiving garbage removal services for the first time during this study.

Finally, to make as salient as possible the connection between tax payments and public goods, the Mayor's Office painted the sides of garbage removal vehicles with the equivalent of, "Property Taxes + Business Taxes = Services". In addition, agents were assigned to walk alongside the vehicles with a megaphone to call to the people to bring out their garbage as the vehicle passed by. These agents were also given brief scripts developed by the Mayor's communications team with messages emphasizing the connection between tax collection and public goods.

6 Results

I test and report findings for the effects of tax collection, public goods and their interaction on tax compliance in 6.1 and on political unrest in 6.2. I then provide additional causal estimates of the intensity of public goods treatment in 6.3 on both tax compliance and political unrest.

6.1 Effects on Tax Compliance

I begin by testing the intent to treat effect using the following regression:

$$\begin{aligned} Y_{ib} = \ & \beta TAX_{ib} \ + \ \delta PUBLIC_b \ + \ \gamma FRAME_{ib} \ + \\ & \mathbf{X}_{ib}\Theta \ + \ \mathbf{Z}_b\Psi \ + \ \phi_s \ + \ \epsilon_{ib} \end{aligned} \tag{6}$$

Where Y_{ib} is the outcome of interest for property i in bloc b. TAX_{ib} is an indicator variable equal to 1 if property i in bloc b is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. $PUBLIC_b$ is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. $FRAME_{ib}$ is the indicator variable for random assignment to receive a 'Mechanism Invoice' with additional framing or

Table 4:
Effects of Tax Collection and Public Goods on Amount of Taxes Paid By Property

	Amount of Tax Paid in USD	
	FY19	FY19-FY20
	(1)	(2)
Tax Collection	−6.29** (3.17)	−10.02** (5.06)
Public Goods	3.55** (1.80)	3.53 (2.51)
Mean	13.1	21.86
F-Stat	38.98	35.13
Observations	40,060	40,060
Adjusted R^2	0.02	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

'nudges' accompanying the property tax invoice. In addition, $\mathbf{X}_{ib}$ is a series of individual covariates including the valuation of the property, the total tax bill, and past tax payment history. $\mathbf{Z}_b$ are bloc covariates incorporating the population density, number of registered properties, and average level of access to services like roads, power, and water. As discussed in the Section 5, the randomization was stratified by mean bloc property value, mean payment history, and geographic presence of other salient public services. As a result, I keep stratum fixed effects, ϕ_s, in all specifications. Furthermore, in all estimations, I cluster the standard errors, ϵ_{ib}, at the bloc-level to correspond to the primary level of randomization.

I start by testing the predictions outlined in my model introduced in 3. I use amount of taxes paid (in USD) per property as the outcome of interest and regress this on the model described in Equation 6. From this equation, the coefficient β on TAX_{ib} can be interpreted as the causal effect of being assigned to the tax collection treatment. As a result, finding an estimate of $\beta < 0$ would provide evidence supporting Prediction 1: In the absence of public goods or credible fines, increasing tax collection effort *increases* evasion. Furthermore, the coefficient δ on $PUBLIC_b$ can be interpreted as the causal effect of being assigned to the public goods treatment. Thus an estimate of $\delta > 0$ would provide evidence supporting Prediction 2: Increasing public good provision *decreases* evasion.

Results are reported in Table 4. In column 1, I present my preferred specification. I test effects on property tax payments made in the first full fiscal year of the experiment, FY 2019 corresponding to the period from October 1, 2018 - September 30, 2019. I find support for my model's prediction on tax collection. Being randomly assigned to have a tax agent from

the mayor's office hand-deliver an invoice for property taxes decreased tax payments by \$6.29 in FY 2019. This effect is statistically significant by conventional standards and represents a sizable economic reduction of 48% on an average tax payment during the period of \$13.10. I also find statistically and economically significant evidence supporting my model's prediction on public goods. Being located in a bloc randomly assigned to receive garbage removal services during this period increased tax payments by \$3.55 or 27% over the mean.

In column 2, I take advantage of an additional 6 months of administrative data collected and test effects over the 18 months including all of FY 2019 (October 1, 2018 - September 30, 2019) as well as the first half of FY 2020 (October 1, 2019-March 30, 2020). Over this longer period, I find further support for Prediction 1 as the tax collection effect is of the same sign and magnitude. Being randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes decreases payments by 47% in this case or \$10.02 on an average payment of \$21.86. For public goods I find an effect over this longer period that is the same absolute increase but much less precise. It suggest a \$3.53 increase from being located in a property randomly assigned to receive garbage removal services. This point estimate is only a 16% increase on the \$21.86 average. Though still economically significant, the effect is not statistically significant by conventional standards. This may suggest that the public goods effect is attenuating with time.

In order to test if the effects of receiving both tax collection and public goods together are different than simply adding their individual effects, I introduce the interactive terms to regression Equation 6:

$$\begin{aligned} Y_{ib} = \; & \beta TAX_{ib} \; + \; \delta PUBLIC_b \; + \; \gamma FRAME_{ib} \; + \\ & \mu TAX_{ib} \times PUBLIC_b \; + \; \nu PUBLIC_b \times FRAME_{ib} \; + \\ & \mathbf{X}_{ib}\Theta \; + \; \mathbf{Z}_b\Psi \; + \; \phi_s \; + \; \epsilon_{ib} \end{aligned} \tag{7}$$

With this new equation, it is important to note that the interpretations of β and δ change. Now β is the effect of being randomly assigned to *only* receive the standard tax collection intervention and δ is the effect of being randomly assigned to *only* receive the public goods treatment. The effect of receiving both together is $\beta \; + \; \delta \; + \; \mu$ with μ identifying any additional effect arising from the combination of the two interventions. So then $\mu = 0$ would indicate that the combined effect is equal to the sum of the individual effects. $\mu > 0$ would indicate that the combined effect is greater than the two separately, in other words that there are resulting complementarities of combining the two. $\mu < 0$ would indicate that in combination the result is less than simply adding the two effects together.

Results are reported in Table 5. Column 1 is again testing the effect in the first fiscal year and Column 2 includes the full 18 months of administrative data available. The estimated effect of tax collection alone is negative, statistically significant and larger than the estimates in Table 4. This is especially true in Column 1 for FY2019 for which the estimated effect from tax collection is nearly 50% larger in magnitude. This indicates that the negative effects

of tax collection estimated in Table 4 are driven by the effect in areas without public goods. Though positive in point estimate,the effects of public goods alone in both columns of Table 5 are much smaller and much less precisely estimated. The interaction effects are also estimated to be positive and much larger. For instance, in FY2019 the interaction effect alone has a larger point estimate than the absolute value of either of the estimated effects in Table 4. This suggests that the positive public goods effects estimated in 4 are largely driven by positive interaction effects in areas also receiving the tax collection treatment. The interaction effects are also imprecisely estimated and not statistically significant meaning that though I cannot reject a null effect, I also cannot rule out potentially large positive complementarities. This lack of precision with estimates involving public goods could be expected. The treatment itself is randomized at the bloc-level, the estimated effects of each intervention alone push in opposite directions so the combination tends toward zero and the estimate is losing power by dividing the sample. As noted in the lower section of the table, the estimated linear combination of the two individual effects and the interactive effect are negative though noisy for both the first fiscal year and the full 18 months indicating that I cannot reject that the total effect is equal to zero. Taken as a whole, this table provides additional support for Prediction 1: In the absence of public goods or credible fines, increasing tax collection effort *increases* evasion, and in particular the nuance of the condition that the negative effect of tax collection is strongest in the absence of public goods and attenuates as public goods are introduced.

Table 5:
Interaction Effects of Tax collection and Public Goods on Amount of Taxes Paid By Property

	Amount of Tax Paid in USD	
	FY19	FY19-FY20
	(1)	(2)
Tax Collection	-9.22^{**} (4.53)	-12.47^{*} (7.40)
Public Goods	0.55 (1.54)	1.37 (2.56)
PUBLIC X TAX	6.96 (5.66)	5.76 (9.89)
Linear Combination	-1.71 (3.43)	-5.33 (5.75)
Mean	13.1	21.86
F-Stat	30.24	31.22
Observations	40,060	40,060
Adjusted R^2	0.02	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. "Linear Combination" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", "Public Goods", and "PUBLIC X TAX" equaling zero, and the corresponding clustered standard errors are presented in parentheses. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

6.2 Effects on Political Unrest

I now turn my attention to measuring how these interventions affect political unrest.[5] In this section I first provide evidence for survey responses and then propose two novel measures of political participation.

I again estimate Equations 6 and 7, but now with survey responses regarding protest participation as the outcome of interest. Results are reported in Table 6. In Columns 1 and 2, I estimate effects on responding "yes" to the question, "Have you or any member of your household participated in a protest in the past 30 days?". In Column 1, I find that public goods reduced the likelihood of a household answering "yes" by 0.10, more than a 50% reduction on an average of 0.19. Neither tax collection in Columns 1 and 2 nor the interaction in Column 2 have a statistically significant affect on the likelihood of a household reporting to have participated in protests. However, when controlling for the interaction, I find the effect from public goods to be statistically indistinguishable from that of Column 1. I interpret this as evidence that the reduction in protesting from public goods is primarily occurring in areas without tax collection.

Columns 3 and 4 test effects on responses to the question, "How many hours in the previous month have you or any member of your household spent protesting?" In Column 3, I find evidence that tax collection increases the amount of time spent protesting by 1.2 hours which is a 176% increase over the mean value of 0.68 hours. Column 4 controls for the interaction and finds that tax collection alone increases reported time protesting by 0.87 hours or 127%. Consistent with my findings from Columns 1 and 2, I find public goods alone reduces time reported protesting by 0.43 hours or 60%. The interaction effect in this case is large and positive, and though noisy, it does suggest the potential for a compounding effect wherein providing providing public goods while attempting to increase tax collection does not move more households to protest (Column 2), but for those that do protest, the interaction may increase the amount of time those households spend protesting.

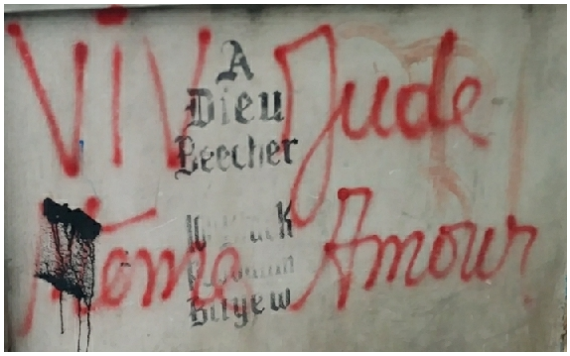
[5] This RCT was conceived of and designed to also directly test effects on voter turnout and electoral outcomes in the next election in Haiti. However, due to a confluence of events, the elections have been delayed into the new year. Though it is my intention to continue following the effects of these interventions into the election, those results will be the subject of a companion paper.

Table 6:
Effects of Tax Collection, Public Goods and their Interaction on Self-Reported Protest Participation

	Protest Participation in the Past Month (Self-Reported)			
	Indicator		Hours	
	(1)	(2)	(3)	(4)
Tax Collection	0.07 (0.04)	0.07 (0.06)	1.27* (0.70)	0.95* (0.52)
Public Goods	−0.11*** (0.04)	−0.11** (0.04)	−0.16 (0.46)	−0.44** (0.22)
PUBLIC X TAX		−0.003 (0.09)		0.74 (1.19)
Linear Combination		-0.04 (0.06)		1.24 (1.2)
Mean	0.19	0.19	0.68	0.68
F-Stat	3.04	2.7	2.44	2.27
Observations	445	445	445	445
Adjusted R^2	0.02	0.02	0.01	0.01

"Indicator" is a binary indicator equal to 1 if a household answered "yes" to the question, "Have you or any member of your household participated in a protest in the past 30 days?". "Hours" is the amount of time in hours given in response to the question, "How many hours in the previous month have you or any member of your household spent protesting?" "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. "Linear Combination" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", "Public Goods", and "PUBLIC X TAX" equaling zero, and the corresponding clustered standard errors are presented in parentheses. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Figure 7: Partisan Graffiti



The two most common forms of political graffiti in the city. The left saying"Down with Mayor Jude" and the right saying"Long Live Jude, One Love".

Political participation can often take the form of public speeches and debates. My qualitative research indicated that in this setting political graffiti (Figure 7) is a principal forum for this expression. Research in other disciplines support this interpretation, notably the analysis of De Ferrari (2020) who finds graffiti in this particular setting to serve as "public poetry" leading to "political praxis" by providing a medium for political debates to evolve over time across the walls of the city. I conducted a census of all streets in my study city and geo-tagged all graffiti identified resulting in a bloc-level outcome of interest for which I can test the effects of tax collection and public goods using the bloc-level randomization discussed in 5. I estimate effects using Equations 8 and 9, and control for bloc-level propensity for political graffiti by controlling for the presence of messages referencing the previous election which took place prior to the start of interventions. In this way I introduce geo-tagged political graffiti as a novel independent metric of political engagement and unrest.

$$Y_b = \beta TAX_b + \delta PUBLIC_b + \mathbf{Z}_b\Psi + \phi_s + \epsilon_b \tag{8}$$

$$Y_b = \beta TAX_b + \delta PUBLIC_b + \mu TAX_b \times PUBLIC_b + \mathbf{Z}_b\Psi + \phi_s + \epsilon_b \tag{9}$$

Before testing my outcomes of interest, I first conduct two placebo tests. I begin by testing for effects on graffiti in general to address concerns that any estimated effects on political graffiti are driven by the overall prevalence of graffiti in a given bloc. Second I test effects on graffiti focused on a social movement unrelated to local government or the interventions being tested (see Appendix D for a discussion). Table 16 in the Appendix provides the results of both tests, and in summary I do not find any statistically significant effects for either outcome of interest.

Table 7:
Effects of Tax Collection, Public Goods and their Interaction on Political Graffiti

	Graffiti					
	Political		Positive Political		Negative Political	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax Collection	0.046	0.117**	−0.014	−0.058*	−0.008	0.044
	(0.043)	(0.057)	(0.027)	(0.032)	(0.033)	(0.044)
Public Goods	−0.058	0.034	0.014	−0.043	−0.040	0.028
	(0.040)	(0.060)	(0.029)	(0.043)	(0.031)	(0.052)
PUBLIC X TAX		−0.183**		0.112*		−0.135*
		(0.087)		(0.066)		(0.069)
Linear Combination		-0.032 (0.049)		0.012 (0.05)		-0.063 (0.03)
Mean	0.116	0.116	0.054	0.054	0.066	0.066
F-Stat	2.74	2.8	1.78	2.25	1.29	1.97
Observations	241	241	241	241	241	241
Adjusted R^2	0.018	0.034	0.005	0.016	−0.015	−0.002

"Political" is a binary indicator equal to 1 if any graffiti coded with a political message was identified in the bloc and zero otherwise. "Positive Political" is a binary indicator equal to 1 if the message was identified as political and was positive in content (generally calling for support with 'long-live') was identified in the bloc and zero otherwise. "Negative Political" is a binary indicator equal to 1 if the message was identified as political and was negative in content (generally condeming with 'down with') was identified in the bloc and zero otherwise. "Tax Collection" is an indicator variable equal to 1 if a bloc is randomly assigned to have tax agents from the mayor's office hand-deliver invoices for property taxes in the bloc and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a census bloc is randomly assigned to receive garbage removal services and equal to zero otherwise. "Linear Combination" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", "Public Goods", and "PUBLIC X TAX" equaling zero, and the corresponding clustered standard errors are presented in parentheses. Analysis is limited to only census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions control for propensity for political graffiti with an indicator variable equal to 1 if messages referring to the previous election completed prior to the start of interventions was identified in the bloc and zero otherwise. All regressions also include additional bloc controls and strata fixed effects. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Estimated effects of tax collection and public goods on political graffiti are reported in Table 7 and provide further evidence of tax collection increasing political unrest and public goods mitigating this effect. The estimates in Column 1 are not statistically significant by conventional measures, but the point estimates are consistent with the effect of tax collection increasing political graffiti while public goods decrease political graffiti. Introducing the interaction in Column 2 provides clearer evidence that the interventions impact political graffiti in this setting. Tax collection alone has the statistically significant effect of doubling the probability of finding political graffiti in a bloc by increasing 0.117 on a mean of 0.116. Though the effect of public goods alone is quite small and imprecise if positive, the interaction is a relatively large and statistically significant reduction in the probability of finding political graffiti in a bloc by 0.183 providing evidence that public goods significantly mitigates the large positive effect of tax collection on political graffiti. As seen in the lower section of the table, the estimated linear combination of the total effect is negative though not statistically different from zero itself.

Columns 3-6 estimate effects on the tone of political graffiti finding results consistent with tax collection increasing political discontentment and additional evidence of public goods mitigating those effects. In Columns 3-4, I estimate effects on the probability in a given bloc of finding political graffiti expressing support, generally in the form of "long-live" followed by the named official or policy. In Column 3, I find suggestive though imprecise evidence that tax collection is reducing the prevalence of positive messages while public goods are increasing such messages. This interpretation is consistent with Column 4 where, with the interaction added, tax collection alone essentially eliminates positive political graffiti in treated areas with an estimated coefficient of –0.058 on a mean of 0.054. Moreover, in the interaction, I find a statistically significant two-fold increase in positive political messages of 0.112 though the linear combination of the full effect is not statistically different than zero. In Columns 5-6, I estimate the effects on political messages condemning or expressing dissatisfaction with the government or individual officials. For the most part these estimates are smaller and even less precise than the previous two columns. The estimated effect of public goods in Column 5 suggests that it reduces the probability of negative political graffiti in the bloc. This interpretation is again supported by adding the interaction in Column 6. The result is a statistically significant and economically large reduction of 0.135 on a mean of 0.066. Furthermore, as can be see in the lower section of the table, the linear combination of all three coefficients in Column 6 results in a statistically significant reduction in the probability of finding negative graffiti in a bloc of 0.063 on the mean of 0.066. So in summary, I find evidence that tax collection increases political graffiti overall while also causing the tone to be more negative. I also find evidence that public goods mitigate this effect from tax collection in the interaction.

Figure 8: Barricade Construction

A relatively common form of protest over the past two years, in particular during the period of 'Lock-Down'.

Haiti had, and in many ways continues to, experience a particularly difficult period which culminated in all major cities including my study city being shut down for two months in the fall of 2019. One of the primary means of protest in Haiti is filling the streets with debris and other hazards. In the more extreme forms, neighbors construct barricades that significantly inhibit or completely bloc traffic from entering an area. Commonly built from old vehicles, used construction materials and worn out tires that are pilled together and often set on fire, these barricades are a very public and destructive form of asserting independence and dissatisfaction with the state (Figure 8). Having mapped the universe of barricades during this period (and their relative magnitude), I introduce my second independent metric for measuring political engagement and unrest: barricade construction.

I estimate Equations 8 and 9 with barricade construction as the outcome of interest. Table 8 presents the results providing further evidence of tax collection causing increased unrest which is reduced by the introduction of public goods. Columns 1-2 estimate affects on the probability that "Any" protest construction was identified in the bloc. Though none of these estimates are statistically significant, they do all provide point estimates with signs consistent with tax collection increasing unrest while public goods mitigate this engagement with more of the effect coming through the interaction. Columns 3-4 focus only on the most extreme forms of barricade construction that completely bloc the road from being accessed. I find statistically significant and large effects of tax collection increasing barricade construction by 0.099 or about a 135% increase over the mean of 0.071. Furthermore, I find that public goods reduce barricade construction by 0.066 or more than 90%. Column 4 indicates that the tax collection effect is driven by those blocs only receiving tax collection which increase barricade construction by 0.135 or 190% over a mean of .071. Though the estimated effects of both the public goods only and interaction are not statistically significant,

Table 8:
Effects of Tax Collection, Public Goods and their Interaction on Protest Barricade Construction

	Protest Construction					
	Any		Barricades		Access Blocked	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax Collection	0.060	0.083	0.099***	0.135***	0.051	0.105
	(0.038)	(0.051)	(0.032)	(0.047)	(0.060)	(0.078)
Public Goods	−0.056	−0.026	−0.066**	−0.019	−0.149**	−0.078
	(0.038)	(0.052)	(0.029)	(0.029)	(0.058)	(0.083)
PUBLIC X TAX		−0.060		−0.093		−0.139
		(0.081)		(0.063)		(0.117)
Linear Combination		-0.002 (0.054)		0.023 (0.039)		-0.112 (0.082)
Mean	0.124	0.124	0.071	0.071	0.357	0.357
F-Stat	2.05	1.93	2.75	2.46	4	3.84
Observations	241	241	241	241	241	241
Adjusted R^2	0.190	0.188	0.108	0.112	0.129	0.130

"Any" is a binary indicator equal to 1 if any form of protest construction was identified in the bloc and zero otherwise. "Barricades" is a binary indicator equal to 1 if significant protest construction to bloc the flow of vehicles was identified in the bloc and zero otherwise. "Access Blocked" is an indicator variable equal to 1 if a the primary road into a bloc was blocked by a "Barricade" and equal to zero otherwise. "Tax Collection" is an indicator variable equal to 1 if a bloc is randomly assigned to have tax agents from the mayor's office hand-deliver invoices for property taxes in the bloc and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a census bloc is randomly assigned to receive garbage removal services and equal to zero otherwise. "Linear Combination" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", "Public Goods", and "PUBLIC X TAX" equaling zero, and the corresponding clustered standard errors are presented in parentheses. Analysis is limited to only census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions also include bloc controls and strata fixed effects. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

they are both negative and the interaction's point estimate is much larger in absolute value suggesting that the positive public good effect from Column 3 is largely the result of public goods mitigating the positive effect of tax collection. Finally, in Columns 5 and 6, I estimate effects on having access blocked by barricade construction. Though less precise, I find point estimates consistent with tax collection, especially tax collection by itself, increasing the probability of having primary access blocked due to barricade construction. I find statistically significant evidence that public goods reduce this probability by 0.149 on a mean of 0.357 or about a 40% reduction.

Though the effects estimated are derived from the exogenous variation induced by the experimental design, there may be concerns regarding the causal pathway and interpretation of effects from public goods. For instance, the garbage removal service itself resulting from the public goods treatment may be simply eliminating the prevalence of materials for barricade construction and thus mechanically reducing barricades. It would still be an effect of interest for the local government regardless as it indicates a reduction in a particularly salient and difficult to respond to form of disruptive and often violent protest. However, this mechanical explanation is unlikely as the overwhelming majority of garbage collected on the routes are household refuse loaded by the neighbors themselves while barricades are generally constructed from tires, old vehicles and old construction materials (see Figure 8). Another potential explanation for the effect may be that the community wants to avoid losing access to this service and as a result tempers their proclivity to build barricades not out of a sign of gratitude but out of a desire to maintain services. Again, from the perspective of the government, the pathway would not change the value of the finding. However, in this case the emphasis would be on not merely providing public goods, but providing public goods that specifically and transparently require road access into order to be provided.

6.3 Public Goods Intensity on Tax Compliance and Political Unrest

I now test the effects of the relative intensity of the public goods treatment on my outcomes of interest. To do this I leverage the quasi-random variation in the number of times garbage vehicles passed through each bloc on removal routes. The schedule for these routes, including the vehicles and personnel, was randomly assigned over a two week schedule at the start of the intervention. As a result, the variation in routes passing through a given bloc arise from the lack of availability of assigned vehicles or personnel. The majority of missed routes were due to vehicle breakdowns, with flat tires being the primary issue. Longer periods were missed on a few occasions when gangs hijacked vehicles for fuel or ransom. Occasional staffing issues also interfered with route completion due to sick days or unexcused absences.

$$INTENSITY_b = \rho_1 TAX_{ib} + \rho_2 PUBLIC_b + \rho_3 FRAME_{ib} + \mathbf{X}_{ib}\Lambda + \mathbf{Z}_b\Pi + \phi_s + u_{ib} \quad (10)$$

$$Y_{ib} = \beta TAX_{ib} + \delta \widehat{INTENSITY}_b + \gamma FRAME_{ib} + \mathbf{X}_{ib}\Theta + \mathbf{Z}_b\Psi + \phi_s + \epsilon_{ib} \quad (11)$$

$$INTENSITY_b = \rho_1 TAX_b + \rho_2 PUBLIC_b + \mathbf{Z}_b\Pi + \phi_s + u_b \quad (12)$$

$$Y_b = \beta TAX_b + \delta \widehat{INTENSITY}_b + \mathbf{Z}_b\Psi + \phi_s + \epsilon_b \quad (13)$$

In Table 9, I present the results of testing the effects of public goods intensity on my independently observed outcomes of interest. The results are consistent with my primary findings. The resulting estimates of the effect of tax collection remaining qualitatively similar in sign and magnitude as before, and the estimated effects of public goods intensity are consistent with the previous public goods estimates. In the odd numbered columns are the reduced-form estimates of Equation 6 in Column 1 and Equation 8 in Columns 3, 5 and 7 replacing the public goods indicator variable for bloc b with the public goods intensity count variable for bloc b. Assuming the variation in intensity to be as-good-as-random, the interpretation of the resulting coefficient estimates is the marginal causal effect of one additional removal route passing through the bloc. If the variation were not as-good-as-random, the interpretation would be one of correlation instead. I also estimate more flexible forms but do not find increased explanatory value. To ensure I am identifying effects from random variation, in the even numbered columns, I present the second stage of the instrumental variables regression described for Column 2 in Equation 11 and for Columns 4, 6 and 8 in Equation 13. Both of these regressions make use of my randomly assigned public goods indicator in a first-stage to instrument the public goods intensity per Equations 10 and 12 respectively. This instrument satisfies the exclusion restriction as it was randomly assigned, and it also serves as a very strong instrument as can be seen in the first stage regression output (Appendix Table 17). The interpretation of the resulting coefficient estimates for these models is the local average treatment effect (LATE) of one additional removal route passing through the bloc induced by the random assignment of the bloc to the public goods treatment.

Columns 1 and 2 present findings consistent with public goods increasing tax compliance. Both columns indicate an economically and statistically significant increase in tax payments caused by an additional garbage removal route passing through the bloc. The reduced form OLS estimate indicates an additional \$0.18 per property in tax revenue while the IV estimate indicates an LATE of \$0.13 per property. This range of 1-1.3% increase from the mean value suggests between \$5,210 - \$7,210 in additional revenue could be earned from one route passing through the entire city. As I will discuss more Section 9, this far

Table 9:
Effects of Tax Collection and Intensity of Public Goods on Amount Tax Paid and Political Unrest

	Amount Tax Paid USD in FY19		Political Graffiti		Barricades		Access Blocked	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tax Collection	−6.22*	−6.18*	0.0442	0.0442	0.0974***	0.0967***	0.0472	0.0467
	(3.18)	(3.19)	(0.0425)	(0.0424)	(0.0320)	(0.0321)	(0.0600)	(0.0601)
Public Good Intensity	0.18**		−0.0023**		−0.0011		−0.0046**	
	(0.09)		(0.0011)		(0.0010)		(0.0019)	
IV Intensity		0.13**		−0.0022		−0.0025**		−0.0056**
		(0.065)		(0.0015)		(0.0011)		(0.0022)
Mean	13.1	13.1	0.12	0.12	0.071	0.071	0.36	0.36
F-Stat	38.39	38.9	2.87	2.75	2.4	2.71	3.86	4
Observations	40,060	40,060	241	241	241	241	241	241
Adjusted R^2	0.02	0.02	0.02	0.02	0.10	0.09	0.13	0.13

"Amount of Tax Paid USD in FY19" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "Political" is a binary indicator equal to 1 if any graffiti coded with a political message was identified in the bloc and zero otherwise. "Barricades" is a binary indicator equal to 1 if significant protest construction to bloc the flow of vehicles was identified in the bloc and zero otherwise. "Access Blocked" is an indicator variable equal to 1 if a the primary road into a bloc was blocked by a "Barricade" and equal to zero otherwise. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "public goods Intensity" is an count variable of the number of times a garbage collection vehicle entered the census bloc based upon the random assignment of route schedule and the quasi-random variation arising from the interruptions to that schedule (vehicle maintenance, insufficient staffing and the hijacking of vehicles by local gangs). "IV Intensity" is the estimated value of "public goods Intensity" instrumented by the "Public Goods" random assignment variable in a firststage regression. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

exceeds both current and projected operating costs and suggests the activity could provide substantial positive revenue for the city.

Columns 3-8 present findings consistent with public goods reducing political unrest. Columns 3 and 4 both indicate a reduction in political graffiti caused by an additional garbage removal route though only the reduced form estimate is statistically significant by conventional standards indicating a 2% reduction from the mean. Though Column 5 is imprecisely estimated, it suggests an additional garbage removal route leads to a reduction in Barricade construction, and the IV estimate in Column 6 is statistically significant with an additional route induced by the exogenous variation from my RCT decreasing the probability of localized barricade construction by 3.5% from the mean value. Finally both Columns 7 and 8 find statistically significant negative effects on the probability of road access to an area being blocked by barricades with the IV estimate indicating an LATE of a 1.5% reduction from the mean value.

Taken together, my results show that tax collection, especially in the absence of public goods, reduces tax payments and increases negative political engagement, political protesting

and the more destructive and violent forms of unrest. Furthermore, all of these adverse effects are found to be largely mitigated by public goods which, both as a whole treatment and on the margin, lead to increased tax payments as well as positive or supportive political engagement while reducing protesting - especially the more destructive or violent unrest.

7 Robustness

Leveraging several experimental elements of my design, I first test for and find evidence against tax collection spillovers in 7.1. I then make use of geographic variation to test for spillovers resulting from the public goods treatment in 7.2 and find evidence suggesting that my preferred specification may be an underestimate.

7.1 Testing for and Finding Evidence against Tax Collection Spillovers

Continuing with the examination of mechanisms, I next test for spillovers potentially affecting the interpretation of my estimated negative effect of tax collection on tax payments. It may be an underestimate if the taxpayer who is not visited by a tax agent also reduces tax payments in response. This might be because the taxpayer interprets the lack of visit as a further sign of state weakness or if a sense of solidarity or sympathy causes a similar crowding-out of tax morale in taxpayers who are not visited. The negative tax collection effect could also be an overestimate though if this taxpayer that was not visited responds by increasing payments. This might be because she interprets the lack of visit as a signal of trust from the state and thus her feeling of obligation or tax morale increases and with it her payments. It could also be in response to anticipating that her neighbors will reduce payments and potentially attract additional attention of the state to the area increasing the chance of facing a financial penalty if found in noncompliance.

I first test for tax collection spillovers by leveraging the multiple levels of randomization in my experimental design. Specifically, I compare the responses of properties that are located in blocs randomly assigned at the bloc-level to not receive tax invoices to those randomly assigned into tax collection blocs but further randomly assigned at the property-level to not receive a visit from the mayor's tax agent. I call this second group my 'spillover' properties as their response is the combination of not receiving a visit and also being close to someone who does receive a visit.

Table 10 presents the results of these first spillover tests. In all regressions I use bloc fixed effects so that my identifying variation is the property-level randomization allowing me to compare the responses of 'spillover' properties to those randomly assigned to receive a tax collection visit. In Columns 1 and 3, I find estimated effects of tax collection that are statistically significant and almost 50% larger than my preferred specification in 4. The coefficients are similar in magnitude to the estimated effects of tax collection only in Table 5. This is especially the case for FY2019 while over the full 18 months the point estimate here is both larger and more statistically significant. In Columns 2 and 4, I divide the sample

Table 10:
Testing for Spillovers of Tax Collection on Amount of Taxes Paid By Property with Bloc Fixed Effects

	Amount of Tax Paid in USD			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−9.16** (3.56)		−14.83** (5.82)	
ONLY TAX		−9.75*** (2.81)		−14.02*** (4.65)
TAX + PUBLIC		−8.33 (5.94)		−15.96 (9.80)
Mean	13.1	13.1	21.86	21.86
Bloc FE	Yes	Yes	Yes	Yes
Observations	40,060	40,060	40,060	40,060
Adjusted R^2	0.02	0.02	0.03	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "ONLY TAX" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned only to the tax collection treatment and equal to zero otherwise. "TAX + PUBLIC" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned both the tax collection and public goods treatment and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

between blocs only receiving tax collection and those receiving both tax collection and public goods. The resulting coefficients for only tax collection are even more precisely estimated and of the same magnitude. While the point estimates for blocs receiving tax collection and public goods together are of the same magnitude, they are much less precisely estimated and not statistically significant by conventional standards. To further test if the 'spillover' properties are lowering the absolute value my estimates, in Table 11, I drop them from my estimates. The results in Columns 1-4 do not vary significantly from my estimates in Tables 4 and 5. Taken together, these estimates are consistent with my preferred specification suggesting that the estimates are neither an under nor an overestimate.

In Columns 5-6 of Table 11 I also test the effect of a placebo tax letter delivery treatment I developed with the Mayor's office (See F.7). This treatment isolates the effect of a government agent visit from the information effect specifically caused by discussing taxes. Though the effect is imprecisely measured, it is positive and large. This is evidence that the visit by the government itself is not driving the large negative effect of tax collection on tax compliance. This suggests the effect is instead driven by the discussion of taxes or the information only about taxes contained within the invoice or discussed during the visit. In fact, it indicates that the tax information treatment may be considerably larger in magnitude than the estimate suggests as it also includes the noisy large positive effect from the government visit. This is consistent with my model in that the visit itself communicates state capacity and thus increases the perceived probability of facing a fine. When the state specifically uses that visit to discuss tax payments, this triggers the crowding-out of tax morale and results in a net negative effect.

7.2 Testing for Spillovers from Public Goods and Finding an Underestimate

I now test for spillovers from the public goods treatment by leveraging geographic variation. One might expect the estimated positive public goods treatment effects on tax compliance to be an underestimate due to 'contaminated' control properties that are located on roads that form boundaries with treated blocs and as a result are actually as-good-as-treated. Similarly, as the trash routes reach into the city, they likely reduce the total distance needed to travel to a garbage removal point so that many control properties are effectively being treated with a lesser intensity. Another potential reason to expect an underestimate is that there was a net increase in garbage removal services across the city and this anecdotally resulted in an overall reduction in the amount of trash in the center of the city and along the main roads of the city while also increasing the overall visibility of government services. As a result even properties in control areas might be treated to some extent by this net improvement. However, if instead, households in control areas resent being left out or otherwise respond to their status of being randomly assigned to control by reducing their tax payments, then the estimated effect on public goods would be an overestimate.

To test for spillovers, I first drop all boundary properties from my analysis. In Table 12, I present results. In all Columns I find public goods effects that are larger and in most cases more statistically significant than in my preferred specification. In Column 1, I

Table 11:
Testing for Spillovers of Tax Collection on Amount of Taxes Paid By Property by Dropping 'Spillover' Properties

	Amount of Tax Paid in USD					
	FY19		FY19-FY20		FY19-FY20	
	(1)	(2)	(3)	(4)	(5)	(6)
Tax Collection	−6.07*	−10.16**	−9.71*	−14.29*	−9.49*	−11.56
	(3.45)	(5.07)	(5.40)	(8.31)	(5.03)	(7.26)
Public Goods	3.81**	0.74	4.01	2.13	3.54	2.68
	(1.84)	(1.62)	(2.45)	(2.47)	(2.50)	(2.39)
Placebo					12.12	23.19
					(12.79)	(20.96)
PUBLIC X TAX		9.50*		10.38		4.63
		(5.72)		(9.70)		(9.77)
PUBLIC X Placebo						−28.39
						(21.36)
Mean	12.98	12.98	12.98	21.62	21.86	21.86
Drop 'Spillover' Properties	Yes	Yes	Yes	Yes	No	No
Observations	37,903	37,903	37,903	37,903	40,060	40,060
Adjusted R^2	0.02	0.02	0.03	0.03	0.03	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "ONLY TAX" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned only to the tax collection treatment and equal to zero otherwise. "TAX + PUBLIC" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned both the tax collection and public goods treatment and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 12:
Testing for Spillovers of Pubic Goods on Amount of Taxes Paid By Property by Dropping all Boundary Properties

	Amount of Tax Paid in USD			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−3.91 (3.26)	−6.16 (4.23)	−4.07 (4.52)	−6.69 (6.33)
Public Goods	5.32** (2.53)	2.03 (1.88)	5.21* (3.16)	3.25 (2.73)
PUBLIC X TAX		5.55 (4.47)		6.10 (6.84)
Mean	12.82	12.82	21.22	21.22
Drop all boundary properties	Yes	Yes	Yes	Yes
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.03	0.03	0.05	0.05

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

estimate that public goods increase tax payments by $5.32 or more than 40% over the mean as compared to my preferred specification showing only a 27% increase in Table 4. Column 3 is similarly larger and in this specification statistically significant. Of note, in both Columns 2 and 4 the point estimates on interaction effects are smaller both in terms of comparing to my preferred specification as well as when comparing their relative magnitudes to that of the public goods alone in their respective regressions. This provides suggestive evidence of potentially heterogeneous effects of public goods on interior properties as compared to boundary properties with the former directly increasing their tax payments in response to public goods while the later may be more likely to mitigate their negative tax collection responses in the interaction. In the Appendix in Table 18, I also drop just contaminated properties from control and find qualitatively similar results. Taken together, this evidence suggests my estimated public goods treatment effect on tax compliance is an underestimate of the real effect.

To further test for spillovers, I restrict analysis to just properties that have been randomly assigned to not receive the public goods treatment. I then create an indicator variable equal to 1 if the bloc borders any of the blocs randomly assigned to receive public goods (and thus border some length of the assigned garbage removal routes) and zero otherwise. The resulting distribution is 101 blocs that border routes in some way and 42 blocs that do not. The intuition for this test is that if there are either positive or negative spillover effects from being just outside of the public goods treatment which dissipate over distance or if being

close to treatment is as-good-as being treated and only the response farther removed would be the true control response, then there should be statistically significant differences in tax payments between those properties just beyond treatment and those even farther away.

Results are presented in Table 13. Most notably, none of the coefficients estimated from the "Bordering Public Goods Blocs" indicator are statistically significant. Thus I fail reject the null hypothesis of there being no spillovers in either direction resulting from differential responses of untreated properties close to and far away from treatment. All of the coefficient estimates are relatively small in magnitude, for instance in Column 1 the point estimate of \$0.21 is less than 2% the mean value of \$13.10. Finally, the sign of the border effect flips when controlling for the interaction in Columns 2, 4 and 6 which is inconsistent with a large spillover effect in one direction or the other.

Notably in Column 2 when controlling for the interaction, the point estimate increases by an order of magnitude, and though it is still not statistically significant, it is the only estimate larger than its cluster-adjusted standard error in magnitude. Then Columns 3-4, when I drop properties that are actually on the roads that are receiving services, the point estimate for this same isolated set of blocs reduces to just \$0.68. These findings are consistent with the evidence above of my public goods estimate being an underestimate as the border properties appear to be increasing payments in response to treatment. This effect appears to diminish quickly as the rest of the bloc does not respond in kind. This can be seen in Columns 5-6 with all boundary properties dropped and the coefficients on the border indicator remaining similar in magnitude. In summary, I do not find evidence of significant negative spillovers driving up my estimate, but I do find some suggestive evidence that I have underestimated the true public goods effect on tax compliance[6].

[6] For estimated effects on alternative measures of tax compliance, see Appendix Tables starting with 19.

Table 13:
Testing for Spillovers of Pubic Goods on Amount of Taxes Paid By Property by Comparing Blocs Bordering Public Goods Blocs to those Not Bordering Public Goods Blocs

	Amount of Tax Paid in USD in FY19					
	(1)	(2)	(3)	(4)	(5)	(6)
Tax Collection	−10.39**	−6.97	−10.81**	−7.93	−9.95***	−15.84
	(4.11)	(6.86)	(4.36)	(6.97)	(3.62)	(9.97)
Bordering Publc Goods Blocs	−0.21	2.06	−1.07	0.68	−1.45	0.43
	(1.62)	(1.89)	(1.62)	(1.77)	(1.69)	(1.57)
Bordering Publc Goods Blocs X TAX		−4.92		−4.31		8.84
		(8.90)		(9.40)		(10.71)
Linear Combination		-9.83 (5.27)		-11.56 (5.49)		-6.58 (3.05)
Mean	13.1	13.1	13.1	13.1	13.1	13.1
All Blocs are Control Bloc for Public Goods	Yes	Yes	Yes	Yes	Yes	Yes
Dropped Contaminated	No	No	Yes	Yes	Yes	Yes
Dropped All Boundary Properties	No	No	No	No	Yes	Yes
F-Stat	34.07	27.89	32.58	26.25	41.66	32.65
Observations	24,374	24,374	22,093	22,093	17,131	17,131
Adjusted R^2	0.06	0.06	0.07	0.07	0.07	0.07

"Amount of Tax Paid in USD in FY19" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization and further restricted to only those census blocs randomly assigned to NOT receive the public goods intervention. "Bordering public goods Blocs" is an indicator variable equal to 1 if a property is located in a census bloc bordering at least one census bloc that was randomly assigned to receive garbage removal services and equal to zero otherwise. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

8 Mechanisms

Following the additional predictions of my model as elaborated in A, I first discuss my social exposure experiment and present the resulting evidence of an antisocial punishment mechanism in 8.1. Then in 8.2, I present additional experimental evidence that making penalties or obligations salient mitigate and, in the case of obligations, reverse the negative tax collection effect.

8.1 Testing Social Exposure and Finding Evidence of Antisocial Punishment

As discussed with respect to the existing Literature in 2 and Modeled in 3, social norms and in particular the possibility of your neighbor becoming aware of your tax compliance decision may be quite important in maintaining the stability of the low-tax equilibrium in low-compliance settings. To experimentally test my model's prediction[7] A.1: Increasing social exposure of tax compliance *increases* evasion, I introduce a novel treatment for increasing probability of social exposure of tax compliance (Figure 9). Within blocs randomly assigned to tax collection, fifty percent (50%) of households randomized to receive a tax invoice were also randomly assigned to be notified that if they paid their taxes in the subsequent 30 days, the government would come and install a sign on their house's exterior-facing door indicating payment (Appendix F.8). To identify these effects I now estimate:

$$\begin{aligned} Y_{ib} = \; & \beta TAX_{ib} \; + \; \delta PUBLIC_b \; + \; \kappa SOCIAL_{ib} \; + \; \gamma FRAME_{ib} \; + \\ & \mathbf{X}_{ib}\Theta \; + \; \mathbf{Z}_b\Psi \; + \; \phi_s \; + \; \epsilon_{ib} \end{aligned} \tag{14}$$

$$\begin{aligned} Y_{ib} = \; & \beta TAX_{ib} \; + \; \delta PUBLIC_b \; + \; \kappa SOCIAL_{ib} \; + \; \gamma FRAME_{ib} \; + \\ & \mu TAX_{ib} \times PUBLIC_b \; + \; \lambda PUBLIC_b \times SOCIAL_{ib} \; + \; \nu PUBLIC_b \times FRAME_{ib} \; + \\ & \mathbf{X}_{ib}\Theta \; + \; \mathbf{Z}_b\Psi \; + \; \phi_s \; + \; \epsilon_{ib} \end{aligned} \tag{15}$$

Where $SOCIAL_{ib}$ is the indicator variable for random assignment to be notified that the mayor would place a sign on their door. Of note, $FRAME_{ib}$ is now the indicator variable for random assignment to also receive additional framing or 'nudges' accompanying the property tax invoice other than the social exposure treatment.

[7] See A for predictions from the model related to mechanisms.

Figure 9: Stickers Indicating Tax Compliance

(a) Sticker design

The center is the seal of the city saying "Mayor of Carrefour" and the edges indicate "Paid" and the corresponding fiscal year "2018-2019".

(b) Example in the field

See middle pane window just to the right of the red door.

Half of all properties in invoice treatment blocs are informed that if they pay their taxes in the 2018-2019 fiscal year, the mayor will affix a sign indicating payment on the outside of their house facing the road. I designed this treatment in partnership with the Mayor's office. For the 2019-2020 fiscal year the design is the same but the color of the background is now red. This paper only presents the effects of being informed of this treatment, in a subsquent companion paper I will present the effects of receiving these stickers as well as the effects of seeing your neighbors receive them on tax compliance.

Table 14, presents the results. Columns 1 and 3 indicate that the Social Exposure treatment does not have a statistically significant additional effect on tax payments either as an interaction or in a linear combination with tax collection. However, Columns 2 and 4 provide evidence that this is because the sign of the social exposure treatment is determined by the interaction (or lack thereof) with public goods. For properties that are only receiving tax collection, the additional Social Exposure treatment is economically and statistically significant causing an additional reduction in tax payments of $2.68 in the first 12 months and $9.23 over the full 18 months. The linear combination with the tax collection effect is also statistically significant as indicated in the lower panels for both Columns 2 and 4. I interpret this as evidence of antisocial punishment and these findings are consistent with A.1.

Where public goods are also received, the Social Exposure effect changes signs and causes a large, and in the case of Column 4 statistically significant, increase in tax payments as an interactive effect. The full linear combination however for both Columns 2 and 4, though positive and large, are imprecise. They provide suggestive evidence that as public goods are introduced, people become more willing and perhaps even interested in having their neighbors know their compliance decision thus mitigating the antisocial punishment mechanism and potentially leading toward the pro-social signaling previously documented

Table 14:
Effects of Social Exposure of Tax Compliance on Amount of Taxes Paid By Property

	Amount of Tax Paid in USD			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−4.01* (2.31)	−5.16* (2.85)	−5.02 (3.76)	−5.11 (5.14)
Public Goods	3.57* (1.97)	0.47 (1.61)	3.18 (2.68)	1.34 (2.64)
Social Exposure	2.56 (3.23)	−2.68** (1.26)	−2.31 (4.39)	−9.23** (4.35)
PUBLIC X TAX		3.19 (4.47)		0.90 (7.40)
PUBLIC X SOCIAL		15.56 (9.77)		19.50* (11.57)
TAX + SOCIAL	-1.45 (3.98)	-7.84 (3.04)	-7.32 (5.05)	-14.34 (5.32)
Full Linear Combination		11.39 (9.48)		7.4 (10.47)
Mean	13.1	13.1	21.86	21.86
Observations	40,060	40,060	40,060	40,060
Adjusted R^2	0.02	0.02	0.03	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. "TAX + SOCIAL" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", and "Social Exposure" equaling zero, and the corresponding clustered standard errors are presented in parentheses. "Full Linear Combination" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of all coefficients shown equaling zero, and the corresponding clustered standard errors are presented in parentheses. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

in the literature.

8.2 Testing Penalties and Obligations on Tax Compliance with Frames

Finally, I am also able to test additional channels motivating tax compliance indirectly with additional frames that I included in the mechanism invoice randomization for the tax collection intervention. I did this by adding brief messages or 'nudges' to the top of a subset of invoices and having the tax agents repeatedly reference the given frame during in-person delivery to the taxpayer (see Appendix F).

I experimentally test the first of the two remaining predictions from my model[8], A.3: Increasing the salience of penalties *decreases* evasion, with a frame making salient the additional 5% financial penalty for each month that one is late in making payment on one's property taxes (Figure F.5). Though this does not resolve the fundamental issue of the lack of state capacity discussed previously, by making salient a more manageable and lesser known penalty of noncompliance, I am able to test if this standard prediction and finding in the literature can hold even in this setting where increased tax collection effort is leading to decreased payments. I then experimentally test the remaining prediction from my model, A.2: Increasing the salience of obligations *decreases* evasion, with a frame making salient one's obligation to pay taxes by using culturally appropriate proverbs that call for sharing and cooperation: 'Many hands make light work' and 'You can't eat soup with just one finger' (Figure F.4).

I then estimate the following:

$$Y_{ib} = \beta TAX_{ib} + \delta PUBLIC_b + \gamma_j \sum_{1}^{J} FRAME_{jib} + \mathbf{X}_{ib}\Theta + \mathbf{Z}_b\Psi + \phi_s + \epsilon_{ib} \tag{16}$$

$$Y_{ib} = \beta TAX_{ib} + \delta PUBLIC_b + \gamma_j \sum_{1}^{J} FRAME_{jib} + \mu TAX_{ib} \times PUBLIC_b + \nu_j PUBLIC_b \times \sum_{1}^{J} FRAME_{jib} + \mathbf{X}_{ib}\Theta + \mathbf{Z}_b\Psi + \phi_s + \epsilon_{ib} \tag{17}$$

Where $FRAME_{jib}$ is now the full set of different indicator variables for random assignment to also receive additional framing or 'nudges' accompanying the property tax invoice. Each j representing a specific frame or 'nudge' and $j = 1$ being the social exposure treatment from

[8] See A for predictions from the model related to mechanisms.

the previous section[9].

Table 15 presents the results which are consistent with my model's predictions. The point estimates throughout for both types of frames are positive and relatively large in Columns 1 and 3. This is also the case in Columns 2 and 4 in the areas that are only receiving tax collection. The interaction with public goods is again shown to be important as the evidence indicates that the nudges are most effective and statistically significant in the areas only receiving tax collection treatment. Though imprecisely estimated, the effect of penalties is reversed in areas receiving public goods. In contrast, the point estimates of the obligations effect remains positive though quite noisy. Furthermore, though penalties mitigate the negative tax collection effect, the linear hypothesis tests in the lower panel indicate that only obligations result in net positive and statistically significant effects. Most notably in Column 4 for areas not receiving public goods, framing tax collection by making salient cultural appeals to obligations caused an increase in tax payments of \$9.82 or about 40% more than the mean of \$21.86. These findings also provide further evidence supporting the conditions on Prediction 1: both that public goods and the salience of penalties mitigate the expected negative effect of increased tax collection in this setting and also that the crowding-out can be reversed with a sufficiently strong appeal to obligations.

[9] For results testing the remaining frames included in this study, please see Table 30 and examples discussed starting in F.

Table 15:
Effects of Framing Tax Collection on Amount of Taxes Paid By Property

	Amount of Tax Paid in USD			
	FY19		FY19-20	
	(1)	(2)	(3)	(4)
Public Goods	3.6* (2.0)	0.5 (1.6)	3.3 (2.7)	1.3 (2.6)
Tax Collection	−3.9* (2.3)	−5.1* (2.8)	−4.3 (3.7)	−4.5 (5.0)
Penalties	4.0 (2.9)	6.4** (3.0)	6.0 (4.2)	8.7* (4.7)
Obligation	12.9 (8.2)	9.3* (4.8)	17.2* (8.8)	14.3** (7.1)
Public X TAX		3.4 (4.4)		1.1 (7.4)
Public X PENALTIES		−5.9 (5.8)		−6.4 (8.2)
Public X OBLIGATION		10.3 (22.6)		8.2 (23.3)
TAX + PENALTIES	0.01 (2.29)	1.3 (2.12)	1.65 (3.42)	4.21 (4.04)
PENALTIES Linear Combination		-0.77 (4.29)		0.21 (5.67)
TAX + OBLIGATION	8.93 (7.51)	4.12 (2.81)	12.85 (7.98)	9.82 (4.62)
OBLIGATION Linear Combination		18.25 (20.03)		20.42 (20.15)
Mean	13.1	13.1	21.86	21.86
Observations	40,060	40,060	40,060	40,060
Adjusted R^2	0.02	0.02	0.03	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. "TAX + COST" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", and "Cost of Evasion" equaling zero, and the corresponding clustered standard errors are presented in parentheses. "COST Linear Combination" presents the estimated coefficient of all coefficients contributing to the full interaction effect of PUBLIC, TAX and COST equaling zero, and the corresponding clustered standard errors are presented in parentheses. "TAX + OBLIGATION" presents the estimated coefficient of "Tax Collection", and "Obligation" equaling zero, and the corresponding clustered standard errors are presented in parentheses. "OBLIGATION Linear Combination" presents the estimated coefficient of all coefficients contributing to the full interaction effect of PUBLIC, TAX and OBLIGATION equaling zero, and the corresponding clustered standard errors are presented in parentheses. All regressions also control for the remaining $j-2$ frames separately tested. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

9 Discussion: Costs and Benefits

The cost effectiveness of implementing these interventions is also relevant for policy. The tax collection intervention was approximately $2 per delivery inclusive of printing, transport and labor. Per Column 1 of Table 4, each delivery resulted in approximately $6.29 of lost tax payments in the first year. This points to an initial net loss of approximately $8.29 per taxpayer. The Social Exposure treatment in areas not receiving public goods further compounded the losses to a total of $9.84 per taxpayer - equivalent to losing about 3/4 of property tax revenue for the city or nearly $400,000 if this least effective approach had been universally implemented.

One variation on the tax collection implementation was able to mitigate the negative effects and resulted in a gain even in the absence of public goods. Per Column 4 of Table 15 and taking into account the costs of implementation mentioned above, making obligations salient resulted in an average net increase in revenue of $7.82 per delivery over the first 18 months of data collected thus far. This is a 35% increase in revenue per taxpayer. If the city had implemented a well-framed tax collection intervention like this one, the results suggest the potential for an average of $210,000 in additional revenue in the first year.

The public goods intervention has cost about $3,000 per month for maintenance, fuel, and labor. This is equivalent to approximately $2.30 per property per year in variable costs - significantly lower than the average property tax bill of about $90, and not even 1/5 the average annual tax payment of $13.10. Returning to Column 1 of Table 4, this suggests that public goods resulted in a net increase of $1.25 per property in the first year. This implies the treatment may have been able to net more than $50,000 in increased returns had it been implemented across the entire city. However, much of that positive estimated effect came from mitigating the negative tax collection effect in the interaction. Taking instead the point estimate from only public goods in Column 1 of Table 5, then just in terms of variable costs the intervention would have lost more than $70,000 for the city in that first year.

This is all before considering the initial start-up costs of implementation - specifically the 6 vehicles donated at the start of the experiment. Each of these would cost on the order of $250,000 if purchased new for a total of $1.5 million. However, the vehicles that were donated had been used by the sister city in Martinique for several years. It is possible that had the Mayor attempted to purchase something equivalent on the secondary market he could have done so for something closer to $50,000 per vehicle or $300,000 total. Furthermore, with better management and more efficient routes, I believe that it could be possible to accomplish most of what the current team is managing with even just one well-functioning vehicle which would suggest a minimum investment on the order of $50,000. Assuming that the resulting increases in tax revenue are something closer to my preferred specification which would be in keeping with my findings in Table 12 that this is an underestimate. The result from Table **??** testing the intensity of public goods also indicated that an additional garbage collection route across the city, which generally takes two weeks, would yield between 170% and 240% of the monthly implementation costs. Even before accounting for the potential savings of running more efficient, non-randomized routes, these findings imply a plausible pathway for

sufficiently increasing net revenue to repay the initial start-up investment by the end of the first year of implementation.

Even if the intervention were a net accounting loss for the government in terms of revenue directly raised for money directly spent, the investment in public goods would likely still result in a net benefit for several reasons. As discussed in 4.2, the primary predictor for paying taxes this year is having paid taxes last year. Public goods may then be able to move more people to pay, and pay more, thus establishing a habit of payment potentially leading to positive revenue over time. Furthermore, evidence from Table 14 indicated that public goods interacted with Social Exposure of tax compliance to mitigate the antisocial punishment mechanism that likely is contributing to the stability of the low-tax equilibrium. Furthermore, suggestive evidence indicated that public goods may also shift the effect of Social Exposure to one of pro-social signaling. Lastly, Tables 6-8 as well as **??** all indicate public goods decrease or mitigate political unrest including the more violent and destructive actions. Given that my study city was completely shut down by protests for the equivalent of 3 months over the 24 months of implementation, periods of unrest come at significant financial costs for the citizens and the government. So insofar as public goods provision can promote pro-social behaviors resulting in increased revenue while reducing the productive losses due to protesting, they will likely generate net benefits - and this is before taking into account the potential welfare gains from public goods provision in this context.

10 Conclusion

The literature has indicated state capacity building to be a central challenge that all developed countries historically had to overcome and poor countries today must face. It requires raising revenue without inciting political backlash. Two tools discussed by the literature and available to policy makers that are relevant in this context are increased tax collection effort and public goods provision. In this paper, informed by qualitative research and economic theory, I provide evidence through an RCT at-scale and lasting multiple years covering the urban area of one of the largest cities in one of the world's weakest states. I cross-randomize both tax collection and public goods in the same setting and follow results in administrative data, panel surveys and through novel metrics of political unrest that I introduce.

Contrary to the established literature from more developed settings, I find evidence of statistically and economically significant negative effects on individual tax payments caused by increased tax collection effort. This is consistent with theory and my modeling predictions based on a setting wherein at baseline compliance levels are low and there are few-if-any external incentives. I also find that increasing the probability of social exposure of tax compliance due to a novel treatment I introduce further reduces payments to almost zero. This second finding is consistent with antisocial punishment, a new mechanism from psychology that I bring to this literature. I find that tax collection also increases political unrest across several measures: both the prevalence of political graffiti as well as the negativity of its tone, and participation in political protests, especially the more destructive and violent types of unrest resulting in the construction of protest barricades which are often then set on fire.

I find that effectively framing tax collection interventions with appropriate 'nudges' that make salient key elements of the tax compliance decision can mitigate the negative tax collection effect on tax revenue and even increase tax payments. Most notably, making culturally appropriate appeals to obligations in this setting significantly increases revenue.

With respect to public goods, I find they cause statistically and economically significant increases in tax payments while also reducing political unrest, especially barricade construction during protests. I find further evidence that public goods mitigate many of the adverse tax collection effects both on tax revenue as well as on protest behavior. I also find that the presence of public goods changes the sign of the effects from social exposure of tax compliance potentially leading to pro-social signaling.

Though my findings provide a cautionary tale for governments seeking to increase tax revenue, an accounting exercise on the interventions suggests that both types, if well managed and adapted, may be able to generate positive revenue for the government and even potentially recover initial investments. However, with the lack of financing available in these settings, it is likely that international aid, multilateral or bilateral lending or even intragovernment transfers may be required to provide short-term initial investments.

Open questions and next steps for this research lie in identifying what new equilibrium for tax compliance is possible as well as identifying if and under what circumstances these effects on political unrest caused by tax collection might lead to increased democratic accountability. Most directly, I will be following this and other governments as they attempt to leverage public goods provision to balance purse and peace.

References

Acemoglu, Daron, "Politics and economics in weak and strong states," *Journal of Monetary Economics*, 2005, *52* (7), 1199–1226.

Allingham, Michael G. and Agnar Sandmo, "Income tax evasion: a theoretical analysis," *Journal of Public Economics*, 1972.

Bates, Robert H. and Da-Hsiang Donald Lien, "A Note on Taxation, Development, and Representative Government," *Politics & Society*, 1985.

Becker, Gary S, *Crime and Punishment : An Economic Approach*, Vol. I 1974.

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

Bernheim, B Douglas and Dmitry Taubinsky, "Behavioral Public Economics," in "Behavioral Economics" 2018.

Besley, Timothy, "State capacity, reciprocity and the social contract," *Econometrica*, 2020, *88* (4), 1307–1335.

_ **and Torsten Persson**, "The origins of state capacity: Property rights, taxation, and politics," *American Economic Review*, 2009, *99* (4), 1218–1244.

_ **and** _ , "The logic of political violence," *Quarterly Journal of Economics*, 2011, *126* (3), 1411–1445.

_ **and** _ , "Taxation and Development," in Alan Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, eds., *Handbook of Public Economics*, Vol. 5 2013, pp. 51–110.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi, "Sources of Variation in State Effectiveness and Consequences for Policy Design *," 2016.

Blattman, Christopher, Donald Green, Daniel Ortega, and Santiago Tobón, "Pushing Crime Around the Corner? Estimating Experimental Impacts of Large-Scale Security Interventions," *NBER Working Paper Series*, 2017, p. 29.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini, "The Political Resource Curse The Political Resource Curse," *American Economic Review*, 2013, *103* (5), 1759–1796.

Buchanan, James M, *The Power to Tax: Analytical Foundations of a Fiscal Constitution,* 1980.

Bursztyn, Leonardo, Michael Callen, Bruno Ferman, Ali Hasanain, and Noam Yuchtman, "Political Identity : Experimental Evidence on Anti-Americanism in Pakistan," 2017.

Cantoni, Davide, David Yang, Noam Yuchtman, and Y. Jane Zhang, "Are Protests Games of Strategic Complements or Substitutes? Experimental Evidence from Hong Kong's Democracy Movement," 2017.

Carrillo, Paul, Dina Pomeranz, and Monica Singha, "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," *Quarterly Journal of Economics*, 2016, (September), 273–314.

Castro, Lucio and Carlos Scartascini, "Tax compliance and enforcement in the pampas evidence from a field experiment," *Journal of Economic Behavior and Organization*, 2015, *116* (December), 65–82.

Cohen, Molly, Noam Lupu, and Elizabeth Zechmeister, "The Political Culture of Democracy in the Americas, 2016/17: A Comparative Study of Democracy and Governance," 2017.

De Ferrari, Guillermina, ""Le poème tué": Public Poetry as Political Praxis," 2020.

Dubois, Laurent, *Haiti: The Aftershocks of History*, first edit ed., New York, NY: Metropolitan Books, 2012.

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke, "Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in Germany," *American Economic Journal: Economic Policy*, 2016, *8* (3), 203–232.

Erard, B and J S Feinstein, "The Role of Moral Sentiments and Audit Perceptions in Tax Compliance," *Public Finance/Finances Publiques*, 1994, *49* (August), 70–89.

Farmer, Paul, *The Uses of Haiti*, third edit ed., Monroe, Maine: Common Courage Press, 2005.

Frey, Bruno S., "A Constitution for Knaves Crowds out Civic Virtues," *The Economic Journal*, 1997, *107* (443), 1043–1053.

Fukuyama, Francis, *State-Building: Governance and World Order in the 21st Century*, Ithaca, NY: Cornell University Press, 2004.

Girard, Philippe, *Haiti: The Tumultuous History - From Pearl of the Caribbean to Broken Nation*, first edit ed., New York, NY: St. Martin's Griffin, 2010.

Hallsworth, Michael, "The use of field experiments to increase tax compliance," *Oxford Review of Economic Policy*, 2014, *30* (4), 658–679.

Hernandez, Marco, Julian Jamison, Ewa Korczyc, Nina Mazar, and Roberto Sormani, "Applying Behavioral Insights to Improve Tax Collection: Experimental Evidence from Poland," Technical Report, World Bank Group 2017.

Herrmann, Benedikt, Christian Thoni, and Simon Gachter, "Antisocial Punishment Across Societies," *Science*, 2008, *319*, 1362–1367.

Holz, Justin, John A. List, Alejandro Zentner, Marvin Cardoza, and Joaquin Zentner, "The $100 Million Nudge: Increasing Tax Compliance of Businesses and the Self-Employed Using a Natural Field Experiment," 2020.

Huntington, Samuel P., "The Third Wave: Democratization in the Late 20th Century," *Journal of Democracy*, 1991.

Janvry, Alain De and Jean-Jacques Dethier, "The World Bank and Governance: The Bank's Efforts to Help Developing Countries Build State Capacity," *World Bank Policy Research Working Paper*, 2012, (November).

Kettle, Stewart, Mateo Hernandez, Simon Ruda, and Michael Sanders, "Behavioral Interventions in Tax Compliance - Evidence from Guatemala," *World Bank Policy Research*, 2016, (June), 1–40.

Khan, Adnan, Asim Khwaja, and Benjamin A Olken, "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics*, 2015, pp. 1–53.

Khan, Adnan Qadir, Asim Ijaz Khwaja, and Benjamin A. Olken, "Making Moves Matter : Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings," *American Economic Review*, 2019, *109* (1), 237–270.

_ , _ , _ , Osman Haq, and Mahvish Shaukat, "Rebuilding the Social Compact: Urban Service Delivery and Property Taxes in Pakistan," Technical Report April, 3ie Impact Evaluation Report 117, New Delhi 2020.

Kirchler, Erich, *The economic psychology of tax behaviour* 2007.

Levi, Margaret, "Of Rule and Revenue," in "Of Rule and Revenue" 1988.

Luttmer, Erzo F P and Monica Singhal, "Tax Morale," *Journal of Economic Perspectives*, 2014, *28* (4), 149–168.

Maguire, Robert and Scott Freeman, eds, *Who Owns Haiti? People, Power and Sovereignty*, Gainesville, FL: University Press of Florida, 2017.

Manacorda, Marco, Edward Miguel, and Andrea Vigorito, "Government transfers and political support," *American Economic Journal: Applied Economics*, 2011, *3* (3), 1–28.

Martin, Lucy, "Taxation and Accountability : Experimental Evidence for Taxation's Effect on Citizen Behavior." PhD dissertation 2013.

Morrison, Kevin, "Oil, Non-Tax Revenue, and Regime Stability: The Political Resource Curse Reexamined," 2005.

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.

Pande, Rohini, "Can democracy work for the poor?," *Science (New York, N.Y.)*, 2020, *369* (6508), 1188–1192.

Pomeranz, Dina, "No taxation without information: Deterrence and self-enforcement in the value added tax," *American Economic Review*, 2015, *105* (8), 2539–2569.

_ **and Jose Vila-Belda**, "Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities," *Annual Review of Economics*, 2019, pp. 1–39.

Reckers, Philip M J, Debra L Sanders, and Stephen J Roark, "the Influence of Ethical Attitudes on Taxpayer Compliance," *National Tax Journal*, 1994.

Ross, Michael L., "Does Taxation Lead to Representation?," *British Journal of Political Science*, 2004.

Shimeles, Abebe, Daniel Zerfu Gurara, and Firew Woldeyes, "Taxman's dilemma: Coercion or persuasion? Evidence from a randomized field experiment in Ethiopia," *American Economic Review*, 2017, *107* (5), 420–424.

Slemrod, Joel, "Tax Compliance and Enforcement," *Journal of Economic Literature*, 2019, *57* (4), 904–954.

_ **, Obeid Ur, and Rehman Mazhar Waseem**, "Pecuniary and Non-Pecuniary Motivations for Tax Compliance: Evidence From Pakistan," 2019.

Srinivasan, T. N., "Tax evasion: A model," *Journal of Public Economics*, 1973.

Thomas, M A, *Govern Like Us: U.S. Expectations of Poor Countries*, first edit ed., New York, NY: Columbia University Press, 2015.

Tilly, Charles, "States and their Citizens," in "Coercion, Capital, and European States, AD 990-1990" 1990.

Torgler, Benno, "Tax morale and direct democracy," *European Journal of Political Economy*, 2005, *21* (2), 525–531.

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

Appendices

A Secondary Model Predictions

The model as presented in Equation 1, in addition to the predictions for primary research questions regarding tax compliance as presented in 3.2, also yields predictions which I test when exploring mechanisms in Section 8. Turning to the effect on increasing social exposure, q, on tax evasion, e. Note that in a low-tax compliance setting, unlike as is generally modeled in the literature (Shimeles et al., 2017), the social incentive here is assumed to be strictly negative $\Sigma < 0$. This is consistent with antisocial punishment mechanism wherein when most are not following the rules, desire for social conformity will incentivize being perceived as a rule breaker (Herrmann et al., 2008).

$$\frac{\partial e}{\partial q} = -\frac{\Sigma}{2\rho} > 0 \tag{18}$$

which implies:

A.1 Increasing social exposure of tax compliance *increases* evasion.

However, the presence of a mechanism like antisocial punishment in the model does not preclude norms of tax compliance also playing a role in decreasing evasion. Indeed, the sense of obligation, ρ, functions like this and is the only incentive for tax compliance in the most extreme low-capacity case of no enforcement, no credibility and no public goods. As a result, even in the general case, we have the following:

$$\frac{\partial e}{\partial \rho} = -\frac{e}{\rho} < 0 \tag{19}$$

Though this is not something I will be able to directly test, I will make salient this obligation in a variety of ways when exploring mechanisms, and these comparative statics imply:

A.2 Increasing the salience of obligations *decreases* evasion.

Finally, above I make the simplifying assumption that in a low-compliance setting the level of perceived fine is zero, $f = 0$, due to the state's lack of capacity and credibility. However, the model also predicts that efforts to increase the perceived fine, even if starting from essentially zero, should decrease evasion. This can be seen in:

$$\frac{\partial e}{\partial f} = -\frac{\tau}{2\rho} < 0 \tag{20}$$

This similarly lends itself to making salient the fine so to test the prediction:

A.3 Increasing the salience of penalties *decreases* evasion.

B Additional Figures

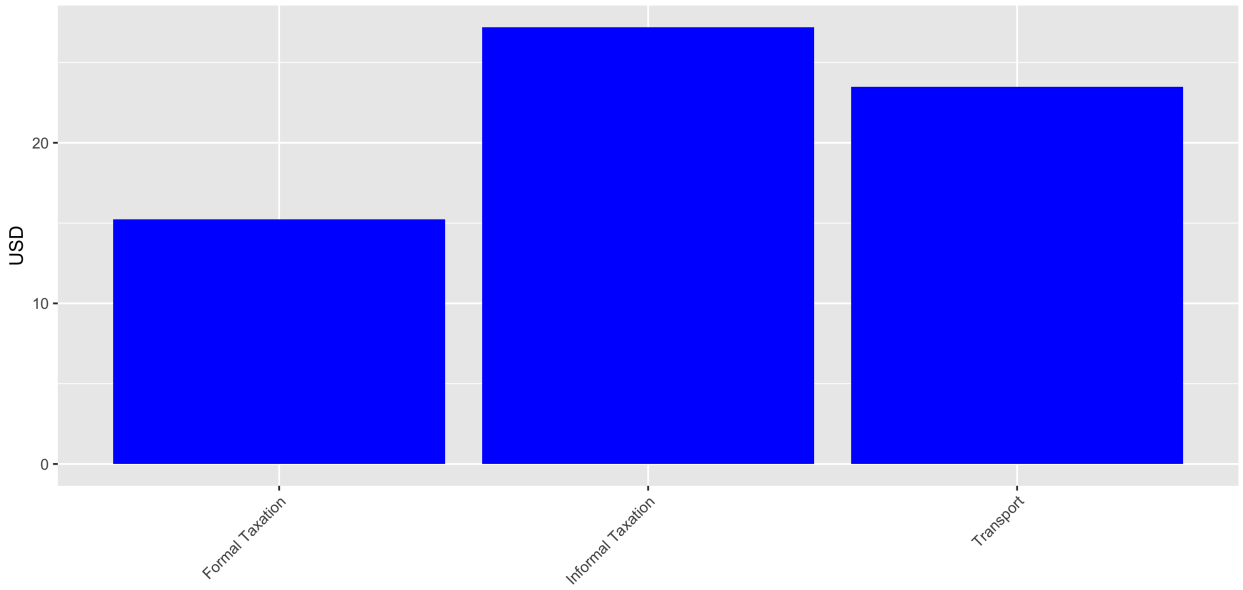


Figure 10: Formal vs Informal Taxation

Contributions at baseline to formal and informal taxation with transport expenditures also provided as a benchmark. Baseline self-reported measures from 1,500 representative households. All amounts in USD per month.

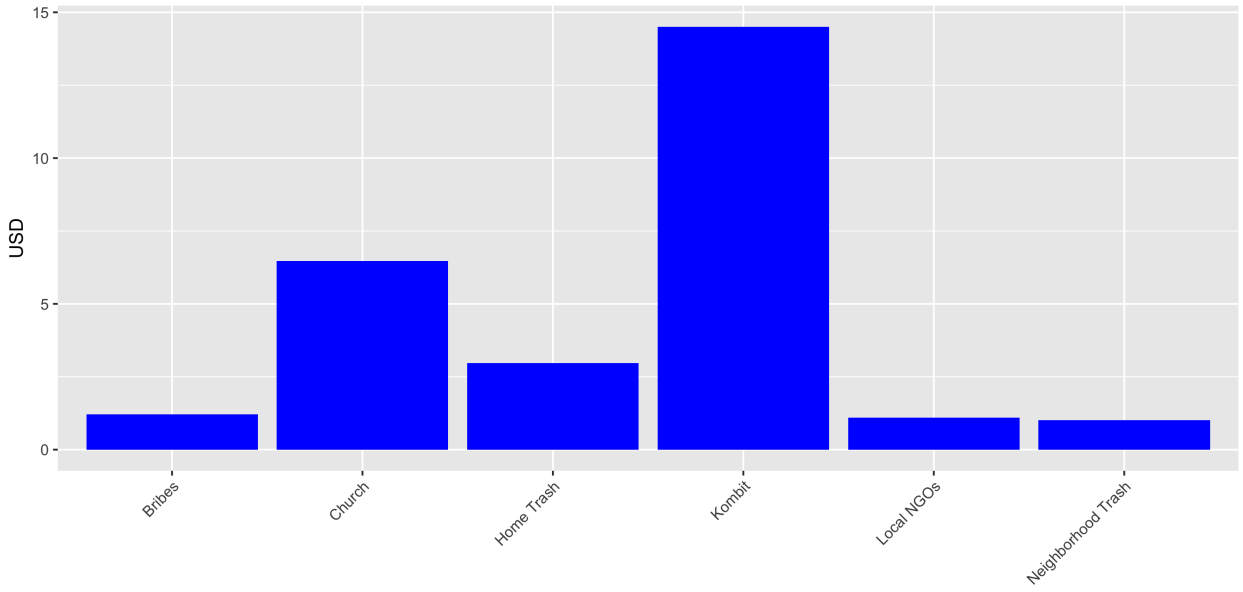


Figure 11: Composition of Informal Taxation

Composition of contributions to informal taxation. "Kombit" is the combination of all community-based collective action activities. Value of in-kind contributions estimated by survey respondents. Value of labor contributions calculated by multiplying time worked by $\frac{1}{2}$ national minimum wage. Baseline self-reported measures from 1,500 representative households. All amounts in USD per month.

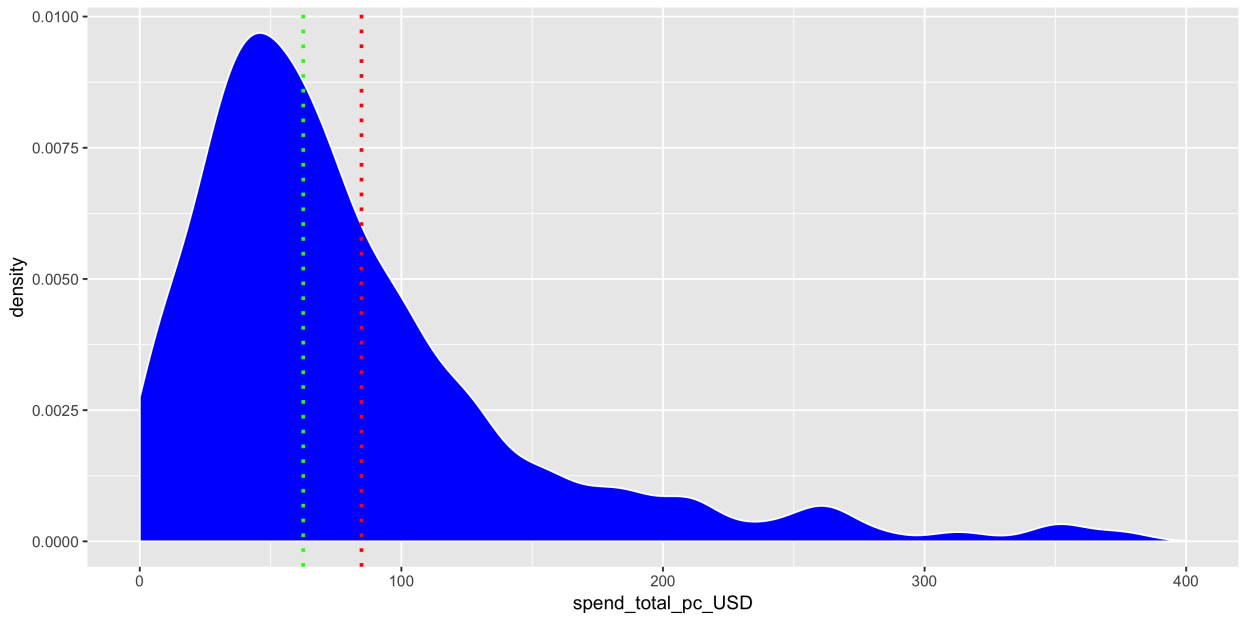


Figure 12: Monthly per capita Expenditures

Self-reported household monthly per capita expenditures in USD as reported by the head of household in aggregate and then divided by the number of household members. Green dotted line is the median and the red dotted line is the mean.

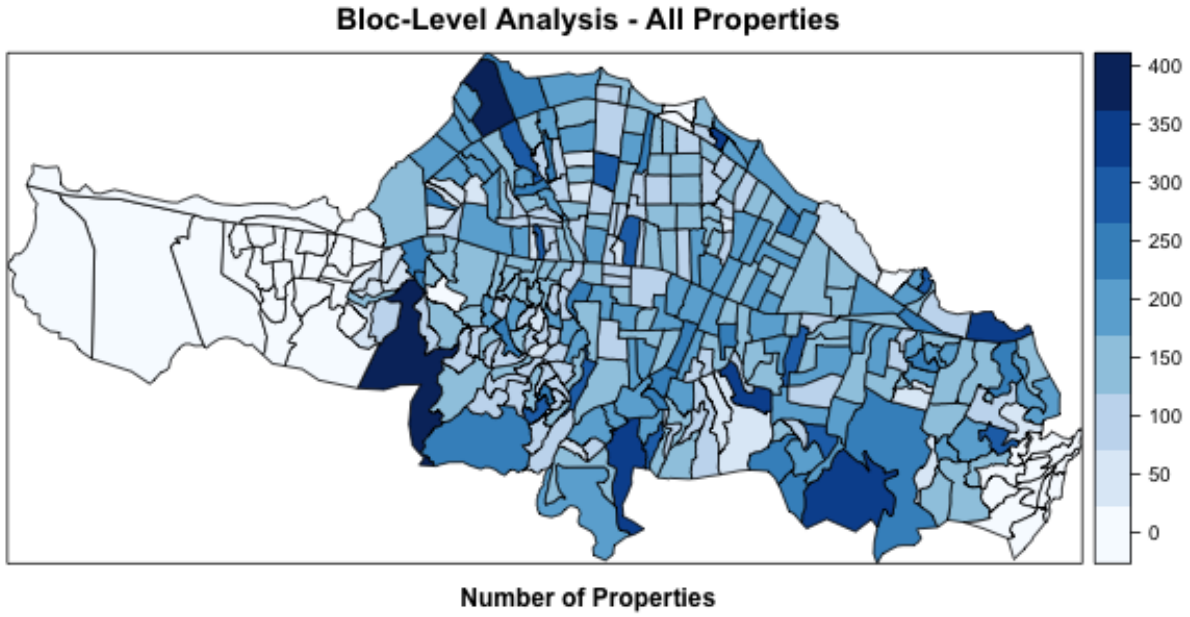


Figure 13: Universe of Properties by Bloc

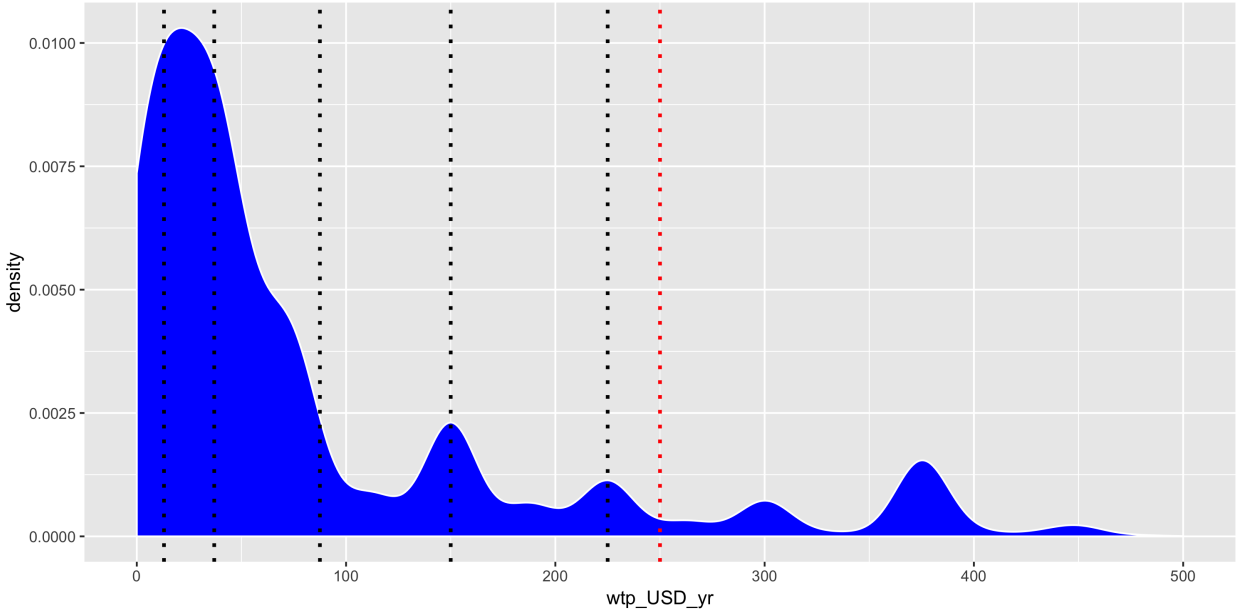


Figure 14: Willingness to Pay for garbage removal Services

The results of an incentive compatible Becker–DeGroot–Marschak willingness to pay solicitation for garbage removal services. X-axis indicates values in terms of USD. Dotted lines indicate the equivalent amounts for each of the various steps in the property tax schedule.

C Baseline Data and Modeling Assumptions

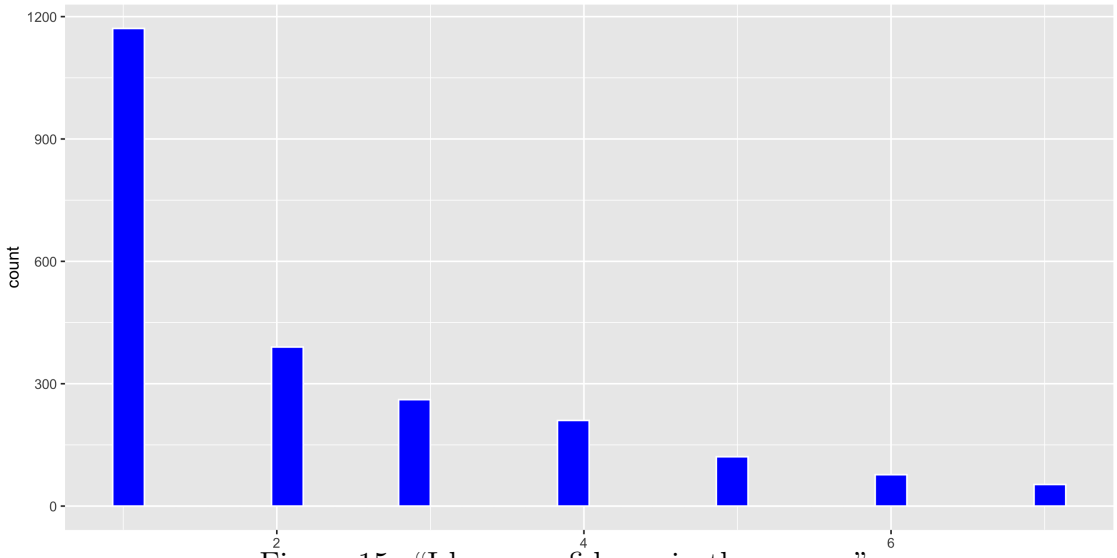


Figure 15: "I have confidence in the mayor."

1 = Completely Disagree 7 = Completely Agree

Baseline Survey Responses from a representative sample of households. Respondents were asked to indicate how strongly they agreed with the statement "I have confidence in the mayor" on a 7-point scale with 1 indicating "Completely Disagree" and 7 corresponding to "Completely Agree".

My data collection further support the intuition discussed in Section 3 that incentives to pay taxes in this setting are quite different than has generally been modeled in the literature. I conducted a series of focus group discussions followed by a baseline survey of a representative sample of households prior to the start of interventions. Responses are consistent with the modeling assumptions I have made regarding the importance of state capacity, civic virtue, public goods and social incentives for optimizing the evasion decision.

Respondents repeatedly highlighted the absence of the state in their lives and their limited trust in any state institutions or officials. As an example, survey responses revealed the most trusted public official of those asked about to be the mayor, but still a large majority of people 'completely disagree' with the statement, 'I have confidence in the mayor' (Figure 15).

Focus groups and survey respondents consistently were able to identify the potential penalties for noncompliance - the most significant being property seizure. However, none reported having experienced or having heard of others experiencing this penalty. This is consistent with the Mayor's own position at the start of this study that even if he had the capacity to enforce the prescribed penalties (which he indicated to not be the case), because so few people pay, any attempt to enforce the punishment would be perceived as arbitrary or as politically motivated retribution. Officials reported that the frequency of protests and concerns of civil unrest prevent the mayor's office from taking any action in this regard. The dynamic of political protesting limiting the lawful action of the mayor's office was directly observed on several occasions during this study, most notably in the repeated closure of the

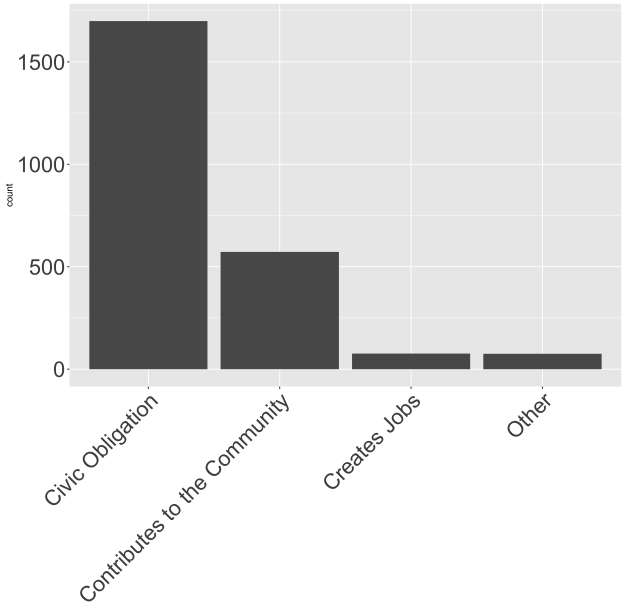


Figure 16: "Why do you believe your property taxes are just?"
Baseline Survey Responses from a representative sample of households. Respondents responding "yes" when asked, "Do you believe your property taxes are just?" were asked the open ended follow-up question, "Why do you believe your property taxes are just?"

city dump due to protests by those living in close proximity to it. This is consistent with my modeling assumption of the perceived fine, f, being very small or even zero in this context.

At the same time, a large majority of survey respondents indicated a sense of civic obligation toward paying property taxes (Figure 16). The lack of external incentives and the strong internal motivation to pay are consistent with the set-up outlined by Frey (1997) and formalized in my model which anticipates a crowding-out effect from increased tax collection efforts.

With respect to public goods in my study setting, consistent with information from the mayor's office and observed in the field, survey respondents indicated that there were essentially no services on offer, and they were not satisfied with this level of public goods (Figure 17). Furthermore, the lack of public goods provided is the most common reason given for the perceived lack of fairness of property taxes (Figure 18). This is consistent with the level of public goods determining the perceived fairness of the tax, and tax compliance decreasing as public goods are increased.

Finally, focus group discussions indicated an aversion to discussing tax compliance with neighbors, particularly if one had paid one's taxes. Survey responses were more polarized with most respondents divided between the extreme positions of "Completely Disagree" and "Completely Agree" in response to the statement "If I paid my taxes, I would want my neighbors to know" (Figure 19). Though this evidence does not fully support the modeling assumption of the social incentive depressing tax compliance, it does question the more common assumption that social exposure unambiguously encourages tax compliance.

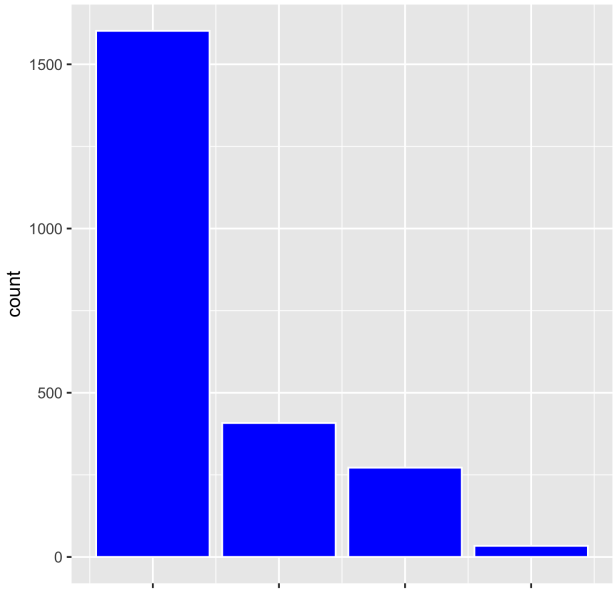


Figure 17: "How satisfied are you with local government services?"
1= Not Satisfied at All 4 = Completely Satisfied
Baseline Survey Responses from a representative sample of households. Respondents were asked, "How satisfied are you with local government services? Not at All Satisfied, Somewhat Not Satisfied, Somewhat Satisfied, Completely Satisfied"

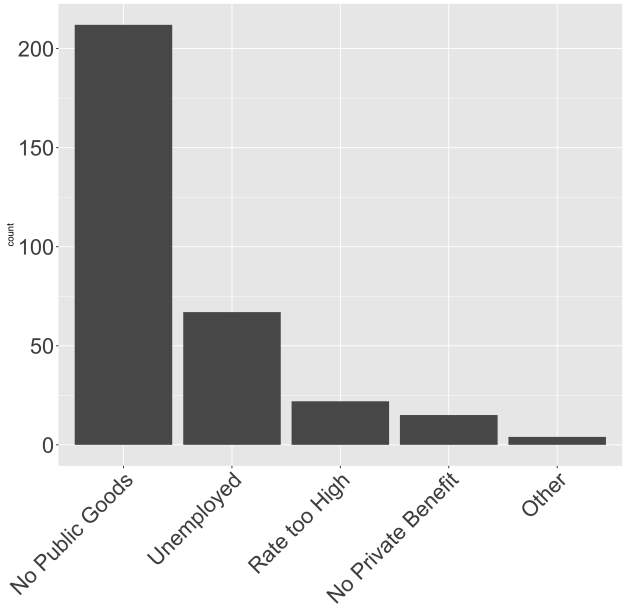


Figure 18: "Why do you believe your property taxes are NOT just?"
Baseline Survey Responses from a representative sample of households. Respondents responding "no" when asked, "Do you believe your property taxes are just?" were asked the open ended question, "Why do you believe your property taxes are not just?"

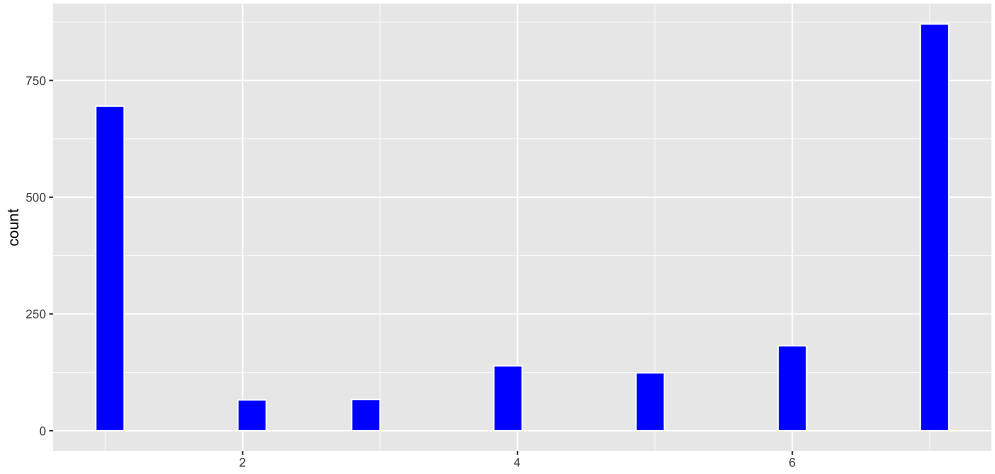


Figure 19: "If I paid my taxes, I would want my neighbors to know."
1 = Completely Disagree 7 = Completely Agree
Baseline Survey Responses from a representative sample of households. Respondents were asked to indicate how strongly they agreed with the statement "'If I paid my taxes, I would want my neighbors to know." on a 7-point scale with 1 indicating "Completely Disagree" and 7 corresponding to "Completely Agree".

D Placebo Graffiti Tests

Figure 20:
Social Movement Graffiti

Examples of the principal message for a social movement calling for an end to violence in the country. Roughly translated as, “The Killed Poem”, I use this message as one of the placebo tests for my analysis of the effects of tax collection and public goods on political graffiti.

Table 16:
Placebo Tests for Effects of Tax Collection, Public Goods and their Interaction on Graffiti

	Graffiti			
	Any		Social Movement	
	(1)	(2)	(3)	(4)
Tax Collection	−0.011	0.022	−0.048	−0.070
	(0.051)	(0.066)	(0.055)	(0.068)
Public Goods	−0.026	0.016	−0.048	−0.075
	(0.052)	(0.072)	(0.050)	(0.070)
PUBLIC X TAX		−0.084		0.055
		(0.105)		(0.105)
Linear Combination		-0.05 (0.08)		-0.09 (0.08)
Mean	0.75	0.75	0.2	0.2
F-Stat	6.78	6.1	2.91	2.64
Observations	241	241	241	241
Adjusted R^2	0.191	0.190	0.068	0.065

"Any" is a binary indicator equal to 1 if any graffiti was identified in the bloc and zero otherwise. "Social Movement" is a binary indicator equal to 1 if a message identified as part of a nascent social movement calling for an end to violence was identified in the bloc and zero otherwise. "Tax Collection" is an indicator variable equal to 1 if a bloc is randomly assigned to have tax agents from the mayor's office hand-deliver invoices for property taxes in the bloc and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a census bloc is randomly assigned to receive garbage removal services and equal to zero otherwise. "Linear Combination" presents the estimated coefficient of a two-sided general linear hypotheses test of the linear combination of the coefficients on "Tax Collection", "Public Goods", and "PUBLIC X TAX" equaling zero, and the corresponding clustered standard errors are presented in parentheses. Analysis is limited to only census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. Columns 3 and 4 control for propensity for political graffiti with an indicator variable equal to 1 if messages referring to the previous election completed prior to the start of interventions was identified in the bloc and zero otherwise. All regressions also include additional bloc controls and strata fixed effects.

During the graffiti census a strange message continued to turn-up, ultimately comprising nearly 10% of the total amount of graffiti in the city. The message, most commonly in the form"Le Poeme Tue" or "La Religion Tue" ("The Killed Poem" or "The Killed Religion") was not known to any of my field team and turned out to be a nascent nonpartisan social movement calling for the end of violence in the country. I first test all graffiti as a placebo in Columns 1-2 of Table 16 and find no statistically significant results. Then, given the movement's national focus and the orthogonal relationship of its aim and the interventions of this study, I use these specific graffiti as an additional placebo test in Columns 3-4, and I do not find any evidence of effects for either specification.

E Supplemental Tables

Table 17:
First Stage Estimates for Instrumenting Public Goods Intensity

	First Stage Regression to Instrument Public Goods Intensity for Models:			
	Taxes Paid	Political Graffiti	Barricades	Access Blocked
	(1)	(2)	(3)	(4)
Tax Collection	−0.81	−0.72	−0.72	−0.72
	(1.01)	(1.15)	(1.16)	(1.16)
Public Good Intensity	27.01***	26.26***	26.43***	26.43***
	(1.25)	(1.31)	(1.28)	(1.28)
Mean	12.86	13.08	13.08	13.08
F-Stat	72.83	73.71	83.3	83.3
Observations	40,060	241	241	241
Adjusted R^2	0.66	0.69	0.69	0.69

All regressions have "public goods Intensity" as the outcome of interest. The model labeled "Taxes Paid" is the first stage for the second stage estimating "Amount of Tax Paid USD in FY19'. The models labelled "Political Graffiti", "Barricades" and "Access Blocked" each are the first stages for the second stages estimating the outcome of interest of the same name. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a census bloc is randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 18:
Testing for Spillovers of Pubic Goods on Amount of Taxes Paid By Property by Dropping 'Contaminated' Properties

	Amount of Tax Paid in USD			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−3.91 (3.26)	−6.16 (4.23)	−4.07 (4.52)	−6.69 (6.33)
Public Goods	5.32** (2.53)	2.03 (1.88)	5.21* (3.16)	3.25 (2.73)
PUBLIC X TAX		5.55 (4.47)		6.10 (6.84)
Mean	12.82	12.82	21.22	21.22
Drop all boundary properties	Yes	Yes	Yes	Yes
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.03	0.03	0.05	0.05

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 19:
Effects of Tax Collection, Public Goods and their Interaction on log(1 + Amount of Tax Paid in USD by Property)

	log(1 + Amount of Tax Paid in USD)			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−0.085*** (0.027)	−0.066* (0.036)	−0.083*** (0.032)	−0.061 (0.045)
Public Goods	0.020 (0.020)	0.056* (0.029)	0.026 (0.025)	0.061* (0.036)
PUBLIC X TAX		−0.048 (0.055)		−0.056 (0.064)
Linear Combination		-0.058 (0.037)		-0.056 (0.04)
Drop all boundary properties	Yes	Yes	Yes	Yes
F-Stat	268.5	216.37	297.04	245.63
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.257	0.258	0.293	0.293

"log(1 + Amount of Tax Paid in USD)" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG), then multiplying by an exchange rate of 65HTG:1USD, adding one (1) and then taking the log() of the result. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 20:
Effects of Tax Collection, Public Goods and their Interaction on Change in Amount of Tax Paid from FY 2018 in USD by Property

	Change in Amount of Tax Paid from FY 2018 in USD			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−3.49 (5.46)	−2.92 (7.42)	−7.41 (7.33)	−9.09 (9.20)
Public Goods	6.77** (2.89)	2.48 (1.89)	12.10** (5.15)	4.52 (3.54)
PUBLIC X TAX		−0.40 (6.10)		5.16 (8.24)
Linear Combination		-0.83 (4.17)		0.59 (6.82)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	1.2	1.2	14.02	14.02
F-Stat	15.88	16.31	6.52	5.61
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.03	0.03	0.003	0.003

"Change in Amount of Tax Paid from FY 2018 in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then subtracting from that the sum of all property tax payments for a given property made in FY2018, and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 21:
Effects of Tax Collection, Public Goods and their Interaction on log(1+ Change in Amount of Tax Paid from FY 2018 by Property)

	log(1 + Change in Amount of Tax Paid from FY 2018)			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	-0.085^{***} (0.024)	-0.068^{**} (0.032)	-0.090^{***} (0.029)	-0.078^{**} (0.040)
Public Goods	0.015 (0.017)	0.036 (0.026)	0.018 (0.022)	0.051 (0.031)
PUBLIC X TAX		−0.041 (0.048)		−0.032 (0.058)
Linear Combination		-0.07 (0.03)		-0.06 (0.04)
Drop all boundary properties	Yes	Yes	Yes	Yes
F-Stat	100.882	78.894	451.28	364.44
Observations	26,099	26,099	26,339	26,339
Adjusted R^2	0.074	0.074	0.322	0.322

"log(1+ Change in Amount of Tax Paid from FY 2018)" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then subtracting from that the sum of all property tax payments for a given property made in FY2018, adding one (1), and then taking the log() of the result. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 22:
Effects of Tax Collection, Public Goods and their Interaction on Change in Amount of Tax Paid from Average of Previous Three Years in USD by Property

	Change in Amount of Tax Paid from Average of Previous Three Years			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−6.79 (6.97)	−8.35 (8.69)	−10.70 (8.73)	−14.51 (10.84)
Public Goods	7.56** (3.28)	2.05 (2.00)	12.89** (5.40)	4.09 (3.57)
PUBLIC X TAX		4.49 (5.49)		10.05 (8.40)
Linear Combination		-1.81 (5.22)		-0.38 (7.44)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	1.9	1.9	14.71	14.71
F-Stat	7.32	7.98	7.89	6.5
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.05	0.05	0.01	0.01

"Change in Amount of Tax Paid from Average of Previous Three Years in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then subtracting from that quantity of the sum of all property tax payments for a given property made for each of the previous three fiscal years prior to the start of the project (FY2016, FY2017 and FY2018) divided by 3, and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 23:
Effects of Tax Collection, Public Goods and their Interaction on log(1 +Change in Amount of Tax Paid from Average of Previous Three Years in USD by Property)

	log(1 + Change in Amount of Tax Paid from Average of Previous Three Years)			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	-0.078^{***} (0.024)	-0.048 (0.032)	-0.094^{***} (0.026)	-0.075^{**} (0.036)
Public Goods	0.021 (0.017)	0.045^{*} (0.026)	0.015 (0.019)	0.034 (0.028)
PUBLIC X TAX		-0.070 (0.048)		-0.046 (0.051)
Linear Combination		-0.073 (0.034)		-0.087 (0.037)
Drop all boundary properties	Yes	Yes	Yes	Yes
F-Stat	897.64	700.59	4446.7	3479.24
Observations	24,443	24,443	25,311	25,311
Adjusted R^2	0.358	0.358	0.545	0.545

"Change in Amount of Tax Paid from Average of Previous Three Years in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then subtracting from that quantity of the sum of all property tax payments for a given property made for each of the previous three fiscal years prior to the start of the project (FY2016, FY2017 and FY2018) divided by 3, adding one (1), and then taking the log() of the result. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 24:
Effects of Tax Collection, Public Goods and their Interaction on Paying Tax Bill in Full

	Binary Indicator Paid All of the Tax Bill			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−0.014* (0.007)	−0.005 (0.009)	−0.009 (0.006)	−0.002 (0.007)
Public Goods	0.004 (0.004)	0.013* (0.007)	0.004 (0.003)	0.008* (0.005)
PUBLIC X TAX		−0.023 (0.014)		−0.016 (0.012)
Linear Combination		-0.014 (0.011)		-0.01 (0.01)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	0.09	0.09	0.05	0.05
F-Stat	173.31	135.48	104.73	82.56
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.175	0.175	0.122	0.122

"Binary Indicator Paid All of the Tax Bill" is a binary indicator calculated by first summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then subtracting the total tax bill. The indicator variable equals 1 if the difference is greater than or equal to zero and otherwise the indicator equals to zero. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 25:
Effects of Tax Collection, Public Goods and their Interaction on Percent of Tax Bill Paid

	Percent of Tax Bill Paid			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−0.027** (0.012)	−0.012 (0.013)	−0.014 (0.009)	−0.001 (0.011)
Public Goods	0.012 (0.008)	0.019 (0.012)	0.007 (0.007)	0.013 (0.009)
PUBLIC X TAX		−0.034 (0.023)		−0.030* (0.017)
Linear Combination		-0.027 (0.017)		-0.017 (0.013)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	0.16	0.16	0.13	0.13
F-Stat	185.67	149.65	199.06	163.61
Observations	27,090	27,090	27,216	27,216
Adjusted R^2	0.124	0.124	0.180	0.180

"Percent of Tax Bill Paid" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then dividing it by the total tax bill. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 26:
Effects of Tax Collection, Public Goods and their Interaction on Number of Tax Installments Made by Property

	Number of Tax Installments Made			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−0.015* (0.008)	−0.008 (0.010)	−0.016 (0.012)	−0.005 (0.015)
Public Goods	0.005 (0.005)	0.013 (0.008)	0.006 (0.008)	0.016 (0.012)
PUBLIC X TAX		−0.017 (0.016)		−0.026 (0.023)
Linear Combination		-0.012 (0.011)		-0.015 (0.016)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	0.13	0.13	0.2	0.2
F-Stat	272.96	218.13	261.7	208.41
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.207	0.208	0.262	0.262

"Number of Tax Installments Made" is a count variable. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 27:
Effects of Tax Collection, Public Goods and their Interaction on Average Amount Per Tax Installment in USD by Property

	Average Amount Per Tax Installment in USD			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−5.27** (2.37)	−7.52** (3.83)	−4.92* (2.69)	−7.25 (4.42)
Public Goods	2.55** (1.23)	1.27 (1.24)	2.17 (1.42)	1.79 (1.44)
PUBLIC X TAX		5.16 (4.08)		5.13 (4.73)
Linear Combination		-1.09 (1.79)		-0.33 (2.02)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	10.71	10.71	13.48	13.48
F-Stat	47.07	37.29	48.19	41.92
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.05	0.05	0.06	0.06

"Average Amount Per Tax Installment in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG), then dividing it by the number of installments made during that period, and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

Table 28:
Effects of Tax Collection, Public Goods and their Interaction on log(1 + Average Amount Per Tax Installment in USD by Property)

	log(1 + Average Amount Per Tax Installment in USD)			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	-0.084^{***} (0.026)	-0.067^{*} (0.036)	-0.080^{***} (0.030)	-0.061 (0.042)
Public Goods	0.019 (0.020)	0.055^{*} (0.028)	0.025 (0.023)	0.058^{*} (0.033)
PUBLIC X TAX		-0.045 (0.053)		-0.050 (0.060)
Linear Combination		-0.057 (0.036)		-0.052 (0.036)
Drop all boundary properties	Yes	Yes	Yes	Yes
F-Stat	273.05	220.49	309.82	257.97
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.263	0.264	0.290	0.290

"log(1 + Average Amount Per Tax Installment in USD)" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG), then dividing it by the number of installments made during that period, then multiplying by an exchange rate of 65HTG:1USD, adding one (1) and then taking the log() of the result. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 29:
Effect of Tax Collection, Public Goods and their Interaction on Binary Indicator for Any Tax Payment by Property

	Binary Indicator for Any Taxes Paid			
	FY19		FY19-FY20	
	(1)	(2)	(3)	(4)
Tax Collection	−0.0133** (0.0060)	−0.0081 (0.0079)	−0.0129* (0.0066)	−0.0069 (0.0090)
Public Goods	0.0018 (0.0044)	0.0109* (0.0065)	0.0035 (0.0050)	0.0115 (0.0076)
PUBLIC X TAX		−0.0131 (0.0118)		−0.0147 (0.0130)
Linear Combination		-0.0103 (0.0081)		-0.01 (0.0083)
Drop all boundary properties	Yes	Yes	Yes	Yes
Mean	0.12	0.12	0.14	0.14
F-Stat	339.83	274.29	389.24	318.48
Observations	27,773	27,773	27,773	27,773
Adjusted R^2	0.2709	0.2710	0.2969	0.2970

"Binary Indicator for Any Taxes Paid" equals 1 if the property paid any amount of taxes in the period indicated and 0 otherwise. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions include the $FRAME_{ib}$ indicator which control for additional randomly assigned variations in implementation of the 'tax collection" treatment beyond the simple delivery and explanation. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

F Treatment Invoices

Summary of Tax Invoice Frames

- Pro-social:
  - Nationalism: (opening lines of the national anthem) *For the country, For the ancestors / Let us march united. Let us march united. / Let there be no traitors in our ranks! Let us be masters of our soil. / Let us march united. Let us march united.* (Appendix F.3)
  - Obligation: (commonly referenced proverbs) *Remember, "Many hands make light work," and, "You can't eat soup with just one finger."* (Appendix F.4)
- Individualistic:
  - Penalties: *Property taxes are due at the end of the month of APRIL! Know that for each month of delayed payment, your total tax bill increases by 5%.* (Appendix F.5)
  - Benefits: *Remember that Property Taxes are the principal source of financing for public goods: garbage removal and street cleaning. Your contribution will make these services possible for you and your community.* (Appendix F.6)

Table 30:
Effects of Framing tax collection on Amount of Taxes Paid By Property Showing Results for all Frames

	Amount of Tax Paid in USD			
	FY19		FY19-20	
	(1)	(2)	(3)	(4)
Public Goods	3.6* (2.0)	0.5 (1.6)	3.3 (2.7)	1.3 (2.6)
Tax Collection	−3.9* (2.3)	−5.1* (2.8)	−4.3 (3.7)	−4.5 (5.0)
Benefit of Compliance	4.0 (3.2)	5.3 (3.3)	8.4* (5.0)	9.4 (6.0)
Penalties	4.0 (2.9)	6.4** (3.0)	6.0 (4.2)	8.7* (4.7)
Nationalism	1.3 (2.7)	6.0** (2.8)	10.3 (8.3)	19.3 (12.6)
Obligation	12.9 (8.2)	9.3* (4.8)	17.2* (8.8)	14.3** (7.1)
Public X TAX		3.4 (4.4)		1.1 (7.4)
Public X BENEFIT		−3.0 (6.9)		−2.5 (11.3)
Public X PENALTIES		−5.9 (5.8)		−6.4 (8.2)
Public X NATIONALISM		−12.8** (5.3)		−24.5* (13.9)
Public X OBLIGATION		10.3 (22.6)		8.2 (23.3)
Mean	13.1	13.1	21.86	21.86
Observations	40,060	40,060	40,060	40,060
Adjusted R^2	0.02	0.02	0.03	0.03

"Amount of Tax Paid in USD" is calculated by summing all property tax payments for a given property made within the specified payment period in Haitian Gourdes (HTG) and then multiplying by an exchange rate of 65HTG:1USD. "FY19" is the first full fiscal year of the experiment running from October 1, 2018 - September 30, 2019. "FY19-20" includes the 18 months of administrative data collected during the period October 1, 2018 - March 30, 2020. "Tax Collection" is an indicator variable equal to 1 if a property is randomly assigned to have a tax agent from the mayor's office hand-deliver an invoice for property taxes and equal to zero otherwise. "Public Goods" is an indicator variable equal to 1 if a property is located in a census bloc randomly assigned to receive garbage removal services and equal to zero otherwise. Analysis is limited to only properties within census blocs that were eligible to receive garbage removal services and thus were part of the bloc-level randomization. All regressions also include individual and bloc controls, strata fixed effects, and standard errors clustered at the bloc-level. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$

F.1 Invoice

LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	

VALEUR LOCATIVE APPLIQUÉE : 131,250.0 **LOYER DÉCLARÉ 75,000.00** **TAUX : 8.00%**

INFORMATION SUR LES CONTRIBUTIONS À PAYER

CONTRIBUTIONS COMMUNALES		CONTRIBUTIONS TRESOR PUBLIC	
Montant principal	10,500.00	Taxe additionnelle (10%)	1,050.00
		DSAV (2/1000)	21.00
			0.00
Total contribution communale	10,500.00	**Total contribution Trésor Public**	1,071.00

CONTRIBUTION TOTALE À PAYER EN GOURDES : 11,571.00

AUTRES EXERCICES A PAYER : 2010-2011, 2011-2012, 2012-2013, 2013-2014, 2014-2015, 2015-2016, 2016-2017, 2017-2018

IMPORTANT : La mairie de Carrefour invite les contribuables de la commune de Carrefour à verser le montant mentionné ci dessus, sans addition des surtaxes de retard, à condition que le paiement se fasse avant le 30 Avril 2019. Passé ce délai,le centre des impôts de Carrefour (CDI) calculera le montant des surtaxes à raison de 5% par mois de retard et exigera leur paiement. Tel 22260726

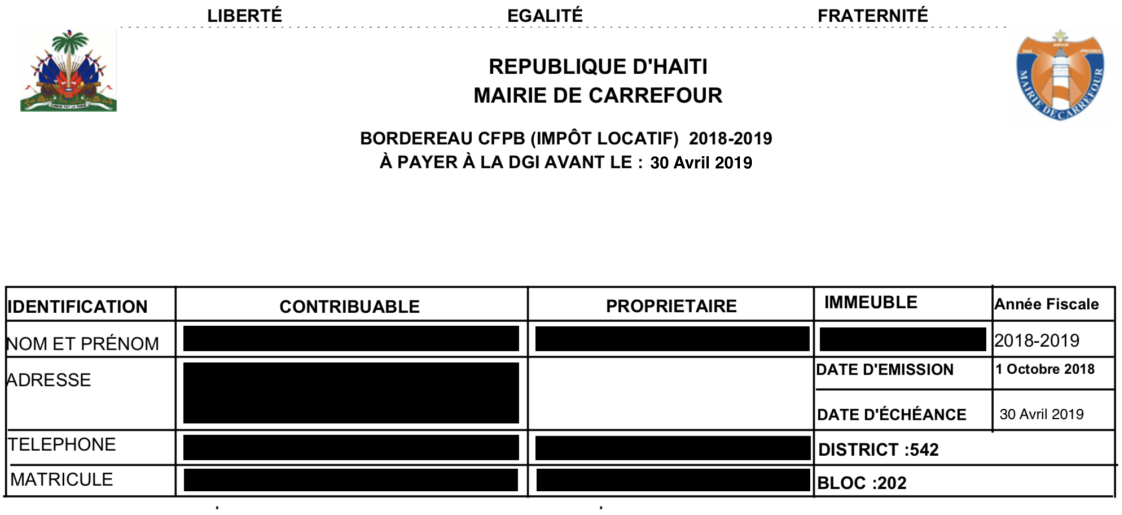
Gagnez du temps et de l'effort! Avec bordereau en main, veuillez-vous présenter directement à la DGI pour faire le paiement. La DGI est située entre Thor 14 et Cote Plage 16 entre 9h00 AM et 4h00 PM du Lundi au Vendredi.

Directeur des affaires Administratives et Financiere

TYPE :**A**

BM-96659C2B42F1 Imprimé par : Admin Imprimé le : 30/3/2019 8:33:44AM

F.2 Invoice (base version with empty heading)



LIBERTÉ EGALITÉ FRATERNITÉ

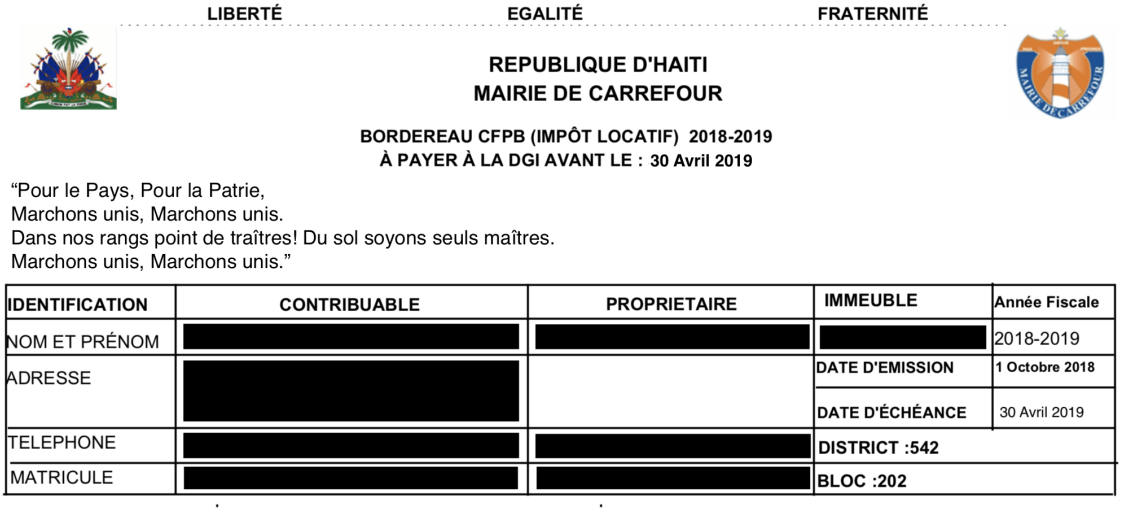
REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	

F.3 Invoice + Pro-social Frame: Nationalism

For the country, For the ancestors / Let us march united. Let us march united. /
Let there be no traitors in our ranks! Let us be masters of our soil. /
Let us march united. Let us march united.



LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

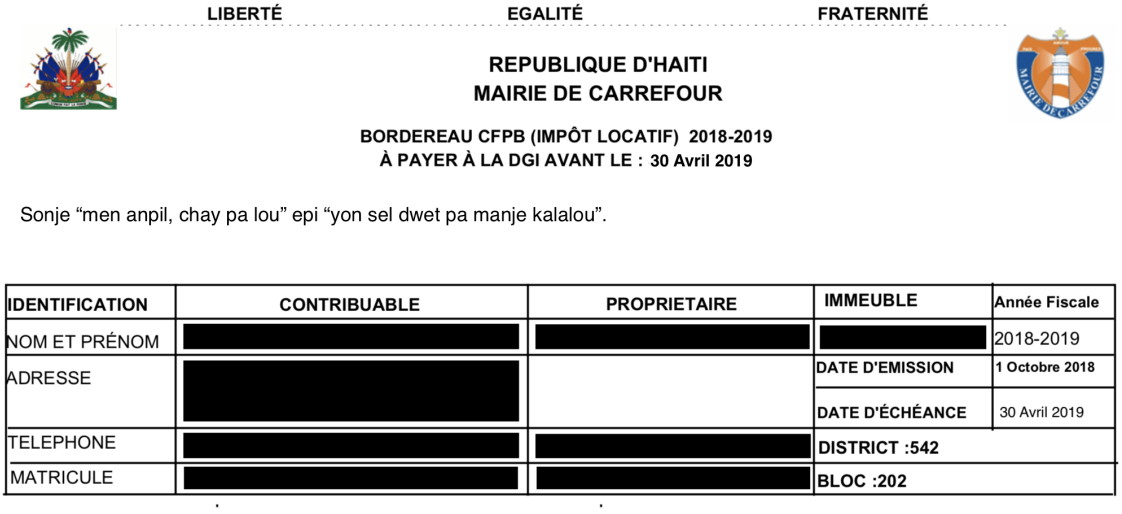
BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

"Pour le Pays, Pour la Patrie,
Marchons unis, Marchons unis.
Dans nos rangs point de traîtres! Du sol soyons seuls maîtres.
Marchons unis, Marchons unis."

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	

F.4 Invoice + Pro-social Frame: Obligation

Remember,"Many hands make light work," and, "You can't eat soup with just one finger."



LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

Sonje "men anpil, chay pa lou" epi "yon sel dwet pa manje kalalou".

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	

F.5 Invoice + Individualistic Frame: Costs

Property taxes are due at the end of the month of APRIL! Know that for each month of delayed payment, your total tax bill increases by 5%.

LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

Les taxes sont dues avant la fin du mois de AVRIL! Sachez que chaque mois de retard de paiement, votre facture fiscale totale augmente de 5%.

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	

F.6 Invoice + Individualistic Frame: Penalties

Remember that Property Taxes are the principal source of financing for public goods: garbage removal and street cleaning. Your contribution will make these services possible for you and your community.

LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

Rappelez vous que l'Impôt locatif est la source principale de financement pour les services publics: l'enlèvement des déchets et le nettoyage des rues. Votre contribution rendra ces services possibles pour vous et votre communauté.

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	

F.7 Government Capacity (Placebo)

LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE
NOM ET PRÉNOM	[redacted]	[redacted]
ADRESSE	[redacted]	
TELEPHONE	[redacted]	[redacted]
MATRICULE	[redacted]	[redacted]

Jours Fériés	2019	2020
Jour de l'Indépendance / Nouvel An	Lundi 1er janvier	Mecredi 1er janvier
Jour des Aïeux	Mercredi 2 janvier	Jeudi 2 janvier
Mardi Gras	Mardi 5 mars	Mardi 25 fevrier
Vendredi Saint	Vendredi 19 avril	Vendredi 10 avril
Pâques	Dimanche 21 avril	Dimanche 12 avril
Fête du travail	Mercredi 1 mai	Vendredi 1 mai
Fête du drapeau	Samdi 18 mai	Lundi 18 mai
Fête Dieu	Vendredi 31 mai	Jeudi 11 juin
Assomption	Jeudi 15 aout	Samedi 15 aout
Mort de Dessalines	Jeudi 17 octobre	Samedi 17 octobre
La Toussaint	Vendredi 1 novembre	Dimanche 1 novembre
Fête des Morts	Samdi 2 novembre	Lundi 2 novembre
Bataille de Vertière	Lundi 18 novembre	Mercredi 18 novembre
Veille de Noel	Mardi 24 décembre	Jeudi 24 décembre
Noel	Mercredi 25 décembre	Vandredi 25 décembre
Veille Nouvel An	Mardi 31 décembre	Jeudi 31 décembre

Directeur des affaires Administratives et Financiere

TYPE :

BM- Imprimé par : Admin Imprimé le : 30/3/2019 8:33:44AM

This letter was designed to be as close to the tax invoice as possible without making any mention of taxes or the tax obligation. Importantly, it is includes all of the same personal information identifying the residents and owners of the property. The content of the message however is simply a calendar of upcoming bank holidays.

F.8 Invoice + Social Exposure

If you pay your property taxes in the next 30 days, the Mayor will place a sign indicating payment on the exterior wall of your house facing the street.

LIBERTÉ EGALITÉ FRATERNITÉ

REPUBLIQUE D'HAITI
MAIRIE DE CARREFOUR

BORDEREAU CFPB (IMPÔT LOCATIF) 2018-2019
À PAYER À LA DGI AVANT LE : 30 Avril 2019

Si vous payez votre impôt locatif dans les 30 prochains jours, la Mairie affichera un panneau indiquant que vous avez payé sur le mur extérieur de votre maison faisant face à la rue.

IDENTIFICATION	CONTRIBUABLE	PROPRIETAIRE	IMMEUBLE	Année Fiscale
NOM ET PRÉNOM	[redacted]	[redacted]	[redacted]	2018-2019
ADRESSE	[redacted]		DATE D'EMISSION	1 Octobre 2018
			DATE D'ÉCHÉANCE	30 Avril 2019
TELEPHONE	[redacted]	[redacted]	DISTRICT :542	
MATRICULE	[redacted]	[redacted]	BLOC :202	