Optimal Assignment of Bureaucrats:

Evidence from Randomly Assigned Tax Collectors in the DRC

Augustin Bergeron, Pedro Bessone, John Kabeya Kabeya, Gabriel Tourek, Jonathan L. Weigel

August 30, 2022

Abstract

The assignment of workers to tasks and teams is a key margin of firm productivity and a potential source of state effectiveness. This paper investigates whether a low-capacity state can increase its tax revenue by optimally assigning its tax collectors. We study the two-stage random assignment of property tax collectors into teams and to neighborhoods in a large Congolese city. The optimal assignment involves positive assortative matching on both dimensions: high (low) ability collectors should be paired together, and high (low) ability teams should be paired with high (low) payment propensity households. Positive assortative matching stems from complementarities in collector-to-collector and collector-to-household match types. We provide evidence that these complementarities reflect in part high-ability collectors exerting greater effort when matched with other high-ability collectors. According to our estimates, implementing the optimal assignment would increase tax compliance by 2.94 percentage points and revenue by 26% relative to the status quo (random) assignment. Alternative policies, such as replacing low-ability collectors with new ones of average ability or increasing collectors' performance wages, are likely incapable of achieving a similar revenue increase.

Affiliations: Bergeron – University of Southern California (augustin.bergeron@usc.edu), Bessone – Uber (pedrobtepedino@gmail.com), Kabeya Kabeya – Direction Générale des Recettes du Kasaï-Central (DGRKAC), Tourek – University of Pittsburgh (gabriel.tourek@pitt.edu), Weigel – UC Berkeley and NBER (jweigel@berkeley.edu). We thank Timothy Armstrong, Oriana Bandiera, Vittorio Bassi, Michael Best, Stéphane Bonhomme, Anne Brockmeyer, Katherine Casey, Arun Chandrasekhar, Isabelle Cohen, Thomas Chaney, Raj Chetty, Lucas de Lima, Esther Duflo, Pascaline Dupas, Marcel Fafchamps, Edward Glaeser, Alan Griffith, Nathan Hendren, Jonas Hjort, Asim Khwaja, Toru Kitagawa, Gabriel Kreindler, Pablo Kurlat, Isabela Manelici, Robert Metcalfe, Joana Naritomi, Nathan Nunn, Paulina Oliva, Benjamin Olken, Imran Rasul, Frank Schillbach, Francesco Trebbi, Jeffrey Weaver, and Noam Yuchtman as well as seminar participants at the NBER Dev SI and elsewhere for invaluable suggestions. For outstanding research assistance, we thank Mert Akan, Manon Delvaux, Samih Ferrah, Arnaud Fournier, Arthur Laroche, Alix Leroy, Florence Oberholzer, Stephen Mathew and David Mast. For excellent program management we thank Elie Kabue Ngindu and the rest of the Odeka Team. We are grateful for the collaboration with the Provincial Government of Kasaï-Central and the Direction Générale des Recettes du Kasaï-Central (DGRKAC). We gratefully acknowledge funding from the EGAP Metaketa II Taxation Initiative and the J-PAL Governance Initiative. This study has been approved by the Harvard Institutional Review Board: Protocol IRB17-0724. The Pre-Analysis Plan was submitted to the AEA RCT Registry on May 25, 2020: AEARCTR-0001316.

1 Introduction

The assignment of workers to tasks and teams is an important margin through which private firms can raise productivity.[1] Less is known, however, about the assignment margin in the public sector, even though ex-ante it may be an attractive tool to raise performance. Indeed, the public sector is often beset by inefficiencies, and many standard tools to boost worker performance, such as promotion incentives, are typically unavailable to governments due to seniority-based civil service regulations.[2] Moreover, there is growing recognition that public-sector workers explain a large share of the variation in state performance across sectors and regions (Finan et al., 2017; Best et al., 2019; Fenizia, 2022). Yet, we have little evidence on whether improving the assignment of public sector employees to postings or teams can enhance state effectiveness.

This paper examines the assignment of frontline public-sector workers as a source of state capacity. We study tax collectors in the Democratic Republic of the Congo (DRC), a fragile state seeking to build a reliable tax revenue base from the ground up. As in many developing countries, field-based teams of tax collectors solicit payment of the property tax directly from households. During the six-month 2018 property tax campaign, the Provincial Government of Kasaï Central randomized tax collectors to teammates and neighborhoods to minimize opportunities for collusion. Our design exploits the two-stage random assignment of (*i*) 34 tax collectors into new two-person teams each month, and (*ii*) collector teams to 180 neighborhoods (19,600 properties) in the city of Kananga. Collector teams first went door to door registering properties and then returned to collect the property tax. During the property tax campaign, the median collector worked with 6 different teammates and in 12 different neighborhoods (covering over 1,500 properties).

We use this two-stage randomization to estimate the optimal assignment — of collectors to teammates and teams to households — and its impact on tax compliance, i.e., the probability that households pay taxes.[3] First, we partition households into high and low types according to their tax payment propensity. To measure households' payment propensity, we rely on estimates of each property owner's ability to pay the property tax provided by the neighborhood chief before tax collection in 78 randomly selected neighborhoods (our *analy-*

[1] See, e.g., Becker (1973); Crawford and Knoer (1981) on the role of assignment theoretically and Graham (2011); Graham et al. (2014); Bonhomme (2021) on estimation for different types of assignment problems. See, e.g., Rotemberg (1994); Ichino and Maggi (2000); Mas and Moretti (2009); Bandiera et al. (2010) on peer effects and social incentives in the workplace.

[2] Bertrand et al. (2020) show that rigid promotion rules constrain the performance of public-sector workers.

[3] The approach we adopt adapts and extends Bhattacharya (2009), and Graham et al. (2020a).

sis sample).[4] Chiefs' estimates are highly correlated with subsequent tax compliance during the campaign, providing a convenient pre-treatment measure of household type. Similarly, we partition tax collectors into two types.[5] Because we lack a pre-treatment measure of collector ability, we use a sample-splitting approach, estimating collector type in the randomly selected sample of 102 neighborhoods for which we don't have information about household payment propensity (our *holdout sample*). Specifically, using a fixed effects model, we estimate the average compliance achieved by each collector in the neighborhoods to which they were randomly assigned. We then split at the median to define high and low types.

Having defined collector and household types, we use the analysis sample to estimate the average tax compliance function — i.e., the expected tax compliance conditional on collector and household types — non-parametrically (Bhattacharya, 2009; Graham et al., 2020a). We then use our estimates to find the counterfactual optimal assignment function: the assignment of collectors to teammates and households that would maximize tax compliance subject to status quo constraints in the marginal distributions of collector type and collector-household type. Finally, we estimate the effect of implementing the (counterfactual) optimal assignment on tax compliance and revenue relative to the status quo random assignment.

It is not ex-ante obvious what assignment would be optimal.[6] Negative assortative matching of collectors into teams could be justified if the essential tasks can be done by a single (skilled) worker, while positive assortative matching could be optimal in the presence of complementarities in effort or skill. Similarly, if high-type households pay taxes with minimal effort on the part of collectors, it could be optimal to assign them to low-type collectors; but the opposite would be true if it takes conscientiousness and persuasiveness to elicit payment even from high-type households. What assignment function maximizes tax compliance is thus an empirical question.[7]

We find that the optimal assignment involves positive assortative matching on both dimensions. To maximize tax compliance while holding tax collection staff constant, the government should (*i*) form teams of exclusively high- or low-type collectors (i.e., homogeneous

[4] These chiefs are locally embedded leaders with a high degree of local information about each neighborhood's residents. After property registration but before collection, state collectors consulted with the city chief in the neighborhood to ask about the ability to pay of each resident (Balan et al., 2022).

[5] In the main analysis, we use two collector types because it appears to be the optimal number of types according to unsupervised machine learning methods (cf. Section A3. The results are also robust to allowing for more collector types (cf. Section 8.2).

[6] Past empirical work on optimal matching reaches divergent conclusions depending on the context and production function of interest (cf. Section 7).

[7] Importantly, by estimating the tax compliance function non-parametrically, our empirical approach allows us to detect complementarity (supermodularity), substitutability (submodularity), or neither.

teams), and (*ii*) assign high-type teams to households with high payment propensity and low-type teams to households with low payment propensity. Positive assortative matching stems from complementarities in collector-to-collector type and collector-to-household type in the average tax compliance function. We provide evidence that these complementarities reflect high-type collectors exerting greater effort when matched with other high types, collecting taxes on more distinct days and for more total hours. They also focus their higher enforcement effort towards high-type households, in neighborhoods where cash-on-hand constraints are less likely to bind, and at times of day when property owners are likely to be cash "rich." High-type teams thus appear to raise more revenue by working longer hours, which increases the probability that they visit property owners at times they have the cash on hand to pay.

We estimate that implementing the optimal assignment would increase tax compliance by 2.941 percentage points relative to the status quo random assignment. This amounts to a 37% increase in tax compliance relative to the status quo average of 8%. The optimal assignment would also lead to a 54 Congolese Francs (CF) increase in tax revenue per owner, a 26% increase from the status quo assignment average of 206 CF. Each dimension of the optimal assignment — collector-to-collector and collector-to-household — contributes roughly equally to the total effect. Specifically, only optimizing the assignment of collectors to teammates would increase compliance by 16%, while only optimizing the assignment of collectors to households would increase compliance by 10%. The increase in tax compliance under the optimal assignment would be progressivity-enhancing, largely falling on wealthier households with more valuable properties.

We consider a range of robustness checks, including implementing an alternative non-linear methodology, varying the number of collector types, using alternative definitions of collector and household type, considering alternative government maximands, optimizing at the neighborhood level (rather than at the household level), and providing estimates robust to overfitting and the "winner's curse." None of these exercises qualitatively change the main results. An important concern is potential endogenous responses to the implementation of the optimal assignment, motivated by past evidence that counterfactual "optimal" assignments can sometimes backfire when implemented in the real world (Carrell et al., 2013).[8] The main

[8] As discussed in Section 8, the specific issue encountered in Carrell et al. (2013) — endogenous formation of subgroups of low- and high-ability students within mixed cohorts — is less concerning in our setting because (*i*) teams are of size two, which prevents endogenous social group formation within teams, (*ii*) in the RCT, we directly observe tax compliance for all possible combinations of types — including *L*-*L* and *H*-*H* teams, which characterize the optimal assignment — while Carrell et al. (2013) relied on predictions outside of their RCT's support, and (*iii*) we can use the monthly reassignment of collectors to test specific implementation issues of concern (e.g., effort or learning), while Carrell et al. (2013) only observed individuals assigned to a single cohort.

issues in our setting are that changing collectors' assignments could directly impact their effort levels or opportunities for learning by match type over the course of the campaign. According to evidence from the repeated re-assignment of collectors each month of the campaign, low-type collectors would be unlikely to differentially exert less effort or lose learning opportunities under the optimal assignment.

We also investigate potential unintended consequences of implementing the optimal assignment policy on other margins, such as bribery, payment of other taxes, and citizens' views of the tax authority. Although there is a marginally significant increase in total bribes, this reflects the fact that high-type collector teams make more visits to households and thus have more opportunities to collect bribes. Indeed, we find no increase in bribes *per visit* — the relevant policy parameter — under the optimal policy. According to our estimates, the optimal policy would also not undermine citizens' compliance with other taxes, their view of the government, or their tax morale.

Finally, to benchmark our estimated effects, we consider two counterfactual collector selection policies: reallocating households assigned to low-type collectors to high-type collectors (*reallocation policies*) or to newly hired collectors (*hiring policies*). To achieve the same increase in tax compliance as the optimal assignment, the government would have to reallocate 63% of the households assigned to low-type collectors to high-type collectors. Alternatively, reallocating households to newly hired collectors of average ability would never achieve compliance gains comparable to those from the optimal assignment, even if all low-type collectors' households were reallocated.

As a further benchmark, we compare the impact of the counterfactual optimal assignment to the effect of performance-based financial incentives for tax collectors. Leveraging random variation in collectors' piece-rate wages during the 2018 tax campaign, we find that the government would have to increase collector compensation by 69% to increase tax compliance as much as the optimal assignment. However, such a policy would actually reduce tax revenue net of wages by 6%, due to the mechanical increase in the wage bill. The cost-ineffectiveness of this policy underscores a crucial advantage of the optimal assignment policy: it would increase state effectiveness while holding constant existing financial and human resources.

We contribute to three strands of literature. First, we provide some of the first estimates of the importance of public-sector worker assignment in shaping state effectiveness. While past work examines the importance of selection (Dal Bó et al., 2013; Callen et al., 2018; Hanna and Wang, 2017; Xu, 2018; Ashraf et al., 2020; Dahis et al., 2021), incentives (Ashraf et al., 2014; Khan et al., 2016, 2019; Rasul and Rogger, 2018; Bertrand et al., 2020; Bandiera et al., 2021), monitoring (Duflo et al., 2012; Dal Bó et al., 2020), and management practices (Rasul and

Rogger, 2018; Bandiera et al., 2021), less attention has been paid to the assignment margin.[9] Two closely related papers are Best et al. (2019) and Fenizia (2022), which exploit the rotation of bureaucrats across sites to study the role of bureaucrat quality in explaining public sector performance.[10] We build on these studies by exploring the optimal assignment of public sector workers to teams and postings, leveraging the random assignment of tax collectors and studying more objective performance measures (tax compliance and revenue) than are typically available for bureaucrats.[11] Finally, we advance this literature by complementing our analysis of administrative data with rich survey data to explore the mechanisms explaining the optimal assignment of collectors and to consider other policy-relevant response margins, such as bribery, payments of other taxes, and citizens' views of the tax authority.

Second, we contribute to the literature on optimal tax administration in developing countries. Given that low-income countries with weak states are characterized by imperfect tax enforcement (Besley and Persson, 2013; Pomeranz, 2015), tax administration is a crucial dimension of their tax policy (Keen and Slemrod, 2017). Past work in developing countries focuses on performance incentives for tax collectors (Khan et al., 2016, 2019), the type of agent hired as tax collectors (Balan et al., 2022), and the use of large taxpayer offices to increase the staff-to-taxpayer ratio (Basri et al., 2021).[12] We contribute to this literature by examining whether governments can, holding other inputs constant, raise revenue simply by improving the assignment of collectors to teammates and of teams to taxpayers. Importantly, this optimal assignment policy aims at improving tax administration using available tax collectors — i.e., without incurring additional costs — which makes it particularly attractive in weak state settings.

Third, we contribute to the optimal matching literature. Recent applied work has stud-

[9] Khan et al. (2019) study the incentive effects of performance-based postings. By contrast we focus on the direct effects of assignment on bureaucrat performance, which Khan et al. (2019) are unable to assess because they cannot disentangle assignment effects (top collectors are assigned to larger tax jurisdictions) from mechanical persistence (once a property is added to the tax roll, the owner pays taxes in subsequent years).

[10] Best et al. (2019) analyze the importance of bureaucrat quality in explaining public procurement prices in Russia. Fenizia (2022) studies the productivity impacts of managers in the public sector in Italy.

[11] Fenizia (2022) includes a similar optimal assignment analysis with three key differences: (*i*) the focus is on the assignment of managers rather than front-line bureaucrats; (*ii*) it studies the uni-dimensional assignment of managers to offices, while we study the bi-dimensional assignment of collectors to teammates and to households; and (*iii*) the optimal assignment analysis assumes ex-ante that the production function is supermodular in office and manager fixed effects, thereby potentially magnifying the extent of positive assortative matching. By contrast, we estimate the production function non-parametrically, which allows us to potentially identify both positive and negative assortative matching.

[12] Beyond tax administration, the literature on public finance in developing countries has primarily focused on tax enforcement (Pomeranz, 2015; Carrillo et al., 2017; Naritomi, 2019), tax instruments (Best et al., 2015), and tax rates (Basri et al., 2021; Bergeron et al., 2021; Brockmeyer et al., 2021).

ied the impact of optimally matching teachers to students (Graham et al., 2020a; Aucejo et al., 2019; Bhattacharya, 2009), students to classmates (Carrell et al., 2013), and financial advisers to clients (Bessone, 2020). While these papers consider uni-dimensional assignment problems, we study the bi-dimensional problem of assigning collectors to teammates and households. In our context, only considering one of the two dimensions would reduce the impact of the optimal assignment by more than half. Moreover, this is (to our knowledge) the first optimal matching paper to exploit the random assignment of workers to postings *and* teammates.[13] Perhaps the closest paper in this vein is Marx et al. (2021), which explores the impact of ethnic diversity among NGO workers canvassing voters in Kenya and similarly uses multiple dimensions of random assignment variation.[14] Finally, we make a small methodological contribution by applying the median-unbiased estimators developed by Andrews et al. (2021) to address possible "winner's curse" upward bias that can arise in optimization problems like those considered in this literature.

2 Setting

The DRC, one of the poorest countries in Africa, is a paradigmatic fragile state with one of the lowest tax-GDP ratios in the world.[15] Kananga, the capital of the province of Kasaï Central, has a population of nearly 1 million and an average monthly household income of \$106 (PPP\$168). The tax revenue of the Provincial Government of Kasaï Central, roughly \$0.30 per person per year in 2015, comes primarily from business licenses and fees, trade and transport taxes, and property taxes. In keeping with international best practices for local revenue mobilization, the provincial government has turned to the property tax to increase tax revenue (Franzsen and McCluskey, 2017), conducting a series of citywide door-to-door collection campaigns since 2016 (Weigel, 2020; Balan et al., 2022).

Although the provincial government is charged with maintaining local roads and infrastructure, public transportation, and trash collection — all of which should ostensibly be paid for with property tax revenues — such services are woefully under-provided. Only the city's

[13] Carrell et al. (2009) study peer effects using the random assignment of students to peer groups. Graham et al. (2020a) use the random assignment of teachers to classrooms to study teacher-to-classroom assignment.

[14] The key differences with our paper include (*i*) their focus is ethnic diversity, while ours is optimal assignment and its impacts (though we also examine matching by ethnicity or other sources of horizontal differentiation in Section A6), (*ii*) they examine teams of NGO workers, while our interest is in assignment of public-sector workers, and (*iii*) the larger number of teams we observe — 132 v. 30 — allows us to explore a broader range of match types, potential mechanisms, and dynamic effects of assignment (Section 8.3). Despite these differences, the papers provide complementary evidence that effort — and specifically the number and duration of visits — is a key mechanism explaining variation in field-based teams' effectiveness.

[15] The DRC's tax-GDP ratio ranks 188 out of 200 countries, including oil-rich countries (OECD, 2020).

main arteries are paved, and they are in severe disrepair or threatened by erosion. In sum, Kananga closely resembles the kind of low-equilibrium trap noted by Besley and Persson (2009), with low state capacity, low tax compliance, and low service provision.

2.1 The 2018 Property Tax Campaign

The experiment we study was embedded in the 2018 property tax campaign, implemented in Kananga by the Provincial Government of Kasaï Central. Before describing the experimental design, we outline key details and procedures of the tax campaign.

Tax Collectors. State tax collectors were contractors hired specifically by the provincial ministry to work on the 2018 property tax campaign.[16] They were drawn from a pool of aspiring bureaucrats who frequently perform contract work for different arms of the provincial government.[17] They did not receive a regular salary outside of the piece-rate compensation for working as a tax collector (noted below). Collectors were on average 30 years old, 94% male, and 70% of them had some university education. Their average household monthly income prior to being hired to work on the tax collection campaign was $110. During the property tax campaign none had full-time jobs in addition to their tax collector work, but 67% of them had some other informal income-generating activities (e.g., leasing out a motorbike to a taxi driver or various forms of petty commerce).

Tax collectors worked in teams of two (which we also refer to as collector pairs), a practice adopted by the provincial tax ministry for all types of tax collection for two reasons. First, it provides a measure of security given that collectors handle state money in the field. Second, it may reduce collusion between collectors and households because hiding illegal transactions is potentially harder when another tax collector is present. Collection by teams could then also inspire confidence among households that their taxes would reach the state rather than collectors' pockets. In many developing countries, working in teams is common among frontline agents in the public and private sectors.[18]

Campaign Stages. In each neighborhood, collectors had one month to complete two tasks: property registration and tax collection (as summarized in Table A1). First, collector teams mapped the neighborhood and constructed a property register. In the absence of an up-to-date property valuation roll, this property register identified those liable for the property tax in each neighborhood. When registering properties, collectors assigned a unique tax ID to each property and issued official tax notices showing the tax liability and other infor-

[16]In some neighborhoods, which are excluded from this analysis, tax collection was conducted by the neighborhood chiefs, as described in Balan et al. (2022).

[17]Such contract work typically consists of public administration tasks like tax collection or health campaigns.

[18]See, e.g., Burgess et al. (2010); Khan et al. (2016); Ashraf and Bandiera (2018).

mation about the tax. Collectors assessed each property's tax liability based on the principal house's construction, as described below, or whether it was exempt.[19] Independent surveyors equipped with GPS devices accompanied collectors during property registration, recording properties' locations, tax IDs, and other household characteristics. Collectors were also instructed to demand payment of the tax during the registration step, or make appointments for future visits.

Second, after completing the property register, the collector team spent the rest of the month making further in-person tax collection visits. They had printed copies of the register containing each property owner's name, tax ID, rate, and exemption status. When they visited a property, they were instructed to record the date of the visit in chalk on the wall of the house. Collectors used handheld receipt printers to issue receipts to taxpayers. The transaction-level receipt data was recorded in the device's memory and downloaded weekly to the government's tax database when they deposited tax revenue. Collectors were required to account for discrepancies with the receipt data (rare in practice). The in-person nature of tax collection left much to the discretion of the collectors: which properties to revisit, how many times, at what time of the day, what persuasion tactics to use to convince property owners to pay, etc. This high degree of discretion for frontline state agents in this and many developing countries motivates our investigation into collector assignment as a source of state effectiveness.[20]

Collector Compensation. Collectors earned piece-rate wages with two components. First, they received 30 CF per property registered. Second, they earned compensation proportional to the amount of tax they individually submitted to the state account.[21] Individual compensation diminished incentives for free-riding.[22] Collectors were also reimbursed for one round trip per day from the tax ministry to their assigned neighborhoods. Beyond monetary compensation, collectors also had career incentives to work hard: after the previous property tax campaign, the tax ministry hired the best performers for full-time positions.

Timing. The campaign began in May 2018 and ran through December. Collector teams worked in two neighborhoods simultaneously, alternating between them during the assigned

[19]Exempt properties constitute 14.27% of total properties in Kananga. They include: (1) properties owned by the state; (2) school, churches, and scientific/philanthropic institutions; (3) properties owned by widows, the disabled, or individuals 55 years or older; and (4) properties with houses under construction.

[20]Field-based visits from tax collectors/inspectors are a cornerstone in tax authorities' enforcement arsenal in many developing countries (e.g., Khan et al., 2016; Cogneau et al., 2020; Krause, 2020; Okunogbe, 2021).

[21]Performance pay for tax staff is used in Pakistan, Brazil, and elsewhere (Khan et al., 2016). In Kanaga, the compensation scheme varied randomly on the property level between (*i*) 30% of the amount of tax collected, and (*ii*) 750 CF per owner who paid the tax. We explore this variation in Section 8.

[22]Collectors rarely worked alone (unless their teammate was sick). They were instructed to alternate which collector on the team took the payment and followed this norm closely according to the tax data.

month. They completed the property register in the first few days of the month and then conducted tax collection visits for the remainder. The average neighborhood comprised 124 properties, and the collectors had ample time to return to properties in both neighborhoods multiple times within the month-long period.

Tax Rates. The property tax in Kananga is a simplified instrument: a fixed fee due once per year that is determined by the value band of a property. Houses made of non-durable materials (e.g., mudbricks) constitute the low-value band with an annual tax liability of 3,000 CF ($2). In contrast, houses made of durable materials (bricks or concrete) constitute the high-value band with a tax liability of 13,200 CF ($9). Although these rates may seem low, they correspond to an average tax rate of roughly 0.32% of estimated property value,[23] not far from the property tax rates in certain U.S. states, which range from 0.27% to 2.35%. Across Kananga, 89% of the properties are classified in the low-value band and 11% are classified in the high-value band.[24] Simplified property tax schemes like the one used in Kananga are common in developing countries, including India, Tanzania, Sierra Leone, Liberia, Malawi, and elsewhere (Franzsen and McCluskey, 2017).

Enforcement. Properties that do not pay the property tax by the year's end in theory owe 250% of the original liability and face a possible court summons. Although sanctions are rarely enforced among the residential property owners who comprise our sample, the majority of citizens at baseline believed that the government would be "likely" or "very likely" to sanction tax delinquents. The ability to shape citizens' perceptions regarding the probability of enforcement is thus a potential mechanism through which some collector teams may prove more effective at collecting taxes than others, which we consider in Section 7.2.

3 Design

3.1 Tax Collector Assignment

To study the optimal assignment of tax collectors, we leverage the random assignment of collectors to teammates and to neighborhoods by the provincial government during the 2018 property tax collection campaign. Every month of the six-month tax campaign, teams of two tax collectors were randomly formed.[25] These teams were then randomly assigned to two neighborhoods, where they would collect taxes for the month. The median assignment load of collectors included 6 different teammates in 12 different neighborhoods spanning 1,200

[23]We estimate property value using machine learning as described in Bergeron et al. (2022).

[24]A separate group of high-value properties, classified as villas, were taxed according to a different schedule and by different collectors and thus are excluded from our analysis.

[25]Each month is an independent draw; i.e., across months the random sampling is with replacement.

properties.

Our analysis focuses on the 180 neighborhoods of Kananga in which a set of 34 state tax collectors were randomly assigned to teams and then to neighborhoods.[26] These 180 neighborhoods span two randomly selected sub-samples where the same state tax collectors worked. In 78 neighborhoods (6,904 properties), which we call the *analysis sample*, the resident city chief went through the property register with collectors and estimated each household's economic ability to pay the property tax before tax collection.[27] We use the chiefs' predictions to define household type (cf. Section 6.1). In the 102 remaining neighborhoods (11,732 properties), which we call the *holdout sample*, we estimate collector types using a fixed effects model (cf. Section 6.2). After defining household and collector types, we then estimate the average tax compliance function and the optimal assignment in the *analysis sample* (cf. Section 6.3). This sample-splitting approach allows us to estimate collector types and the average tax compliance function in different samples to minimize overfitting (i.e., estimating collector type and the average tax compliance function partly based on noise).

The provincial tax ministry has relied on the randomized assignment of tax collectors to teammates and neighborhoods since it began large-scale property tax collection in 2016. The government's logic behind random assignment is twofold. First, as elsewhere, the provincial tax authorities seek to evaluate the impact of policies seeking to raise revenue and have embraced randomization to this end.[28] Second, the tax authorities seek to prevent the development of collusive bribe-paying arrangements between collectors and property owners that could arise if the same collector teams worked in the same neighborhoods each year.[29] By randomly reassigning collectors to teammates monthly and teams to neighborhoods, the government sought to minimize such collusion.

Many tax authorities deliberately reshuffle collectors in a similar fashion to prevent collu-

[26]The tax campaign was active in the 364 neighborhoods of Kananga. We exclude 184 neighborhoods from the analysis: (*i*) the 8 neighborhoods where the logistics pilot took place, (*ii*) 111 neighborhoods where city chiefs collected taxes ("Local" neighborhoods in Balan et al. (2022)), (*iii*) 50 neighborhoods where city chiefs and a different group of state agents teamed up to collect taxes ("Central X Local" neighborhoods in Balan et al. (2022)), (*iv*) 5 neighborhoods with no door-to-door collection (the pure control in Balan et al. (2022)). We exclude these neighborhoods from our analysis because tax collectors were not randomly assigned to neighborhoods or to teammates (*i* - *iii*) or no citizens paid taxes (*iv*). Additionally, we exclude 10 neighborhoods where one of the tax collectors never worked with another teammate, preventing us from estimating these collectors' fixed-effects in Section 5.

[27]The neighborhoods assigned to this treatment arm are called "Central + Local Information" in Balan et al. (2022) where the treatment arm — aimed at comparing city chiefs as tax collectors to state collectors provided with local information — is described in further detail.

[28]For example, the tax authority compared state agents to city chiefs as property tax collectors during the 2018 property tax campaign (Balan et al., 2022).

[29]Khan et al. (2016) document that this form of collusion exists in property tax collection in Pakistan.

sion. For instance, the random assignment of tax collectors to postings resembles the policy of "removes" that was used in 18th-century England (Brewer, 1990) as well as settings like India (Bertrand et al., 2020), China (Chu et al., 2020), Haiti (Krause, 2020), Senegal (Cogneau et al., 2020), and Malawi (Martin et al., 2021) today. Moreover, random assignment has the advantage of being clearly defined, especially compared to opaque assignment mechanisms observed in some contexts.[30] When we compare the optimal assignment and the status quo assignment, the impacts we estimate are thus well-defined quantities that policymakers from other contexts can easily interpret.

3.2 Balance

Table A2 summarizes a series of balance checks. Panel A considers property characteristics, drawing on geographic data, midline survey data on house quality, and estimated property values from Bergeron et al. (2022). Panel B considers property owner characteristics that were collected at midline but are unlikely to have been affected by the assignment of tax collectors. Panel C considers additional owner characteristics collected at baseline, including attitudes about the government and tax ministry. Panel D considers neighborhood characteristics.

Overall, 2 of the 52 differences reported in Panels A–D of Table A2 are significant at the 5% level, and 6 are significant at the 10% level based on t-tests that do not adjust for multiple comparisons.[31] This is in line with what one would expect under random assignment. Table A2 also reports tests of the omnibus null hypothesis that the treatment effects are all zero using parametric F-tests for bilateral comparisons. We fail to reject the omnibus null hypothesis for property, property owner, and neighborhood characteristics. The results are reassuring that the assignment of collector pairs was orthogonal to household characteristics.

4 Data

We use administrative data from property registration and tax collection as well as three household surveys and one survey with tax collectors (Table A1).

4.1 Administrative Data

We have data from property registration on the set of potential taxpayers in each neighborhood. Registration data, covering 19,600 properties in the neighborhoods of interest, include

[30] For instance, Khan et al. (2019) describe the process of assigning tax inspectors to regions of Pakistan as opaque and political (until the government implemented an incentive-based posting mechanism).

[31] Roof quality and having electricity are significant at the 5% level. Distance to education institutions, having a relative who works for the government, ethnic majority status, having electricity, trust in the national government, and a neighborhood-level conflict indicator are significant at the 10% level.

tax ID numbers, geographic coordinates, property owner names, property classifications (cf. Section 2.1), exemption status, and tax rates.[32] The handheld receipt printers used by tax collectors during both stages of the campaign stored details of each transaction in their memory. These data were integrated directly into the government's tax database. The printers recorded the collector's name, a time stamp, neighborhood number, tax ID, property value band, tax rate, and amount paid. By matching payment records to registration data using tax IDs, we observe property tax compliance and revenues — our main outcomes — for all registered properties included in this study.

4.2 Surveys

Household Surveys. Enumerators working for the research team administered baseline surveys to 1,404 households from July to December in 2017.[33] To obtain a representative sample, enumerators visited every X^{th} house, where X was determined by the estimated number of houses in the neighborhood to yield 12 surveys per neighborhood. We primarily use this survey to examine balance of collector assignments.

Enumerators then administered a midline survey at every compound in Kananga two to four weeks after tax collection had finished in a neighborhood. The midline survey measured characteristics of the property and property owner that we use also to examine balance of the collectors' assignment. It also measured secondary outcomes, such as the number of visits from collectors, bribe payments, contributions to other taxes (formal and informal), and respondents' self-reported tax morale and enforcement beliefs. Enumerators attempted to conduct this survey with the property owner for 16,346 properties. For 4,898 of these properties, enumerators conducted the survey with a family member — when the owner was unavailable — or simply recorded property characteristics — such as the quality of the walls, roof, and fence — in the absence of an available respondent.[34]

Collector Surveys. Enumerators administered a baseline and endline survey with collectors before and after the tax campaign. Both surveys covered demographics, trust in the

[32] There are 45,162 registered properties in Kananga. However, we exclude 184 neighborhoods where state tax collectors do not work or are not randomly assigned (cf. Section 3). We also exclude exempt properties. The number of registered properties (19,600) is higher than the total number of properties in the analysis and holdout samples (18,636) mentioned in Section 4.1 due to missing estimates of household's economic ability to pay the property tax for 964 (12%) properties in the analysis sample.

[33] The baseline survey was conducted with 4,343 respondents. After excluding neighborhoods where state tax collectors did not work or were not randomly assigned and exempt properties, we have 1,404 respondents.

[34] The midline survey was conducted with 36,130 respondents. After excluding neighborhoods where state tax collectors did not work or were not randomly assigned and exempt properties, we have 16,346 midline surveys, 11,448 of which were conducted with the owner. Attrition between registration and the midline survey (16.6%) is balanced across collectors (Table A2).

government, perceived performance of the government, views of taxation, and preferences for redistribution. Enumerators surveyed the 34 collectors considered in our analysis.

5 Conceptual Framework

5.1 Household and Collector Types

We consider an economy with N_h households and N_c tax collectors. Each household h is characterized by an observable type $v_h \in V$ and each collector c by an observable type $a_c \in A$, where A and V are finite ordered sets. In the context of tax collection, a household's type measures its likelihood of paying the property tax and a collector's type aims to capture her ability to collect taxes. Each household is assigned to a pair of collectors. Our main specification assumes that households are either low-type (l) or high-type (h) and tax collectors are either low-type (L) or high-type (H), i.e., $v = \{l, h\}$ and $a = \{L, H\}$.[35] A match is a triplet $m = (c_1, c_2, h)$, indicating that tax collectors c_1 and c_2 are assigned to collect taxes from household h. The type of match m is a triplet (a_1, a_2, v_h) that indicates the type of the two collectors and of the household. Since the collectors perform an identical task, their order is arbitrary.

5.2 Optimal Assignment

The government's problem involves picking an assignment function f, which is a probability mass function indicating the distribution of each match type (a_1, a_2, v). The choice of the assignment function f depends on the government's objective and constraints.

Government Objective. We assume that the government chooses the assignment function f that maximizes expected tax compliance, which is given by:

$$\sum_{v \in V} \sum_{a_1, a_2 \in A^2} f(a_1, a_2, v) Y(a_1, a_2, v) \tag{1}$$

$Y(a_1, a_2, v)$ is the average tax compliance function, i.e., the average tax compliance among v-type households assigned to $a_1 - a_2$ pairs of tax collectors:

$$Y(a_1, a_2, v) = \mathbb{E}\left[y_h(c_1, c_2) | a_{c1} = a_1, a_{c2} = a_2, v_h = v)\right] \tag{2}$$

where $y_h(c_1, c_2)$ is the potential tax compliance of household h when assigned to the collector

[35] Unsupervised machine learning methods find that two collector clusters is optimal in this context (cf. Section A3), though we also relax this assumption in Section 8.2.

pair (c_1, c_2).

Government Constraints. We assume that the government faces two constraints when choosing f: a non-overlapping assignment constraint and a workload constraint.

The non-overlapping assignment constraint requires that the number of v-type households to which collector teams are assigned under f equals the total number of v-type households. In other words, this constraint requires the government to assign one and only one collector team to each household. The non-overlapping assignment constraint can be written as:

$$\sum_{a_1, a_2 \in A^2} N_h f(a_1, a_2, v) = N_v \tag{3}$$

with N_h the total number of households and N_v the total number of v-type households.

The workload constraint requires that under the assignment function f, each type of collector is assigned to the same number of households as under the status quo assignment function f^{SQ}. In other words, this constraint requires that the workload of each collector type be kept the same as in the status quo assignment. The workload constraint can thus be written as:

$$\sum_{v \in V} N_f^{asgmt}(a, v) = N_{f^{SQ}}^{asgmt}(a) \tag{4}$$

where $N_f^{asgmt}(a, v)$ is the number of v-type households assigned to a-type collectors under assignment function f, and $N_{f^{SQ}}^{asgmt}(a)$ is the total number of households assigned to a-type collectors under the status quo assignment function f^{SQ}.[36]

Optimal Assignment Problem. The optimal assignment problem thus involves finding the assignment function f^* that maximizes expected tax compliance while keeping the marginal distributions in collector and household type the same as under the status quo assignment. Using the notation above, the optimal assignment problem can be defined as:

[36] $N_f^{asgmt}(a, v) = 2 \times f(a, a, v) \times N_h + \sum_{a' \neq a} (f(a, a', v) + f(a', a, v)) \times N_h$ and $N_f^{asgmt}(a) = \sum_v N_f^{asgmt}(a, v)$. For (a, a, v) matches, two a-type collectors are assigned to a v-type household. The number of households assigned to an a-type collector is thus $2 \times f(a, a, v) \times N_h$. For (a, a', v) or (a', a, v) matches, one a-type collector is assigned to a v-type household. The number of households assigned to an a-type collector is thus $\sum_{a' \neq a} (f(a, a', v) + f(a', a, v)) \times N_h$.

Problem 1. *Optimal Assignment*

$$f^* \equiv \arg\max_f \sum_{v\in V} \sum_{a_1,a_2\in A^2} f(a_1,a_2,v)Y(a_1,a_2,v) \tag{1) \& (2}$$
$$\sum_{a_1,a_2\in A^2} N_h f(a_1,a_2,v) = N_v \qquad \forall v \in V \tag{3}$$
$$\sum_{v\in V} N_f^{asgmt}(a,v) = N_{f^{SQ}}^{asgmt}(a) \qquad \forall a \in A \tag{4}$$

We discuss its uniqueness and asymptotic properties in Sections A2.1 and A2.2.

5.3 Effect of the Optimal Assignment

After identifying the optimal assignment, we can estimate its effect on tax compliance by computing the Average Reallocation Effect (ARE, Graham et al. (2014)), i.e., the difference in average tax compliance between the optimal and the status quo assignment:

$$ARE = \sum_{v\in V} \sum_{a_1,a_2\in A^2} \left[f^*(a_1,a_2,v) - f^{SQ}(a_1,a_2,v)\right] Y(a_1,a_2,v) \tag{5}$$

In our setting, the status quo assignment consists of randomly assigning collectors to teammates and collector pairs to neighborhoods. We can therefore write the status quo assignment function as $f^{SQ}(a_1,a_2,v) = f_a^{SQ}(a_1)f_a^{SQ}(a_2)f_v^{SQ}(v)$, with $f_a(a)$ and $f_v(v)$ the marginal distribution of a-type collectors and v-type households, respectively.[37,38]

6 Estimation

6.1 Household Type

To characterize the optimal assignment function and estimate its effect on tax compliance, we first need to define household and collector types. By household type, we mean its pre-treatment propensity to pay the property tax.[39] We estimate household type using information

[37]The marginal distribution of a-type collectors is given by $f_a(a) = N_f^{asgmt}(a)/N^{asgmt}$ and corresponds to the share of total collector assignments assigned to a-type collectors. The marginal distribution of v-type households is given by $f_v(v) = N_h(v)/N_h$ and corresponds to the share of v-type households.

[38]For our main specification, the household type is characterized by $f_v^{SQ}(l) \approx 1/3$ and $f_v^{SQ}(h) \approx 2/3$ (cf. Section 6.1), the collector type by $f_a^{SQ}(H) = f_a^{SQ}(L) = 1/2$ (cf. Section 6.2), and the status quo assignment by $f^{SQ}(H,H,l) = f^{SQ}(L,L,l) = f^{SQ}(L,H,l) = f^{SQ}(H,L,l) = \frac{1}{4}f_v^{SQ}(l) \approx 1/12$ and $f^{SQ}(H,H,h) = f^{SQ}(L,L,h) = f^{SQ}(L,H,h) = f^{SQ}(H,L,h) = \frac{1}{4}f_v^{SQ}(h) \approx 1/6$.

[39]Unfortunately, we cannot use prior tax compliance because properties' unique tax ID numbers were reassigned during the registration phase of the 2018 property tax campaign (cf. Section 3).

provided by local city chiefs. As described in Section 3, in 78 neighborhoods of Kananga (the analysis sample), the local chief reported each property owner's ability to pay the property tax before tax collection began. During consultations with state collectors, chiefs went line by line through the neighborhood property roll, guided by the owners' names as well as photos of each compound. They reported whether each property owner was "unlikely," "likely," or "very likely" to have the economic ability to pay the property tax. As shown in Figure A1 and explored in more detail in Balan et al. (2022), chiefs' estimates were highly predictive of property tax payment, even controlling for household characteristics.

To maximize power, we primarily consider two household types: low types ($v = l$) are those deemed "unlikely" to be able to pay the property tax according to the chief, and high types ($v = h$) are those deemed "likely" or "very likely" to be able to pay.[40] According to this definition, about $2/3$ of households are high-type and about $1/3$ are low-type. Since we use chiefs' estimates to define household type, our characterization of the optimal assignment relies on the 78 neighborhoods in the analysis sample. Neighborhoods were randomly assigned to one of the two samples, and the 78 neighborhoods in the analysis sample are therefore identical to the 102 neighborhoods in the holdout sample where the same state collectors worked but consultations with the chief did not take place (Balan et al., 2022).

Although the chief predictions are the best available predictor of tax compliance,[41] defining household types using observable house and property owner characteristics might be preferable for some governments.[42] We thus alternatively define household type using such characteristics, and our main results remain robust (Section 8.2).

6.2 Collector Type

As is often true when seeking to assess worker value-added (Chetty et al., 2014), we have no informative pre-treatment measure of collector ability. We therefore estimate collector type in the 102 neighborhoods (covering 11,732 properties) of the holdout sample. This sample-splitting approach allows us to avoid estimating collector type and the average tax compliance function in the same sample, which could lead to overfitting and might mechanically generate complementarity in collector types (Mullainathan and Spiess, 2017).

In our main estimation approach,[43] we define collector c's ability as the average tax com-

[40] This is the most natural definition with two household types since the gap in tax compliance is larger between owners deemed "unlikely" and "likely" to pay than "likely" and "very likely" to pay (Figure A1).

[41] Indeed, the correlation between tax compliance and household type is higher when household type is based on chiefs' estimates (0.102) than when it is based on house characteristics from surveys (0.051).

[42] For instance, in other settings, chief jurisdictions might be too large to have useful information, or they might have a more competitive relationship with the formal state, making them reluctant to share their information.

[43] We also explore alternative approaches in Section 8.2, including partitioning collectors into more than two

pliance achieved across all randomly assigned neighborhoods in the holdout sample, i.e., $\mathbb{E}\left[Y(c_1, c_2, v_h)|c_1 = c\right]$. We estimate a fixed-effects model using OLS:

$$y_{hnt} = \sum_c \alpha_c 1_{[c \in C(n)]} + \lambda_t + \varepsilon_{hnt} \tag{6}$$

where y_{hnt} is an indicator for household h in neighborhood n paying the property tax during tax campaign month t. $C(n)$ is the vector of collectors assigned to work in neighborhood n, and $1_{[c \in C(n)]}$ is an indicator for whether tax collector c was assigned to collect taxes in neighborhood n. We include tax campaign month fixed effects λ_t because these were the randomization strata used during assignment.[44] We cluster standard errors at the neighborhood level (the level of assignment).

The coefficient of interest is $\boldsymbol{\alpha}$, the vector of collector fixed effects.[45] The OLS estimator of $\boldsymbol{\alpha}$ is unbiased but noisy since tax collectors worked with at most 6 teammates and in 12 neighborhoods during the 2018 property tax campaign. We improve the precision of our estimates by shrinking them to the overall mean based on the ratio of signal variance to total variance (Morris, 1983; Kane and Staiger, 2008). We denote $\boldsymbol{\hat{\alpha}^{EB}}$ the vector of Empirical Bayes estimates of collector fixed effects.[46] We report $\hat{\alpha}^{OLS}$ and $\hat{\alpha}^{EB}$ for the 34 collectors in Figure A2 (Panel A).

To motivate our investigation into collector assignments, we illustrate the importance of collectors in shaping tax compliance behavior. According to the $\hat{\alpha}_c^{EB}$ estimates, tax collectors explain 21% of the variance in tax compliance across neighborhoods.[47] By comparison, Best et al. (2019) find that bureaucrats who manage procurement processes in Russia explain 24% of the variation in quality-adjusted prices, and Fenizia (2022) finds that public-sector managers in Italy explain 9% of the total variation in the efficiency of filing insurance claims.

groups and estimating a nonlinear model following Bonhomme (2021).

[44] To identify $\mathbb{E}\left[Y_h(c_1, c_2, v_h)|c_1 = c\right]$, we subtract the average tax compliance across collectors, $\mathbb{E}\left[Y_h(c_1, c_2, v_h)\right]$, when including month fixed effects.

[45] Without time fixed effects, random assignment of collectors to teammates and to neighborhoods implies that $\alpha_c = \mathbb{E}\left[Y_h(c_1, c_2, v_h)|c_1 = c\right]$ in large samples. Because we include month fixed effects, α_c identifies a weighted average of collector c's enforcement ability in different months of the tax campaign (Abadie and Cattaneo, 2018). For simplicity, we assume that collectors' enforcement abilities are fixed over time, but note that we are still identifying a meaningful measure of collector ability if this assumption were violated.

[46] Relying on $\hat{\alpha}^{EB}$ vs. $\hat{\alpha}^{OLS}$ is unlikely to have large effects on the results because we shrink to the overall mean, and the distribution of $\hat{\alpha}^{OLS}$ has little skewness (Figure A2, Panel B). Indeed, 32 of the 34 collectors have the same type when defined based on $\hat{\alpha}^{EB}$ vs. based on $\hat{\alpha}^{OLS}$ for two types of collector.

[47] We compute $Var(2\hat{\beta}_c^{EB})/Var(\overline{Y}_n)$, where $Var(2\hat{\beta}_c^{EB})$ is the sample variance of the Empirical Bayes estimates across collector pairs and $Var(\overline{Y}_n)$ is the sample variance in average tax compliance across neighborhoods. Following Bonhomme (2021) and estimating a non-linear model in collector type also finds that tax collectors explain 21% of the tax compliance variance with two collector types (Table A17, Columns 1–2).

To define collector types, we partition collectors into discrete groups using $\hat{\alpha}_c^{EB}$.[48] The main analysis uses two types because according to a range of unsupervised machine learning approaches two clusters is optimal in our data (Section A3). A collector's type depends on the rank of their $\hat{\alpha}_c^{EB}$ estimate, denoted $r_c = \text{rank}(\hat{\alpha}_c^{EB})/N_c$. Collectors with $r_c < 0.5$ are low-type, while those with $r_c > 0.5$ are high-type.[49]

This non-parametric approach to ranking collectors — based on the tax compliance they achieved across randomly assigned neighborhoods — has the advantage of remaining agnostic about the underlying average tax compliance function.[50] It is possible that equation (6) is misspecified, i.e., tax compliance might not be additive in collector ability. But this would not compromise our objective, which is to define a sensible metric for collector ability and study the effect of the optimal assignment while making as few assumptions as possible.[51] We probe the accuracy of our approach using Monte Carlo simulations (Section A4). Across 1,000 simulated samples, estimating collector type using equation (6) and partitioning collector into two groups using $\hat{\alpha}_c^{EB}$ recovers true collector types about 81% of the time when the average tax compliance function exhibits complementarities in collector type — i.e., when equation (6) is misspecified — vs. about 82% of the time when the average compliance function is linear in collector type — i.e., when equation (6) is correctly specified (Table A13, Panel A, Columns 1–2). Potential misspecifications of equation (6) thus do not much af-

[48] This approach allows us to estimate the average tax compliance function non-parametrically (cf. Section 6.3) and is similar to the methodology used by Bhattacharya (2009) and Graham et al. (2020a).

[49] High-type collectors differ from low-type collectors in many ways beyond their ability to collect taxes (Table A4). They are on average more educated (0.51 more years of schooling) and have higher monthly income prior to the campaign ($61). They are also more likely to believe that taxes are important for development, and less likely to have a relative who works for the provincial government.

[50] For instance, consider the case where tax collectors are horizontally differentiated (e.g., by ethnicity), and matching collectors on ethnicity would increase tax compliance. Under this particular functional form — one of many possible average tax compliance functions — it is possible that the government could do better than our optimal assignment by explicitly matching on ethnicity. However, this functional form would not invalidate our estimates of the optimal assignment based on collectors' observed compliance rank, $\text{rank}(\hat{\alpha}_c^{EB})/N_c$. As randomization ensures that horizontal differences — in this example, ethnicity — are uncorrelated with collector assignments, $\text{rank}(\hat{\alpha}_c^{EB})/N_c$ will still capture a meaningful signal of collector effectiveness and allow us to estimate the impact of the optimal assignment based on this measure. Moreover, which specific trait could be useful for matching will depend on the context, while our approach is more generalizable.

[51] Since the average compliance function exhibits complementarities in collector type (cf. Section 7), we might wrongly attribute some of these complementarities to high-type collectors' abilities. The $\hat{\alpha}_c^{EB}$ estimates would then be upward-biased for high-type collectors. This would be a concern if our objective were to precisely estimate the value added (i.e., fixed effect) associated with each tax collector. But all we seek is a sensible ranking of collectors, and random assignment ensures that any such bias would be distributed uniformly among high type collectors and thus not affect our partitioning of types. Another potential source of bias in this scenario is that high-type collectors could be wrongly classified as low-type if they are only assigned to low-type teammates in small samples. Such misclassification is possible but would lead to underestimating the impacts of the optimal assignment.

fect our estimates of collector type. Misclassifications of collector types, which arise with or without misspecification of equation (6), simply bias our estimates of the effect of optimal matching toward zero. Our estimates should be interpreted as a lower bound.

6.3 Average Tax Compliance Function

Having defined household and collector types, we then estimate the average compliance function $Y(a_1, a_2, v)$ in the analysis sample (78 neighborhoods, 6,904 properties). We follow Bhattacharya (2009) and Graham et al. (2020a) and estimate it non-parametrically using the following regression:

$$y_{hnt} = \sum_{a_1 \in A} \sum_{a_2 \geq a_1} \sum_{v=l,h} \beta(a_1, a_2, v) \cdot 1_{[c(n)=(a_1,a_2)]} \cdot 1[v_h = v] + \lambda_t + \varepsilon_{hnt} \tag{7}$$

where y_{hnt} is an indicator for household h in neighborhood n having paid the property tax during campaign month t. $1_{[c(n)=(a_1,a_2)]}$ indicates whether neighborhood n was assigned to a pair of collectors with types a_1 and a_2, and $1[v_h = v]$ indicates whether household h is of type v. In our preferred specification, equation (7) includes the five dummies $(H, H, h), (L, H, h), (L, L, h), (H, H, l), (L, H, l)$, and the excluded category is (L, L, l), reflecting matches of households of type $V = \{l, h\}$ and collectors of type $A = \{L, H\}$. As before, we include tax campaign month fixed effects λ_t.[52]

Our main specification reports standard errors clustered at the neighborhood level. Because collector type might be imprecisely estimated due to the finite sample size, we also calculate standard errors using Bayesian bootstrap re-sampling (Rubin, 1981) at the neighborhood level (Figure A4). By re-sampling neighborhood weights and using weighted least squares, the Bayesian bootstrap is better suited to our context than the standard bootstrap given that collectors work in pairs.[53]

[52] Section A5 discusses the interpretation of $\hat{\beta}(a_1, a_2, v)$ with campaign month fixed effects.

[53] Our problem can be viewed as part of the class of "pairwise agreement" problems, in which the analyst seeks to estimate the value of an object assessed by multiple judges, each of whom have their own fixed effects. In this class of problems, the standard bootstrap is typically unsuitable because taking random subsamples reduces the number of objects observed across judges and thus impedes one's ability to separate out judge-specific effects. In our setting, a neighborhood is equivalent to a judge. Each neighborhood dropped decreases the precision with which we identify the fixed effects of the two assigned collectors, as well as the fixed effects of other collectors with whom they were assigned (Efron, 1982). By randomly sampling neighborhood weights in each iteration, which does not require dropping neighborhoods altogether, the Bayesian bootstrap is preferable in our setting (Rubin, 1981).

6.4 Counterfactuals: Optimal Assignment and its Effects

We now turn to the estimation of the counterfactual optimal assignment function f^*. Following Bhattacharya (2009) and Graham et al. (2020a), plug the estimated average tax compliance function, $\hat{\beta}(a_1, a_2, v)$, into the empirical analog of the Optimal Assignment Problem (Problem 1):[54]

Problem 2. *Empirical Optimal Assignment*

$$\widehat{f^*} \equiv \arg\max_f \sum_{v \in V} \sum_{a_1, a_2 \in A^2} f(a_1, a_2, v) \widehat{\beta}(a_1, a_2, v) \tag{8}$$

$$\sum_{a_1, a_2 \in A^2} N_h f(a_1, a_2, v) = N_v \qquad \forall v \in V \tag{9}$$

$$\sum_{v \in V} N_f^{asgmt}(a, v) = N_{f^{SQ}}^{asgmt}(a) \qquad \forall a \in A \tag{10}$$

We then use the counterfactual optimal assignment function and average tax compliance function to estimate the ARE:

$$\widehat{ARE} = \sum_{v \in V} \sum_{a_1, a_2 \in A^2} \left[\widehat{f^*}(a_1, a_2, v) - f^{SQ}(a_1, a_2, v) \right] \widehat{\beta}(a_1, a_2, v) \tag{11}$$

As noted, we report standard errors clustered at the neighborhood level and show robustness to Bayesian bootstrapped standard errors.

7 Optimal Assignment

7.1 Characterizing the Optimal Assignment

The optimal assignment of tax collectors to teammates and households is not ex-ante obvious.[55] When forming collector teams, if only one high-type collector is required to ensure that the essential tax collection tasks are completed, pairing high-type collectors with low-type collectors (mixed teams) would likely be optimal. However, complementarities in collector

[54] $\hat{\beta}$ identifies Y up to a constant but the solution to Problem 1 is the same if we substitute Y for $Y + c$. Indeed, $(\mathbf{Y} + c)'\mathbf{f} = \mathbf{Y}'\mathbf{f} + c$ since f is a probability mass function.

[55] Past empirical work reaches mixed conclusions. Carrell et al. (2009) predict that negative assortative matching of students would improve test scores, but Carrell et al. (2013) find the opposite in real life. Bhattacharya (2009) finds that positive assortative matching of students in dorms has little average impact on test scores. Aucejo et al. (2019) and Graham et al. (2020a) find evidence of complementarities between student and teacher attributes. Thakur (2020), Limodio (2021) and Fenizia (2022) provide evidence of negative assortative matching in public bureaucracies.

effort or skill could justify grouping high types and low types together (homogeneous teams). The assignment of collector pairs to households is also ambiguous. If collection from high-type households is easier than from low-type households — e.g., if it only involves showing up and soliciting payment — it could be optimal to assign them to low-type collectors. Alternatively, if it requires conscientiousness in making follow-up visits and skills of persuasion, the government may prefer to assign them to high-type collectors.

According to our estimates, we find that the counterfactual optimal assignment is characterized positive assortative matching on the collector-collector and collector-household dimension.

Collector-Collector Assignment. The average tax compliance function estimated using equation (7) appears to exhibit complementarities in collector type for high-type households (Figure 1). For these households, *H-H* collector pairs achieve 9.5 percentage point higher tax compliance than *L-L* pairs. By contrast, *L-H* pairs achieve only 1.5 percentage point higher compliance than *L-L* pairs. As a more formal test of complementarity (Topkis, 1998), we look for increasing differences. Specifically, we test the hypothesis H_1:$[Y(H,H,h) - Y(L,H,h)] - [Y(H,L,h) - Y(L,L,h)] > 0$ against H_0:$[Y(H,H,h) - Y(L,H,h)] - [Y(H,L,h) - Y(L,L,h)] \leq 0$ and report the associated p-value. This test confirms that the average tax compliance function exhibits complementarities in collector type for high-type households ($p = 0.037$).[56] We find the same pattern of complementarities in collector type when analyzing tax revenue as the outcome ($p = 0.090$, Figure A3).[57] Tests that use standard errors from Bayesian bootstrap re-sampling to account for sampling noise in the estimation of collector type return similar though slightly weaker evidence of complementarity in collector type ($p = 0.109$ for compliance, $p = 0.174$ for revenue, Figure A4).

In light of these complementarities in collector type, the counterfactual optimal assignment function estimated using Equations (8)-(10) involves positive assortative matching of collectors. The government should only form pairs of high-type collectors (*H-H* teams) or pairs of low-type collectors (*L-L* teams). There should be no mixed (*L-H*) teams (Figure 2). This contrasts with the status quo assignment with 25% *L-L* teams, 50% *L-H* teams, and 25% *H-H* teams, due to random assignment. The random reshuffling of collectors into new teams each month, which may have anti-collusion benefits, would be preserved under the optimal

[56] Alternatively, we can test whether the average tax compliance function exhibits non-linearities in collector type. To do so we test H_1: $[Y(H,H,v) - Y(L,H,v)] - [Y(H,L,v) - Y(L,L,v)] \neq 0$ against H_0: $[Y(H,H,v) - Y(L,H,v)] - [Y(H,L,v) - Y(L,L,v)] = 0$. This test confirms that $Y(a_1, a_2, v)$ exhibits non-linearities in collector type for high-type households ($p = 0.074$).

[57] In this context tax revenue is obtained by multiplying tax compliance by the tax liability and thus mechanically results in less precise estimates and slightly weaker evidence of complementarity in collector type.

assignment. For example, the 17 high-type (low-type) collectors would be randomly matched with a high-type (low-type) teammate every campaign month.

Collector-Household Assignment. The average tax compliance function estimated using equation (7) also appears to exhibit complementarities in collector-household type (Figure 1). An *H-H* collector pair achieves 13.5 percentage point higher tax compliance when assigned to high-type household than when assigned to a low-type household. By contrast, an *L-L* pair achieves only 3.5 percentage point higher tax compliance. We again test the hypothesis H_1: $[Y(H,H,h) - Y(L,L,h)] - [Y(H,H,l) - Y(L,L,l)] > 0$ against the hypothesis H_0: $[Y(H,H,h) - Y(L,L,h)] - [Y(H,H,l) - Y(L,L,l)] \leq 0$. There is evidence of increasing differences and thus complementarities in collector-household match type ($p < 0.001$).[58] The same pattern of complementarities in collector-household type applies to tax revenue ($p = 0.004$, Figure A3). These results are robust to calculating standard errors using Bayesian bootstrap re-sampling ($p = 0.004$ for compliance and $p = 0.013$ for revenue, Figure A4).

The optimal assignment function thus also involves positive assortative matching of collector pairs to households. The government should assign *H-H* pairs to high-type households and *L-L* pairs to low-type households. In our context, this would mean allocating all low-type households ($1/3$ of households) to *L-L* pairs and allocating high-type households ($2/3$ of households) first to *H-H* pairs and then to the remaining *L-L* pairs (Figure 2). This contrasts with the status quo (random) assignment where 25% of both household types are allocated to *L-L* pairs, 50% to *L-H* pairs, and 25% to *H-H* pairs.

7.2 Mechanisms

Before turning to the impact of the optimal assignment policy (Section 8), we first explore mechanisms behind the complementarities in collector type and collector-household type. We primarily focus on collector skill and effort but also discuss other possible mechanisms in Section A6.[59]

Collector Skill. A first possible mechanism is that *H-H* pairs were more skillful in convincing households to pay. The in-person mode of tax collection left much to the discretion of the collectors, including what types of messages and other persuasion techniques to use. It could be that high-type collectors were significantly more credible and convincing when paired with other high types. We examine two types of evidence, which ultimately find little support for this mechanism.

[58] We also find more general evidence of non-linearities in collector-household match type ($p = 0.001$).

[59] Section A6 explores homophily and social incentives. We show that these channels are unlikely to explain the complementarities in collector type and collector-household type documented in Section 7.1.

First, we study households' post-taxation beliefs about enforcement and tax morale. If *H-H* pairs were more skilled in shaping property owners' beliefs and thus persuading them to pay, we would expect to find that households randomly assigned to *H-H* pairs would perceive a higher probability of enforcement among delinquent properties after the campaign. Using midline survey data (collected after tax collection was completed in each neighborhood), we find no evidence of complementarities in collector type for citizens' perceived probability of sanctions for tax delinquency (Figure A5, Panel A).[60] Similarly, *H-H* teams do not appear to differentially increase citizens' perceptions that tax revenues are spent on public goods relative to *L-H* and *L-L* pairs (Figure A5, Panel B).[61]

Second, we investigate the specific messages property owners recalled collectors using when trying to convince them to pay. Although recall is likely imperfect, endline survey respondents reported collectors using a range of messaging relating to sanctions, public goods provision, trust in the authorities, social pressure, and legal obligation. We examine if *H-H* pairs relied on different messages than *L-H* and *L-L* pairs, but find no evidence of complementarities in this dimension (Figure A6).[62] It thus appears unlikely that the complementarities we observe reflect differential collector skill in persuading property owners to pay by deploying certain types of messages or otherwise changing their beliefs about tax enforcement or public goods spending (tax morale).

Collector Effort. A second explanation is that high-type collectors exerted greater effort when matched with high-type teammates (e.g., Mas and Moretti, 2009; Brune et al., 2020). To explore this possibility, we investigate the number of days and hours collector pairs worked in assigned neighborhoods by combining two sources of data: (*i*) dated chalk marks that collectors were instructed to leave on the wall of the properties that they visited after registration and were recorded by enumerators during the midline survey, and (*ii*) the date and time of visits that led to a tax payment, which was systematically recorded by the tax receipt data.[63] Although collectors were supposed to work for the entire tax campaign month in each as-

[60] High-type collectors are associated with a higher average perceived probability of sanctions for tax delinquency, but there is no difference between *L-H* and *H-H* pairs. Complementarity tests find no evidence of increasing differences in collector type ($p = 0.964$) or collector-household type ($p = 0.268$).

[61] Complementarity tests again find no evidence of increasing differences in collector type ($p = 0.993$) or collector-household type ($p = 0.183$).

[62] We find no evidence of increasing differences in collector type (p-values between 0.219 and 0.993) or collector-household type (p-values between 0.149 and 0.794).

[63] We do not directly observe the number of hours the tax collectors worked in a neighborhood. Instead, we proxy for it by multiplying the number of days worked by the average number of hours worked per day. We calculate the average number of hours worked per day as the average number of hours between the first and last payment for that neighborhood. We rely on the tax data for this calculation since the dated chalk marks left by tax collectors did not indicate the time of the visit.

signed neighborhood, whether they did so and for how long were left to their discretion. We find evidence that *H*-*H* collector pairs worked on more days and hours than *L*-*H* and *L*-*L* pairs and that days and hours worked exhibit complementarities in collector type (Figure A7).[64]

The number of days and hours worked measured using the chalk marks and tax receipt data offers an objective measure of collector effort. However, low-type collectors might have used chalk marks less regularly than high-type collectors, which would mean that Figure A7 could overstate the extent to which *H*-*H* pairs' performance is explained by effort.[65] For this reason, we also examine midline survey questions about the number of visits by tax collectors after property registration. Although this variable is self-reported and subject to imperfect recall, it provides a useful supplementary measure of collectors' effort. According to this measure, *H*-*H* pairs conducted more visits than *L*-*H* and *L*-*L* pairs, both on the extensive margin (Figure A8, Panel A) — the share of households that received a visit — and on the intensive margin (Figure A8, Panel B) — the number of visits per household — although the increase appears to be linear (rather than convex) in collector type.[66]

Why would collecting taxes on more days and more hours increase tax compliance? One explanation is that it might increase the chances that property owners had the cash on hand to pay the tax when the collectors solicited payment.[67] The impact of liquidity constraints on tax compliance has been well-documented, even in middle- and high-income countries like Mexico and the United States (Brockmeyer et al., 2021; Wong, 2020). If property owners in a low-income setting like Kananga faced time-varying cash on hand constraints, then collector visits on different days and at different times might have increased the probability that property owners had cash on hand when collectors visited.

We provide two pieces of evidence consistent with this interpretation. First, we examine heterogeneity in collector effort by neighborhood employment rate. Property owners with some source of employment are more likely to have cash on hand than the unemployed. If the additional days and hours of tax collection by *H*-*H* pairs boosted tax compliance by relaxing time-varying cash-on-hand constraints, then the increase in collector effort should have been

[64]The coefficients for *H*-*H* pairs are not significantly different from the coefficients for *L*-*H* and *L*-*L* pairs at conventional levels. However, tests of increasing differences in collector type show complementarity in collector type ($p = 0.032$ for days worked and $p = 0.051$ for hours worked) and in collector-household type ($p = 0.078$ for days worked and $p = 0.097$ for hours worked).

[65]Collectors were instructed to leave a dated chalk mark for *all* visits, already helps alleviate this concern.

[66]Complementarity tests find no evidence of increasing differences for visit indicator and number of visits in collector type ($p = 0.520$, $p = 0.131$) or collector-household type ($p = 0.712$, $p = 0.336$).

[67]Another possibility is that receiving more visits from tax collectors affected citizens' beliefs about enforcement. Receiving more frequent visits could have increased owners' perception that the government will sanction tax delinquents. However, this does not appear to be the primary explanation in this setting since taxpayers' enforcement beliefs do not appear to exhibit complementarities in collector type (Figure A5).

concentrated in neighborhoods with higher employment rates where such constraints are less likely to always bind. The data bear out this prediction (Figure A11). Second, making visits at different points in time and later in the day could increase the probability that collectors arrive at a moment when the property has cash on hand. Using the tax receipt data to estimate the average time of collection by collector type, we find suggestive evidence that *H*-*H* collected taxes over a longer period in the day than *L*-*H* or *L*-*L* pairs (Figure A12). *H*-*H* pairs thus appear to raise more revenue because their higher effort levels in effect increase the probability that they visit property owners on days and times when they have the cash on hand to pay.

Why didn't all collector teams work for longer hours if it resulted in higher tax revenue and collector compensation? The observed differences in collector effort are unlikely to reflect a lack of knowledge about collection strategies since tax ministry supervisors stressed that making more and later visits could increase compliance during training. Thus, rather than a knowledge gap, the reason likely concerns a coordination problem. Collectors viewed tax collection as a joint task and strongly preferred working together than alone.[68] Thus, if their partner were unreliable and did not show up for work on time (or at all), even a high-type collector might choose not to work that day. In other words, the tax compliance functions may be O-ring in collector type (Kremer, 1993). Such coordination issues are a common feature of joint production tasks (Alchian and Demsetz, 1972; Olson, 1989), which often characterize the work of public and private sector frontline agents in developing countries (e.g., Burgess et al., 2010; Khan et al., 2016; Ashraf and Bandiera, 2018).

In sum, *H*-*H* pairs appear to have achieved disproportionately higher tax compliance than other pairs by by collecting taxes on more distinct days and for longer total hours. Moreover, they appear to direct their higher enforcement effort toward neighbourhoods where cash-on-hand constraints are less likely to bind and at times of the day when property owners are likely to have cash on hand. This capacity of *H*-*H* pairs likely reflects their ability to solve the coordination problem inherent in team-based tax collection, rather than by overcoming knowledge constraints or other frictions

8 Impact of the Optimal Assignment

We now estimate the increase in tax compliance and revenue under the counterfactual optimal assignment policy. We then examine a series of robustness checks, study potential endoge-

[68] Indeed, very few collectors worked alone according to the data. As noted in Section 2, team preference likely reflects (*i*) feelings of insecurity in handling sums of money on the field when alone, and (*ii*) the belief that property owners will be more likely to pay because in a pair collectors can more credibly assure them the money will reach the state or threaten sanctions.

nous responses to implementing the counterfactual optimal assignment, study impacts on bribes and tax morale, and explore distributional implications.

8.1 Main Results

According to our estimation approach, implementing the optimal assignment policy would result in a 2.941 percentage points ($p = 0.024$) increase in tax compliance, a 37% increase in tax compliance relative to the status quo average of 8% (Table 1, Row 1, Column 1). Implementing the optimal assignment would also lead to a 54.471 CF increase in tax revenue per owner ($p = 0.074$), a 26% increase from the status quo average of 206.213 CF (Row 1, Column 2). These effects remain significant when using standard errors from Bayesian bootstrap re-sampling to account for sampling noise in the estimation of collector type (Table A5, Columns 3 and 4).[69] As discussed in the previous section, the increase in compliance and revenue reflects the complementarities in collector type and collector-household type (Figure 1), which are fully exploited by the optimal assignment policy (Figure 2).

To assess how each margin of the optimal assignment policy — collector-collector and collector-household — would contribute to the total effect of the policy, we estimate the return to optimizing on each of these margins individually (Figure A13 and Table 1, Rows 2–3). Optimizing the assignment of collectors to teammates but assigning teams to households at random would increase compliance by 1.294 percentage points ($p = 0.172$) (Row 2, Column 1) and tax revenue per owner by 21.444 CF ($p = 0.322$) (Row 2, Column 2), a 16% and 10% increase, respectively. Similarly, optimizing the assignment of collectors to households but forming collector teams at random would increase tax compliance by 0.837 percentage points ($p = 0.007$) (Row 3, Column 1) and revenue per owner by 17.156 CF ($p = 0.044$) (Row 3, Column 2), a 10% and 8% increase, respectively. In sum, both assignment margins contribute to raising tax compliance, and the government would do substantially better by jointly optimizing.

8.2 Robustness Checks

We examine a number of alternative estimation approaches and robustness checks.

Number of Collector Types. One potential concern is that our results — i.e., the complementarities in the average tax compliance function and the impact of the optimal assignment — depend on the number of types used in the analysis. For example, if the average tax compliance function does not exhibit increasing differences locally, it might not be optimal to

[69] Bayesian bootstrap re-sampling at the neighborhood level results in slightly higher p-values: 0.080 for tax compliance (Table A5, Row 1, Columns 3) and 0.150 for tax revenue (Table A5, Row 1, Columns 4).

match like types together with more granular collector and household types.

To shed light on this possibility, we show that the results are robust to using more collector types.[70] Since the methodology outlined in Section 5 and 6 relies on sample splitting, we rapidly run into statistical power constraints when using more collector types. For completeness, we still show results with three collector types: low (L), middle (M), and high (H). Despite the low number of observations for some types of collector pairs,[71] we still find evidence of complementarities in collector type and collector-household type (Figure A14).[72] Further validation comes from implementing the non-linear methodology suggested by Bonhomme (2021), which detects complementarities for up to six collector types, as we note in the next subsection.

Relatedly, the estimated impact of the optimal assignment could be affected by the number of types used in the analysis. We investigate this issue in Section A7 by using our conceptual framework to show that the effect of the optimal assignment increases in the number of collector types used in the analysis. The intuition is that the efficiency of optimal matching improves as you increase the number of collector types. We provide evidence consistent with our theoretical result by comparing the effect of the optimal assignment with two and three collector types. With three collector types, implementing the optimal assignment would increase tax compliance by 4.411 percentage points ($p = 0.032$) and tax revenue per owner by 62.212 CF ($p = 0.202$), a 55% and 30% increase, respectively (Table A6).[73] This is higher than the 37% increase in compliance and the 26% increase in revenue associated with the optimal assignment for two collector types (Table 1). In other words, our main specification, which uses two types of collectors and households, can be considered a lower bound on the

[70] We focus on collector types because the number of household types is limited by the categories neighborhood chiefs used when eliciting property owners' ability to pay ("unlikely", "likely", or "very likely"). Although we could partition households into three types, our statistical power is low when estimating the average tax compliance function for the 952 "very likely" households (235 of which were assigned to *L-L* pairs, 582 to *L-H* pairs, and only 135 to *H-H* pairs).

[71] With three collector types $A = \{L, M, H\}$, the average compliance function is estimated in the analysis sample from 2 neighborhoods assigned to *L-L* pairs, 28 to *L-M* pairs, 22 to *L-H* pairs, 4 to *M-M* pairs, 16 to *M-H* pairs, and 6 to *H-H* pairs. By contrast, with two collector types $A = \{L, H\}$ the average compliance function is estimated from 18 neighborhoods assigned to *L-L* pairs, 44 to *L-H* pairs, and 16 to *H-H* pairs.

[72] As is the case with two collector types, pairs with a low-type collector perform considerably worse. The optimal assignment would consists of constituting *L-L* pairs and assigning them to low-type households and constituting *M-H* pairs and assigning them to high-type households. The pairing of middle-type collectors with high-type collectors is explained by *M-H* pairs outperforming *M-M* and *H-H* pairs. However, the comparison of *M-M*, *M-H*, and *H-H* pairs should be interpreted cautiously given the low number of observations: 4 neighborhoods assigned to *M-M* pairs, 16 to *M-H* pairs, and 6 to *H-H* pairs.

[73] While increasing the number of collector types mechanically improves the efficiency of collector assignment (Section A7), it leads to noisier estimates of the average tax compliance function and of the optimal assignment (Table A6).

effect of the optimal assignment with a higher number of types.

Alternative Nonlinear Methodology. As another way to explore robustness to more collector types — and as a robustness check in its own right — we also estimate a non-linear model in collector type following Bonhomme (2021). Specifically, we use a finite mixture approach where the distribution of discrete collector type is modeled using random-effects and estimated using mean-field variational methods. The advantage of this approach is that it does not involve estimating collector type in a first step and therefore does not require splitting the sample into a holdout and analysis sample. As a consequence this alternative approach is more powered to estimate the average tax compliance function and the optimal assignment with a higher number of collector types.[74] As described in detail in Section A8, this approach also finds complementarities in collector type with two and three types of collectors (Figure A21, Panel B). It also finds that positive assortative matching by collector type is optimal (Figure A21, Panel C).[75] In fact, this approach detects complementarities and finds that positive assortative matching is optimal for up to six collector types (Figure A22, Panels B-C).[76] Overall, these results suggest that the complementarities noted in Section 7 are unlikely to depend on the number of collector types — and more generally validates the key results from our preferred estimation strategy.

Alternative Definition of Collector Type. One concern with our approach is that low-capacity governments may lack the human capital or the resources to estimate collector type using a fixed effects model. In practice, the government might instead estimate collector types by observing which collector characteristics are correlated with performance in past tax campaigns. To approximate this simpler approach, we predict collector type based on the relationship between tax compliance and collector characteristics in the holdout sample.[77] With this alternative definition of collector type, we still observe complementarities in collector type and in collector-household type for tax compliance (Figure A15, Panel A) and tax revenue (Figure A15, Panel B), which means that the optimal assignment is the same

[74] This is not our preferred approach because (*i*) it does not allow us to exploit both dimensions of random assignment (collector-collector and collector-household), (*ii*) it is unclear how to assess mechanisms using this approach, and (*iii*) our primary approach is more standard in the applied microeconomics literature on optimal matching (Bhattacharya, 2009; Graham et al., 2020b; Aucejo et al., 2019).

[75] Additionally, Bonhomme (2021)'s methodology estimates type proportions of 47% and 53% with two collector types (and 43%, 32%, and 25% with three types), which is close to the proportions we adopt when using our preferred approach in Sections 5 and 6.

[76] Given the finite sample, there are of course some small departures from complete assortative matching with five and six collector types. Indeed, with five or six collector types there are few observations per collector pair, and the average compliance function is imprecisely estimated (Figure A22, Panel A).

[77] We focus on the collector characteristics described in Panel A of Table A4: gender, age, ethnicity, level of education, math score, literacy, income, and possessions.

as in Figure 2.[78] Implementing the optimal assignment using this alternative collector type definition would be associated with a 2.688 percentage points increase in tax compliance ($p = 0.030$) and a 56.926 CF increase in tax revenue per owner ($p = 0.048$), a 34% and 28% increase, respectively (Table 1, Columns 3 and 4).

Alternative Definition of Household Type. Similarly, in practice the government may lack the ability to consult neighborhood chiefs about each household's ability to pay.[79] Instead, the government might estimate household type using observable characteristics. To approximate this approach, we run an OLS regression of compliance on household characteristics in the holdout sample and use the regression coefficients to predict households' tax compliance in the analysis sample. We then define household type (high or low) based on whether they rank below or above the median in terms of predicted tax compliance.[80] Under this alternative definition, we still find evidence of complementarity in collector type and in collector-household type for tax compliance (Figure A16, Panel A) and tax revenue (Figure A16, Panel B).[81] As a consequence, the optimal assignment again involves positive assortative matching (Figure A17). Implementing the optimal assignment using this alternative household type definition would be associated with a 2.759 percentage points increase in tax compliance ($p = 0.067$) and a 50.417 CF increase in tax revenue per owner ($p = 0.148$), a 34% and 24% increase, respectively (Table A7, Columns 3–4).

Government Objective. Thus far, we have assumed that the government's objective function is to maximize tax compliance.[82] However, a government might instead prefer to maximize tax revenue. The results are similar when focusing on this alternative government objective (Table A8, Column 1). The revenue-maximizing assignment policy would increase tax revenue per owner by 61.014 CF ($p = 0.020$), i.e., a 30% increase, which is larger than the increase in revenue per owner obtained when maximizing tax compliance (54.471 CF),

[78] Formal tests show complementarity in collector type for tax compliance and revenue ($p = 0.069$ and $p = 0.051$) and in collector-household type for the same outcomes ($p = 0.001$ and $p = 0.010$).

[79] Arranging these chief consultations might be prohibitively costly for some low-capacity governments. Alternatively, some settings might not have similarly locally embedded chiefs, or these chiefs might have a more competitive relationship with the formal state (Henn, 2021).

[80] We focus on the characteristics described in Panel A of Table A2: distance to state buildings, to health institutions, to education institutions, to roads, to eroded areas and property value. We omit wall, roof, and fence quality due to the lower number of observations for these characteristics and because they are highly correlated with property value (0.661, 0.510, and 0.260, respectively.)

[81] Formal tests show complementarity in collector type for tax compliance and revenue ($p = 0.099$ and $p = 0.153$) and in collector-household type for the same outcomes ($p = 0.040$ and $p = 0.073$).

[82] We focus on tax compliance rather than revenue as the government's objective because it is more precisely estimated. Indeed, revenue is equal to tax compliance multiplied by a constant (tax liability) and thus a noisier object for the optimization problem.

although the two are not statistically different.[83]

Neighborhood Level Assignment. One concern with the household-level assignment is that sending collectors to different households throughout the city could have high administrative costs (because collectors would need to travel to multiple neighborhoods per day, for instance). Neighborhood-level assignment, which we describe in Section A9, might therefore be more policy relevant. Table A9 shows results for two neighborhood-level optimal assignment policies: categorizing neighborhoods as high- or low-type based on their share of high- and low-type households (Columns 1–2) or their number of high- and low-type households (Columns 3–4). Implementing the optimal assignment would increase tax compliance by 1.764 percentage points ($p = 0.085$) under the first definition (Column 1) and by 2.906 percentage points ($p = 0.048$) under the second definition (Column 3). This latter estimate is just shy of that from our main specification involving household-level assignments (2.941 percentage points). One reason is that using the number of high and low-type households to define neighborhood-type would allow the government to relax the workload constraint by collector type in equation (4). Thus, high-type collectors would be assigned to more households under this neighborhood-level assignment than under the status quo assignment.[84] Taking neighborhoods' size into account thus allows the government to increase the number of high-type households assigned to *H-H* teams and to achieve 99% of the compliance gains of the optimal household-level assignment.

Overfitting and the Winner's Curse. Another concern is that estimating the tax compliance function and the impact of the optimal assignment in the same sample might lead to overfitting, i.e., we might be selecting the optimal assignment based on noise.[85] In par-

[83] We also consider the objective of tax revenue net of bribes in Columns 3–4 of Table A8. Optimizing with this goal reduces the probability of bribe payments but also reduces revenue effects. Given that the optimal assignment would have no increase in bribes per visit (cf. Section A11.1), we think it is unlikely the government would ultimately select this objective.

[84] Differences in the numbers of assignments by collector type are relatively small: (*i*) under the status quo assignment (and the household-level optimal assignment), high and low-type collectors are assigned to 3,452 households in the analysis sample; (*ii*) under the neighborhood-level optimal assignment – defined using the share of high and low-type households in the neighborhood – *H-H* pairs are assigned to 3,344 households (696 low-type and 2,648 high-type), and *L-L* pairs are assigned to 3,560 households (1,610 low-type and 1,950 high-type); (*iii*) under the neighborhood-level optimal assignment – defined using the number of high and low-type households in the neighborhood – *H-H* pairs are assigned to 3,704 households (1,107 low-type and 2,597 high-type), and *L-L* pairs are assigned to 3,200 households (1,309 low-type and 1,891 high-type). Additionally, we find no evidence that collectors face binding time constraints or that they visit a smaller share of households in larger neighborhoods (Section A10.1). These results suggest that the larger assignment load to high-type collectors under the neighborhood-level assignment that considers the number of high and low-type households in each neighborhood is unlikely to cause collector exhaustion, which would lower the effect of implementing the neighborhood-level optimal assignment policy.

[85] This problem should be minor in our context since our model has few variables: five dummies for the different

ticular, because we select the best of many possible assignments using tax compliance by household and collector types, which is imprecisely estimated, the effect of the optimal assignment could be biased upward. This is an example of the "winner's curse" in optimization problems, which we overcome by implementing the methodology introduced by Andrews et al. (2021).[86] Table A10 reports conditional and hybrid median-unbiased estimators and optimal confidence intervals that are valid conditional on the policy selected and so overcome this winner's curse.[87] Reassuringly, the estimated impacts of the policy on tax compliance — 2.897 for the conditional estimator and 2.890 for the hybrid estimator — are similar to our baseline estimate (2.941) and statistically significant at the 10% level. We also find similar results when using tax revenue maximization as the objective.

8.3 Endogenous Responses to Implementing the Optimal Assignment

The impact of the optimal assignment described above implicitly assumes that the average tax compliance function is unaffected by changes in the assignment function. This assumption is essential for the implementation of the optimal policy to have the effects documented in Sections 8.1-8.2. However, past work suggests that this assumption might not always hold and that implementing the predicted optimal assignment sometimes backfires. For instance, Carrell et al. (2013) predicted that mixed squadrons of high and low ability cadets would improve test scores of low ability cadets. Yet, when implemented, mixed squadrons lowered test scores for low-ability cadets because cadets endogenously sorted into segregated subgroups of low and high types within the mixed squadrons. These findings demonstrate the importance of carefully investigating potential endogenous responses from implementing the optimal assignment.

Fortunately, the issues encountered in Carrell et al. (2013) are less concerning in our setting. First, while Carrell et al. (2013) study squadrons of 32 cadets, we study teams of two

combinations of collector and household types, and three campaign month dummies. The small number of variables included in the model restricts the degrees of freedom we have to fit noise.

[86] Another solution would be to split the sample in three (instead of two), enabling us to estimate the impact of the optimal assignment out-of-sample. However, this approach would require splitting our analysis sample of 78 neighborhoods in two, which would be costly in terms of power.

[87] Andrews et al. (2021) only applies to discrete policy spaces. We adapt their approach to the context of our non-discrete optimal assignment policy space in several steps. First, the solution must lie at the intersection of three hyper-planes defined by the two linearly independent constraints in Problem 1 and the requirement that the distribution probabilities sum up to 1. Second, the Fundamental Theorem of Linear Programming (Dantzig, 1951) — which states that if an optimal solution exists, there exists an optimal solution consisting of extreme points on the policy space — allows us to select three points in this 3 dimensional space. We focus on the three solutions in the (finite) set of extreme points that are linearly independent and yield the highest value when applied to the objective function. We are deeply grateful to Toru Kitagawa for helpful discussions on how to adapt Andrews et al. (2021) to our context.

tax collectors, which prevents any subgroup formation. Second, due to the large number of cadets per squadron, Carrell et al. (2013) cannot observe all possible squadron compositions under random assignment. They identify squadrons with only low and high types (no medium types) as 'optimal' by *extrapolating* outside of the experiment's support, an approach that runs the risk of being invalid if the functional form they rely on is incorrect. As an improvement, Booij et al. (2016) propose to cover the range of possible assignments and to *interpolate* between observed assignments. In our context we need to neither extrapolate nor interpolate since we directly observe tax compliance for all possible combinations of types (Figure 1), including (*L,L,l*), (*L,L,h*), and (*H,H,h*), which characterize the optimal assignment (Figure 2). Therefore, we can be more confident in our estimate of the tax compliance that would be achieved under the optimal assignment. Finally, Carrell et al. (2013) are limited in their ability to document potential endogenous responses to implementing the optimal assignment because they only observe a single random assignment of each cadet into one squadron where they remain for multiple years. By contrast, we observe collectors re-randomized into new teams every month. This monthly re-assignment allows us to provide direct evidence about whether being assigned to certain types of teammates or households in the past affects future collector behavior. Specifically, in our context, changes in the assignment could affect collectors' effort or opportunities for learning by match type, resulting in changes in the average tax compliance function. We consider these issues in depth in Section A10 and summarize our findings here.

Endogenous Effort Responses. A first concern is that changing collectors' assignments could impact effort levels by match type, which would affect the average tax compliance function. For our analysis, one worrying scenario would be if collectors target high-type households for tax visits and are time-constrained, i.e., unable to do all the tax visits that have a positive return during the month-long campaign period. Under these conditions, implementing the optimal assignment could lead to lower visit levels and lower tax compliance for (H, H, h) match types than observed under random assignment.[88] However, while we find some evidence that collectors target visits to high-type households, we don't find evidence that the time constraints imposed by the tax campaign limit collectors' ability to do all the visits that have a positive return (Figure A23). Endogenous effort responses of this form are therefore unlikely to be a concern in our context.

Another potentially concerning scenario for our analysis is that low-type collectors could

[88]High-type collectors are assigned to high- and low-type households under the status quo (random) assignment and to high-type households only under the optimal (complete assortative matching) assignment. As a result, if time-constrained, high-type collectors would visit a lower fraction of high-type households under the optimal assignment than under the random assignment. Tax compliance would then be lower for (H, H, h) match types under the optimal assignment than the status quo assignment.

become demoralized under the optimal assignment if they realize that they will only work with low-type teammates and only be assigned low-type households.[89] Implementing the optimal assignment could then lead to lower compliance for (L, L, l) match types than observed under random assignment. We provide evidence by exploring whether low-type collectors assigned by chance to a higher share of low-type teammates and households during the 2018 campaign appear more demoralized at endline according to standard motivation questions from the psychology literature. Low-type collectors assigned to a larger fraction of low-type teammates or of low-type households do not exhibit systematically higher endline demoralization levels (Table A18–A19). According to this evidence, the assignment of low-type collectors to low-type teammates and households under the optimal assignment is unlikely to undermine their motivation and affect the average tax compliance function.[90]

Endogenous Learning Responses. A second concern is that changing collectors' assignment could impact learning by match type, which would affect the average tax compliance function. In particular, collectors might learn from high-type teammates (who might share targeting or persuasion techniques, for example).[91] If low-type collectors learn more than high-type collectors from being paired with a high-type teammate, positive assortative matching would suppress such learning and could lead us to overestimate the impact of the optimal assignment. If, by contrast, high-type collectors learn more from working with a high-type teammate, we might underestimate the impact of the optimal policy — because the optimal assignment would afford more opportunities for such learning. Ultimately, collectors do appear to learn from high-type teammates: past assignment to a high-type teammate has a positive effects on collectors' subsequent performance (Table A21, Column 1–3). However, learning from high-type teammates appears to be more pronounced among high-type collectors. This is consistent with our results *under*estimating the true impact of the optimal assignment (Table A21, Columns 4–5). That said, although the coefficients are large, this

[89]In fact, it is not obvious that the nature of (positive assortative) matching in the optimal assignment would even be salient to collectors, if implemented in the way described below (cf. 'Endogenous Incentive Effects' subsection).

[90]Section A10.1.2 also considers a more extreme case of demoralization: low-type collectors dropping out of the tax campaign entirely. In practice, only three tax collectors (8.82%) did not complete the full 2018 property tax campaign, and their decision to drop out does not appear to have been affected by the fraction of low-type teammates or households they were assigned to (Table A20). Moreover, implementing the optimal assignment would still have a positive and significant impact on tax compliance, even for non-trivial (up to 50%) dropout rates among low-type collectors (Figure A24).

[91]Endogenous learning-by-doing could also affect the average tax compliance function if collectors learn tax collection skills over time and if learning-by-doing varies by match type. However, analyzing exogenous variation in collectors' number of past assignments (and thus opportunities to gain tax collection experience), we find little evidence of learning-by-doing in this context (Table A22).

analysis is underpowered. The most we can confidently infer from this analysis is that our estimates are unlikely to overestimate the impact of the optimal policy due to learning from teammates.

Endogenous Incentive Effects. Finally, we assume that the government does not repeatedly re-evaluate collector performance, re-define types, and re-optimize the assignment, which would be administratively costly and would introduce incentive effects: e.g., collectors could exert higher effort in evaluation years if they know it could impact their future assignments (Khan et al., 2019). A forward-looking government would prefer to shut down these incentive effects by estimating collector types once (without announcing the evaluation) and implementing the optimal assignment in the subsequent years. Moreover, when implementing the optimal policy, the government need not inform collectors about their type or the type of households to which they are assigned. While collectors may be able to tell the difference between low and high types at the extremes of the collector and household type distributions, it is unlikely that they could easily differentiate between the large majority of collectors and households in the middle of those distributions. It is thus not obvious that the positive assortative matching in the optimal assignment would even be salient to collectors.[92]

8.4 Impacts on Bribes and Tax Morale

The optimal policy maximizes tax compliance, but teams of high-type collectors might be more likely to accept bribes as well as taxes,[93] or they might undermine tax morale if they achieve compliance through coercion. We explore the potential costs of implementing the optimal assignment policy in detail in Section A11 and summarize our findings here.

Bribe Payments. In Kananga's door-to-door tax collection system, collectors have discretion over key margins of tax administration and enforcement — the number and timing of tax visits, enforcement intensity, etc. — which could lead to collusive bribery: i.e., households making a smaller payment to collectors directly instead of paying the entire tax liability to the state.[94] As noted in Section 3, the government's choice of randomly assigning tax collectors to teammates and neighborhoods was in part motivated by a desire to minimize collectors' ability to develop collusive relationships with other collectors or with households. Notably, the random assignment of tax collectors to teammates and households would continue to occur under the optimal assignment: high-type (low-type) collectors would be ran-

[92]Note in particular that low-type collectors are still assigned to a non-trivial share of high-type households in the optimal assignment (Figure 2).

[93]Historical accounts in the United States, United Kingdom, and China show that initial expansions of state capacity were associated with increased bribery (Daunton, 2001; Carpenter, 2020; Cui, 2022).

[94]Collusion in property taxation exists in many settings (e.g., Khan et al., 2016).

domly matched with a high-type (low-type) teammate and randomly assigned to work in two neighborhoods every campaign month. Thus, the optimal assignment would preserve any collusion-prevention benefit from random reshuffling.

That said, it remains possible that the propensity to collect bribes varied by collector and household type and that implementing the optimal policy would impact bribery levels. We test this possibility using several measures of bribes described in Section A11. Using measures of extensive margin bribe payment, we find mixed results: a non-significant 0.387 percentage point increase ($p = 0.268$, Panel A of Table 2, Row 1) according to our preferred measure and a marginally significant 2.253 percentage point increase ($p = 0.059$, Panel A of Table 2, Row 5) according to a measure that also captures social desirability. On the intensive margin (i.e., the average amount of bribe paid), we find a marginally significant increase of 13.896 CF ($p = 0.098$, Panel A of Table 2, Row 3).

There is thus suggestive evidence that implementing the optimal assignment could increase bribe payments. The most likely explanation is that *H*-*H* teams conducted more tax visits (Figure A8) which gave them more opportunities to collect bribes. Indeed, we find no evidence that the optimal assignment would increase bribe frequency or amount *per visit* (Panel A of Table 2, Row 2 and 4). Thus, we do not view the suggestive evidence of more bribes as an adverse outcome but as a mechanical reflection of the higher effort exerted by *H*-*H* collector teams. As long as the government tolerates the current rate of bribe payment per visit (1.783%), which it most likely does since it continues field-based property tax collection, then it should prefer to implement the optimal assignment.

Tax Morale. Implementing the optimal assignment could also backfire if it erodes tax morale and thus reduces compliance with other formal or informal taxes. We investigate this possibility using survey data on self-reported contributions to a range of taxes as well as views of the government and of taxation more generally. According to our measures, there is no evidence that the optimal assignment would crowd out payments of other formal or informal taxes (Table 2, Panels B-C). Similarly, the optimal assignment is unlikely to affect views of government (Table 2, Panel D). Suggestively, it may have mixed effects on citizens' view of taxation (Table 2, Panel E), slightly increasing citizen trust in the tax ministry ($p = 0.100$), while slightly reducing the perceived share of tax revenue spent on public goods ($p = 0.106$, respectively). Yet, we find no significant impact of the optimal assignment on property tax morale ($p = 0.491$). In sum, there is little evidence of eroding views of the government or of taxation that might give the government pause in choosing the optimal assignment policy.

8.5 Distributional Impacts

The optimal assignment policy increases tax compliance and revenue *on average*, but does it shift the de facto incidence of the property tax? To investigate the distributional implications of the optimal assignment, we compare the characteristics of taxpayers under the optimal and status quo assignments. Formally, we estimate:

$$\mathbb{E}_f[X_h|Y_h = 1] \tag{12}$$

where X_h denotes household h's characteristics, Y_h is a dummy indicating whether h paid the property taxes, and the subscript f indicates that the expectation is taken with respect to assignment function f. We compare $\mathbb{E}_f[X_h|Y_h = 1]$ with $f = f^*$, the optimal assignment function, and with $f = f^{SQ}$, the status quo assignment function. Section A12 describes the estimation of $\mathbb{E}_f[X_h|Y_h = 1]$.

Using this approach, we find that the taxpayer population includes more high-type households under the optimal assignment — 91% of all payers — relative to the status quo assignment — 83%, a significant difference ($p < 0.001$) (Table 3, Panel A). Because high-type households are themselves wealthier, more likely to be employed or salaried, and more highly educated (Table A3, Panels A–C), we would expect the optimal assignment to shift distribution of the tax burden toward wealthier households. Our estimation bears out this prediction. Taxpayers under the optimal assignment policy would have higher quality house walls ($p = 0.001$), roofs ($p = 0.014$), and overall more valuable properties ($p = 0.084$) compared to the status quo assignment (Table 3, Panel B). They also have higher job security, more education, and higher incomes, though these differences are not statistically significant (Table 3, Panel C).

9 Comparison with Selection Policies and Wage Increases

We now compare the effect of implementing the counterfactual optimal assignment on compliance and revenue with the impact of alternative policies such as collector selection policies and collector wage increases.

9.1 Effects of Selection Policies

To benchmark the effect of the optimal assignment described in Section 8, we first turn to estimating the increase in tax compliance and revenue associated with two types of counterfactual collector selection policies: (*i*) reallocation policies, which involve reallocating a fraction ρ of low-type collector assignments to high-type collectors, and (*ii*) hiring policies, which involve reassigning them to newly hired collectors of average ability instead.[95]

Figure 3 explores the effect of selection policies on tax compliance relative to the status quo assignment when a fraction ρ of the low-type collector assignments are reallocated to high-type collectors (reallocation policies) or to newly hired collectors (hiring policies). Reallocation policies would surpass the optimal assignment only for large values of ρ: the provincial tax ministry would have to reassign at least 63% of low-type collectors' assignments to high-type collectors to achieve the same increase in compliance as under the optimal assignment policy (Panel A).[96] Hiring policies, by contrast, would never rival the optimal assignment (Panel B). At most, the government could increase tax compliance by 2.237 percentage points, which is 0.704 percentage points less than the effect of the optimal assignment, if it were to reallocate all low-type collectors' households to newly hired collectors.[97] We view these estimates of the effects of selection policies as upper bounds given that they assume away other costs, such as the tax on high-type collectors from a larger workload and the search and training costs of hiring new collectors.[98]

9.2 Effects of Collector Financial Incentives

As a second benchmark, we consider another potential intervention: performance-based financial incentives.[99] To assess if the government could increase tax compliance by raising collectors' financial incentives, we exploit random variation in collectors' piece-rate wages between a constant amount — 750 CF per collection — and a proportional amount — 25% of the amount collected — during the 2018 property tax campaign, as described in Section 2.[100] Using this variation, we estimate that the government would have to increase collectors' piece-rate wages by 69% to achieve the same compliance increase as the optimal assignment (Figure 4, Panel A).

[95] Section A13 defines the reallocation and hiring policies using the notation from Section 5. Similar selection policy counterfactuals have been analyzed in the literature on teacher quality (Chetty et al., 2014) and public sector manager quality (Fenizia, 2022).

[96] At most, the government could increase tax compliance by 5.112 percentage points if it were to reassign all the low-type collectors' assignment to high-type collectors.

[97] We find similar results when relying on the predicted collector type from survey characteristics introduced in Section 8.2 (Figure A18).

[98] These costs are unlikely to be large for small values of ρ, since collectors do not appear to be time constrained under the status quo assignment (Figure A23), but they might be important when ρ is large.

[99] Performance-based incentives for tax collectors are used in a number of developing countries, including Brazil and Pakistan. In Pakistan, Khan et al. (2016) find that performance-based incentives for property tax collectors increased tax revenue by 9%.

[100] The piece-rate wage and tax rate associated with each property were written on the property register used by the tax collectors. The piece-rate wage randomization is explored in detail in Bergeron et al. (2021).

While the size of this wage increase might be enough to give the government pause in contemplating this policy, a further consideration is its cost-effectiveness. Specifically, paying collectors a larger share of the tax revenue they collect will only raise revenue if the compliance response to stronger performance incentives is greater than the mechanical increase in the wage bill. We therefore re-run the analysis by examining effects on revenue *net of collector wages* (Figure 4, Panel B). In fact, increasing wages by 69% would result in a 6% decline in net tax revenues. The cost-ineffectiveness of this policy highlights a crucial advantage of the optimal assignment: its cost neutrality. Given the tightness of budget constraints facing governments in low-income countries, the optimal assignment policy is a particularly attractive tool for raising fiscal capacity because it would do so within existing constraints on human and financial resources.[101]

10 Conclusion

This paper explored the role of bureaucrat assignment in government effectiveness in a low-income country with a weak state. Exploiting random assignment of tax collectors to teams and neighborhoods, we found that pairing effective collectors together, as well as assigning effective collector teams to households or neighborhoods with higher payment propensity, would substantially increase tax compliance.[102] According to our counterfactual estimates, implementing the optimal assignment policy would outperform alternative policies such as reallocating collection duties to more effective collectors or increasing the performance-based wages paid to collectors. Ultimately, the optimal assignment of tax collectors to teams and teams to households and neighbourhoods appears to be a promising way for governments to increase tax revenue without increasing the costs of tax administration.

These results build on recent theory (Keen and Slemrod, 2017) and evidence (Khan et al., 2016, 2019; Basri et al., 2021) that improving the efficiency of tax administration is paramount in low-income countries. While much of the literature on the public finance of developing countries focuses on investing in *enforcement* capacity (e.g., Besley and Persson,

[101] We can also compare the effect of the optimal assignment policy with another standard intervention frequently used to stimulate tax compliance: enforcement nudges on tax notices. We leverage the random assignment of enforcement messages on tax notices distributed by collectors during the 2018 property tax campaign, as described in Bergeron et al. (2021). Enforcement messages increased tax compliance by 1.4 percentage points relative to a placebo message (Table A12), which is in line with the effects of enforcement messages found in other settings (e.g., Blumenthal et al., 2001; Fellner et al., 2013; Pomeranz, 2015; Scartascini and Castro, 2007). This is less than half the effect size of the optimal assignment policy.

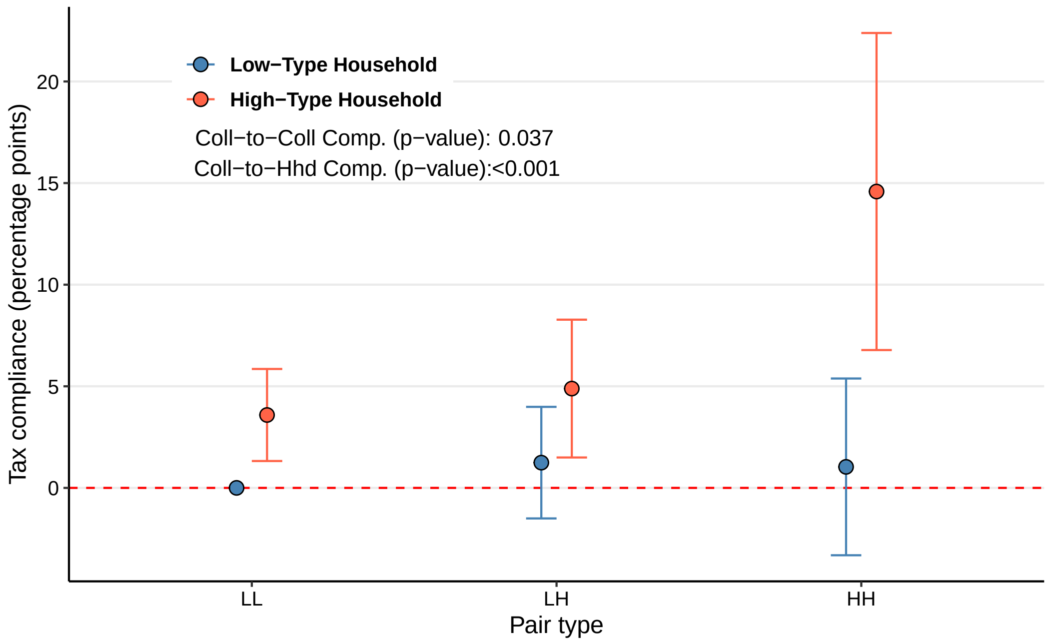
[102] The optimality of positive assortative matching is likely to be specific to the context and objective we study: a low-income country aiming at maximizing revenue from field-based tax collection. In other domains, such as education, the government might have a different objective, such as achieving a baseline level of instruction everywhere, which could justify spreading out high-quality teachers.

2009; Pomeranz, 2015; Naritomi, 2019), which is surely necessary if countries seek to collect 30-40% of their GDPs in tax, there has been perhaps less focus on tax administration as a complementary priority in tax policy, with the exception of Keen and Slemrod (2017), Khan et al. (2016, 2019), and Basri et al. (2021). Particularly in low-income countries with weak states, such as the DRC, raising the efficiency of tax administration is essential if tax authorities are to make the most of enforcement tools like audits and third-party reporting. As Casanegra de Jantscher (1990) put it, "in developing countries, tax administration *is* tax policy."

One natural question is whether tax authorities in low-capacity settings would implement the optimal assignment or would be prevented from doing so by political considerations. For instance, if low-type collectors have powerful patrons, they might lobby in favor of mixed teams, which allow them to free-ride on their productive peers and take home higher revenues. We view understanding how tax authorities respond to information about the returns to implementing the optimal assignment, as well as the role of political constraints in sustaining more idiosyncratic assignments, as fertile ground for future research.

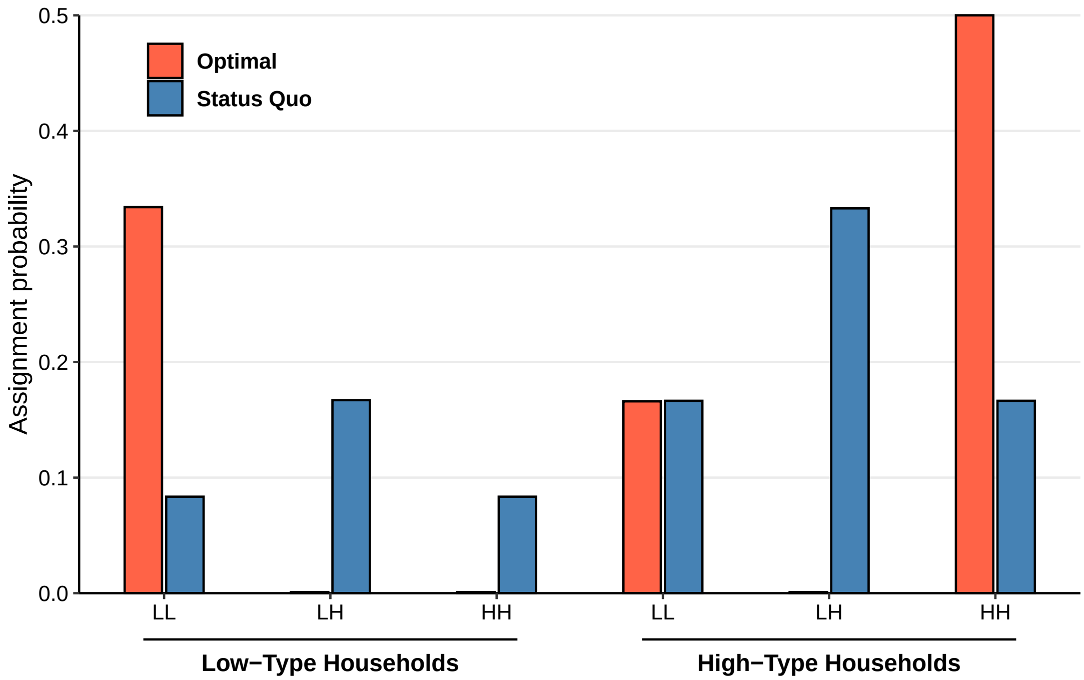
11 Tables and Figures

Figure 1: Tax Compliance By Collector and Household Types



Notes: This figure shows the estimates of average tax compliance when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by household type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, and HH. The y-axis captures the tax compliance probability for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax compliance as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. The figure also reports p-values associated with tests for the average tax compliance function exhibiting increasing differences in collector type and in collector and household type. We report the p-value associated with a test that tax compliance, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.1.

Figure 2: Optimal Assignment Vs. Status Quo Assignment



Notes: This figure shows the counterfactual optimal assignment and the status quo assignment functions. Each bar represents the probability of each match type under the counterfactual optimal assignment (red) and status quo (blue) assignment functions. The first 6 bars show the assignment functions with matches involving low-type households. The 6 subsequent bars show the assignment functions with matches involving high-type households. We discuss these results in Section 7.1.

Table 1: Effects of the Optimal Assignment on Tax Compliance and Revenues

	Household Types: Household Propensity to Pay			
	Collector Types: Fixed Effects Model		Collector Types: Coll. Chars. Model	
	Tax Compliance (in percentage points) (1)	Tax Revenue (in Congolese Francs) (2)	Tax Compliance (in percentage points) (3)	Tax Revenue (in Congolese Francs) (4)
Optimal Assignment	2.941**	54.471*	2.688**	56.926**
	(1.239)	(30.52)	(1.237)	(28.725)
	[0.024]	[0.074]	[0.030]	[0.048]
Collector-to-Collector Only	1.294	21.444	1.097	27.985
	(0.947)	(21.675)	(0.937)	(21.540)
	[0.172]	[0.322]	[0.242]	[0.194]
Collector-to-Household Only	0.837***	17.156**	0.875**	13.371
	(0.312)	(8.520)	(0.369)	(9.232)
	[0.007]	[0.044]	[0.018]	[0.147]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,732	11,732
Observations (Analysis Sample)	6,904	6,904	6,904	6,904

Notes: This table reports estimates of the impact of the counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1 and 3 show estimates for the probability of compliance, i.e., a dummy indicating whether households paid the property taxes (multiplied by 100). The point estimates should be interpreted as percentage point changes. Columns 2 and 4 show estimates for average tax revenue per household in Congolese Francs. Columns 1–2 present results when collector types are estimated using a fixed effects model as described in Section 6.2. Columns 3–4 show results when collector types are estimated from tax collectors' characteristics as described in Section 8.2. Each row present counterfactual results for a different optimal assignment. The first row presents results when optimizing on both the collector-to-household and the collector-to-collector dimension. The second and third rows show results when only optimizing the collector-to-collector and collector-to-household dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax compliance (Columns 1 and 3) and tax revenue (Columns 2 and 4) is reported at the bottom of the table. The final two rows report the size of the holdout and analysis sample. We discuss these results in Sections 8.1 and 8.2.

Table 2: Effects of the Optimal Assignment on Other Outcomes

Dependent variable	ARE (1)	SE (2)	p-value (3)	Mean (4)	Observations (Holdout) (5)	Observations (Analysis) (6)	Sample (Analysis) (7)
Panel A: Bribes							
Paid Bribe	0.387	0.349	0.268	1.718	11,732	4,691	Midline
Paid Bribe Per Visit	-0.037	0.432	0.932	1.783	11,732	1,801	Midline
Bribe Amount	13.896*	8.408	0.098	30.431	11,732	4,691	Midline
Bribe Amount Per Visit	7.583	5.963	0.204	27.206	11,732	1,801	Midline
Gap Self v. Admin	2.253*	1.193	0.059	9.529	11,732	3,543	Midline
Panel B: Informal Labor Taxes							
Salongo	3.890	2.522	0.123	37.495	11,732	3,429	Midline
Salongo Hours	0.187	0.180	0.299	1.601	11,732	3,317	Midline
Panel C: Other Formal Taxes							
Vehicle Tax	-0.144	0.939	0.878	3.138	11,732	541	Endline
Market Vendor Fee	-2.507	2.858	0.380	17.165	11,732	541	Endline
Business Tax	0.772	1.666	0.643	5.492	11,732	541	Endline
Income Tax	-1.710	1.710	0.317	10.635	11,732	538	Endline
Obsolete Tax	0.884	0.780	0.257	1.650	11,732	538	Endline
Panel D: View of Government							
Trust in Government	0.178	0.110	0.106	1.737	11,732	268	Endline
Responsiveness of Government	0.071	0.070	0.315	0.003	11,732	538	Endline
Performance of Government	-0.043	0.062	0.483	0.006	11,732	531	Endline
Panel E: View of Taxation							
Trust in Tax Ministry	0.105*	0.064	0.100	1.685	11,732	270	Endline
Property Tax Morale	0.052	0.075	0.491	-0.036	11,732	540	Endline
Perception of Enforcement	-2.820	2.270	0.214	48.562	11,732	4,074	Midline
Perception of Public Goods Provision	-6.076	3.764	0.106	43.412	11,732	3,733	Midline

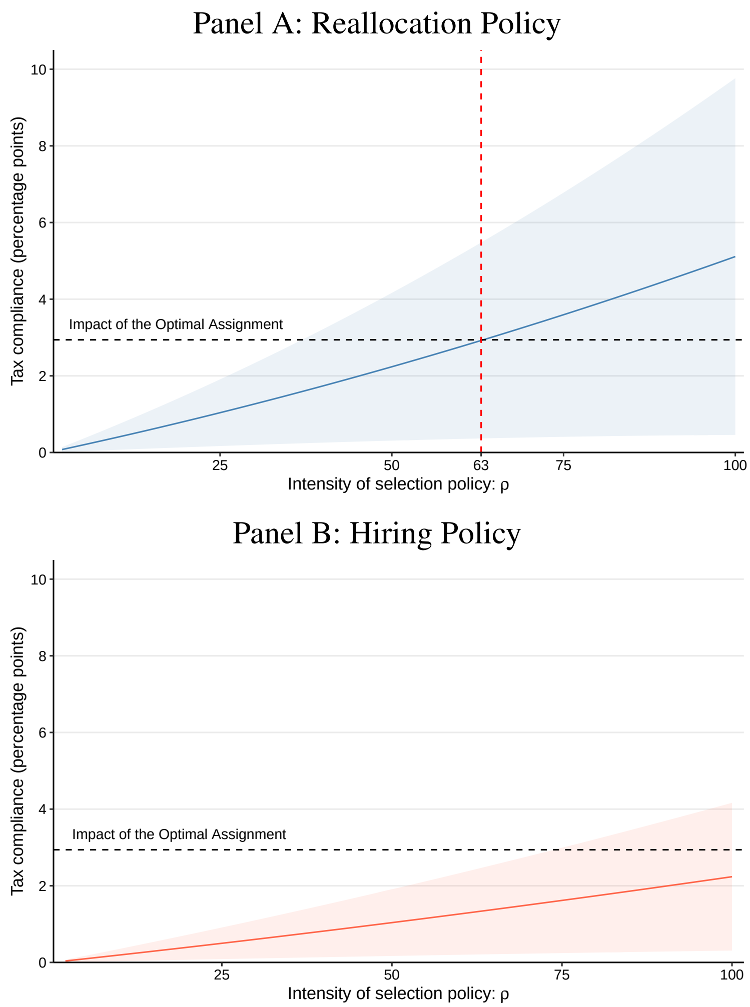
Notes: This table shows the impact of the counterfactual optimal assignment policy on secondary outcomes. In Panel A, the outcomes in rows 1 and 2 are self-reported bribe payment and bribe payment per visit as measured during the midline survey. The outcomes in rows 3 and 4 are self-reported bribe amounts and bribe amounts per visit, as measured during the midline survey. Finally, the outcome in row 5 indicates property owners who reported paying the tax but who were not recorded as having paid in the administrative data. In Panel B, rows 4 and 5 report *salongo* contributions along the extensive and intensive margins of hours, respectively, at midline. In Panel C, rows 6–10 report self-reported payment of other formal taxes at endline. The obsolete tax is a poll tax, which existed in the past but does not currently exist, to test the reliability of self-reports. In Panel D, the outcomes in rows 11–13 are self-reported views of the government: trust, responsiveness, and performance of the government. In Panel E, rows 14–17, we consider self-reported views of taxation: trust in the tax ministry, tax morale, perception of enforcement, and perception that tax revenues are spent on public goods. The ARE estimator for each outcome is shown in Column 1. Standard errors are clustered at the neighborhood level and presented in Column 2 while the corresponding p-values are presented in Column 3 ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average of the outcome variables is shown in Column 4. The number of observations in the holdout sample and the analysis sample are presented in Columns 5 and 6, respectively. The definition of the holdout sample (midline or endline) is given in Column 7. We discuss these results in Section 8.4.

Table 3: Incidence of the Optimal Assignment

	Average Taxpayers Optimal Assignment (1)	Average Taxpayers Random Assignment (2)	Average All (3)	Difference (1) vs. (2) (4)	p-value (5)	Observations Taxpayers (6)	Observations All (7)	Sample (8)
Panel A: Household Type								
High-type Household	0.905	0.826	0.666	0.078***	<0.001	577	6904	Registration
Panel B: Property Characteristics								
Roof Quality	7.000	6.937	6.901	0.063**	0.014	1,296	16,010	Midline
Walls Quality	1.748	1.618	1.497	0.130***	0.001	1,302	16,139	Midline
Fence Quality	1.346	1.380	1.374	-0.034	0.225	1,159	14,862	Midline
Property Value	1689.245	1495.220	1325.137	194.025*	0.084	1,567	19,587	Registration
Panel C: Property Owner Characteristics								
Male Owner Indicator	0.817	0.825	0.800	-0.008	0.739	748	9,400	Midline
Main Tribe Indicator	0.757	0.780	0.802	-0.023	0.343	911	9,555	Midline
Employed Indicator	0.799	0.815	0.802	-0.016	0.452	956	10,302	Midline
Salaried Indicator	0.322	0.311	0.259	0.011	0.691	956	10,302	Midline
Work for Government Indicator	0.194	0.176	0.167	0.018	0.411	956	10,302	Midline
Relative Work for Government Indicator	0.283	0.272	0.245	0.012	0.622	1,056	11,456	Midline
Years of Education	11.122	10.782	10.533	0.341	0.459	185	1,533	Endline
Log Monthly Income	11.012	10.731	10.563	0.281	0.223	185	1,525	Endline

Notes: This table shows the average characteristics of taxpayers under different assignment policies. Columns 1 and 2 show the average for taxpayers under the optimal and the status quo assignments, respectively. Column 3 shows the average for the entire sample of registered properties. Column 4 shows the difference in average characteristics of taxpayers under the optimal and status quo assignment. Column 5 shows the p-value associated with the test that the estimate in column 4 is different than zero ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). Columns 6 and 7 report the number of observations corresponding to each characteristics when focusing on taxpayers (Column 6) and for all observations (Column 7). The analysis sample is listed in Column 8. Panel A considers the household type indicator. Panel B focuses on characteristics of the property measured at midline and the predicted property value estimated using machine learning (Bergeron et al., 2022). Panel C analyzes characteristics of the property owner measured at midline and endline. The variables are described in detail in Section A14. We discuss these results in Section 8.5.

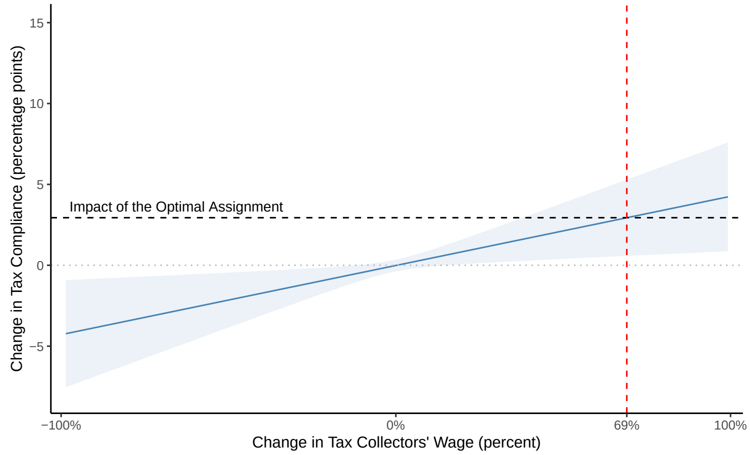
Figure 3: Effects of Selection Policies



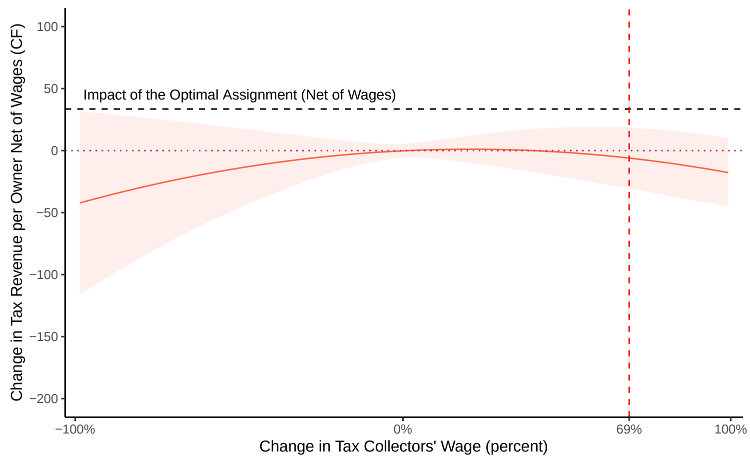
Notes: This figure shows the impact of the selection policies on the probability of tax compliance (y-axis). Selection policies involve reassigning $\rho\%$ (x-axis) of the assignments that a low-ability collector would receive under the status quo assignment to other collectors. Panel A shows the estimated effects of the *reallocation policy*, where the workload is reassigned to existing high-ability collectors in the sample. Panel B shows the estimated effects of the *hiring policy*, where the workload is reassigned to newly hired collectors with types drawn uniformly from {L, H}. In both Panels, the collector types are estimated using a fixed effects model as described in Section 6.2. The shaded areas represent 95% confidence intervals. The dashed horizontal line indicates the counterfactual impact of the optimal assignment policy on tax compliance when collector types are estimated using a fixed effects model, as reported in Column 1 of Table 1. We discuss these results in Section 9.1.

Figure 4: Effects of Wage Increases

Panel A: Effects on Tax Compliance



Panel B: Effects on Tax Revenue Net of Collectors' Wage



Notes: This figure shows the impact of increases in tax collectors' wage on tax compliance (Panel A) and tax revenue net of wages (Panel B). The x-axis shows changes in tax collectors' wage relative to the status quo wage (in percentage). The y-axis in Panel A is the predicted tax compliance for each collectors' wage. It is estimated using the OLS regression of tax compliance on collectors' wage, as shown in Column 1 of Table A11. The y-axis in Panel B is the predicted tax revenue net of collectors' wage by collectors' wage level. It mechanically derives from the predicted tax compliance in Panel A, tax rates, and collectors' wage. In Panel A, the dashed horizontal black line indicates the counterfactual impact of the optimal assignment policy on tax compliance as reported in Column 1 of Table 1. In Panel B, the dashed horizontal black line indicates the counterfactual impact of the optimal assignment policy on tax revenue net of tax collectors' wage. We obtain it by subtracting the predicted increase in collectors' wage associated with the counterfactual optimal assignment policy from the effect on tax revenue reported in Column 2 of Table 1. The shaded areas represent 95% confidence intervals using standard errors bootstrapped (with 1,000 iterations). We discuss these results in Section 9.

References

Abadie, Alberto and Matias D. Cattaneo, "Econometric Methods for Program Evaluation," *Annual Review of Economics*, 2018, *10*, 465–503.

Alchian, Armen A. and Harold Demsetz, "Production, Information Costs, and Economic Organization," *American Economic Review*, 1972, *62* (5), 777–795.

Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey, "Inference on Winners," *Working Paper*, 2021.

Ashraf, Nava and Oriana Bandiera, "Social Incentives in Organizations," *Annual Review of Economics*, 2018, *10*, 439–463.

_ , _ , and Scott S Lee, "Awards Unbundled: Evidence from a Natural Field Experiment," *Journal of Economic Behavior & Organization*, 2014, *100*, 44–63.

_ , _ , Edward Davenport, and Scott S. Lee, "Losing Prosociality in the Quest for Talent? Sorting, Selection and Productivity in the Delivery of Public Service," *American Economic Review*, 2020, *110* (5).

Aucejo, Esteban M., Patrick Coate, Jane Cooley Fruehwirth, Sean Kelly, and Zachary Mozenter, "Teacher Effectiveness and Classroom Composition: Understanding Match Effects in the Classroom," *Working Paper*, 2019.

Balan, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan Weigel, "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the DR Congo," *American Economic Review*, 2022, *112* (3), 762–797.

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

_ , Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat, "The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats," *The Quarterly Journal of Economics*, 2021, *136* (4), 2195–2242.

Basri, M. Chatib, Mayara Felix, Rema Hanna, and Benjamin A. Olken, "Tax Administration vs. Tax Rates: Evidence from Corporate Taxation in Indonesia," *American Economic Review*, 2021, *111* (12), 3827–3871.

Becker, Gary S., "A Theory of Marriage: Part I," *Journal of Political Economy*, 1973, *81* (4), 813–846.

Bergeron, Augustin, Arnaud Fournier, Gabriel Tourek, and Jonathan Weigel, "Using Machine Learning to Improve Property Tax Collection in the DRC," *Working Paper*, 2022.

_ , Gabriel Tourek, and Jonathan Weigel, "The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC," *Working Paper*, 2021.

Bertrand, Marianne, Robin Burgess, Arunish Chawla, and Guo Xu, "The glittering prizes: Career incentives and bureaucrat performance," *The Review of Economic Studies*, 2020, *87* (2), 626–655.

Besley, Timothy and Torsten Persson, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 2009, *99* (4), 1218–44.

_ and _ , "Taxation and Development," *Handbook of Public Economics*, 2013, *5*, 51–110.

Bessone, Pedro, "Can Matching Algorithms Hurt Female Workers' Prospects? Evidence from the Financial Industry in Brazil," *Working Paper*, 2020.

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem, "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan," *Journal of political Economy*, 2015, *123* (6), 1311–1355.

_ , Jonas Hjort, and David Szakonyi, "Individuals and Organizations as Sources of State Effectiveness," *Working Paper*, 2019.

Bhattacharya, Debopam, "Inferring optimal peer assignment from experimental data," *Journal of the American Statistical Association*, 2009, *104* (486), 486–500.

Blumenthal, Marsha, Charles Christian, Joel Slemrod, and Matthew G Smith, "Do Normative Appeals Affect Tax Compliance? Evidence From a Controlled Experiment in Minnesota," *National Tax Journal*, 2001, pp. 125–138.

Bó, Ernesto Dal, Frederico Finan, and Martín A Rossi, "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

_, _, Nicholas Y Li, and Laura Schechter, "Government Decentralization under Changing State Capacity: Experimental Evidence from Paraguay," *Econometrica*, 2020.

Bonhomme, Stéphane, "Teams: Heterogeneity, sorting, and complementarity," *Working Paper*, 2021.

Booij, Adam S., Edwin Leuven, and Hessel Oosterbeek, "Ability Peer Effects in University: Evidence from a Randomized Experiment," *Review of Economic Studies*, 2016, *84* (2), 547–578.

Brewer, John, *The Sinews of Power: War, Money, and the English State, 1688-1783*, Harvard University Press, 1990.

Brockmeyer, Anne, Alejandro Estefan, Juan Carlos Suȑez Serrato, and Karina Ramirez, "Taxing Property in Developing Countries: Theory and Evidence from Mexico," *Working Paper*, 2021.

Brune, Lasse, Eric Chyn, and Jason Kerwin, "Peers and Motivation at Work: Evidence from a Firm Experiment in Malawi," *Journal of Human Resources*, 2020.

Burgess, Simon, Carol Propper, Marisa Ratto, Stephanie von Hinke Kessler Scholder, and Emma Tominey, "Smarter Task Assignment or Greater Effort: the Impact of Incentives on Team Performance," *The Economic Journal*, 2010, *120* (547), 968–989.

Callen, Michael, Saad Gulzar, Ali Hasanain, Yasir Khan, and Arman Rezaee, "Personalities and Public Employee Performance: Experimental Evidence from Pakistan," *Working Paper*, 2018.

Carpenter, Daniel, *The Forging of Bureaucratic Autonomy*, Princeton University Press, 2020.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West, "From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation," *Econometrica*, 2013, *81* (3), 855–882.

_, Richard L. Fullerton, and James E. West, "Does Your Cohort Matter? Measuring Peer Effects in College Achievement," *Journal of Labor Economics*, 2009, *27* (3), 439–464.

Carrillo, Paul, Dina Pomeranz, and Monica Singhal, "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement," *American Economic Journal: Applied Economics*, 2017, *9* (2), 144–164.

Casanegra de Jantscher, Milka, "Administering the VAT," in "Value Added Taxation in Developing Countries," World Bank, 1990, pp. 171–190.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 2014, *104* (9), 2633–79.

Chu, Jian, Raymond Fisman, Songtao Tan, and Yongxiang Wang, "Hometown Ties and the Quality of Government Monitoring: Evidence from Rotation of Chinese Auditors," *Working Paper*, 2020.

Cogneau, Denis, Marc Gurgand, Justine Knebelmann, Victor Pouliquen, and Bassirou Sarr, "Bringing Property Owners into the Tax Net: Avenues of Fiscal Capacity and Local Accountability: Evidence from Dakar, Senegal," *Working Paper*, 2020.

Crawford, Vincent P. and Elsie Marie Knoer, "Job Matching with Heterogeneous Firms and Workers," *Econometrica*, 1981, pp. 437–450.

Cui, Wei, *The Administrative Foundations of the Chinese Fiscal State*, Cambridge University Press, 2022.

Dahis, Ricardo, Laura Schiavon, and Thiago Scot, "Selecting Top Bureaucrats: Admission Exams and Performance in Brazil," *Working Paper*, 2021.

Dantzig, George B., "Linear Programming," in "Problems for the Numerical Analysis of the Future," Institute for Numerical Analysis, U.S. Government Printing Office, 1951.

Daunton, Martin, *Trusting Leviathan: The Politics of Taxation in Britain 1799-1914*, Cambridge University Press, 2001.

Duflo, Esther, Rema Hanna, and Stephen P Ryan, "Incentives Work: Getting Teachers to Come to School," *American Economic Review*, 2012, *102* (4), 1241–78.

Efron, Bradley, *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia, Pa. :Society for Industrial and Applied Mathematics, 1982.

Fellner, Gerlinde, Rupert Sausgruber, and Christian Traxler, "Testing Enforcement Strategies in the Field: Threat, Moral Appeal and Social Information," *Journal of the European Economic Association*, 2013, *11* (3), 634–660.

Fenizia, Alessandra, "Managers and Productivity in the Public Sector," *Econometrica*, 2022.

Finan, Frederico, Benjamin A. Olken, and Rohini Pande, "The Personnel Economics of the Developing State," *Handbook of Economic Field Experiments*, 2017, *2*, 467–514.

Franzsen, Riel and William McCluskey, *Property Tax in Africa: Status, Challenge, and Prospects*, Lincoln Institute of Land Policy, 2017.

Graham, Bryan S., "Econometric Methods for the Analysis of Assignment Problems in the Presence of Complementarity and Social Spillovers," *Handbook of social economics*, 2011, *1*, 965–1052.

__ , Geert Ridder, Petra M. Thiemann, and Gema Zamarro, "Teacher-to-classroom Assignment and Student Achievement," *Working Paper*, 2020.

Graham, Bryan S, Guido W. Imbens, and Geert Ridder, "Complementarity and aggregate implications of assortative matching: A nonparametric analysis," *Quantitative Economics*, 2014, *5* (1), 29–66.

Graham, Bryan S., Guido W. Imbens, and Geert Ridder, "Identification and Efficiency Bounds for the Average Match Function under Conditionally Exogenous Matching," *Journal of Business and Economic Statistics*, 2020, *38* (2), 303–316.

Hanna, Rema and Shing-Yi Wang, "Dishonesty and Selection into Public Service: Evidence from India," *American Economic Journal: Economic Policy*, 2017, *9* (3), 262–90.

Henn, Soeren, "Complements or Substitutes? How Institutional Arrangements Bind Chiefs and the State in Africa," *Working Paper*, 2021.

Ichino, Andrea and Giovanni Maggi, "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm," *The Quarterly Journal of Economics*, 2000, *115* (3), 1057–1090.

Kane, Thomas J. and Douglas O. Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," *Working Paper*, 2008.

Keen, Michael and Joel Slemrod, "Optimal Tax Administration," *Journal of Public Economics*, 2017, *152*, 133–142.

Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken, "Tax farming redux: Experimental evidence on performance pay for tax collectors," *The Quarterly Journal of Economics*, 2016, *131* (1), 219–271.

__ , __ , and __ , "Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings," *American Economic Review*, 2019, *109* (1), 237–70.

Krause, Benjamin, "Taxation Toward Representation: Public Goods, Tax Collection, Social Norms, and Democratic Accountability," *Working Paper*, 2020.

Kremer, Michael, "The O-Ring Theory of Economic Development," *The Quarterly Journal of Economics*, 1993, *108* (3), 551–575.

Limodio, Nicola, "Bureaucrat Allocations in the Public Sector: Evidence from the World Bank," *The Economic Journal*, 2021, *131* (639), 3012–3040.

Martin, Lucy, Brigitte Seim, Luis Camacho, and Simon Hoellerbauer, "Marketing Taxation? Experimental Evidence on Enforcement and Bargaining in Malawian Markets," *Working Paper*, 2021.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and Team Performance in a Kenyan Organization," *Journal of Public Economics*, 2021, *197*, 104332.

Mas, Alexandre and Enrico Moretti, "Peers at Work," *American Economic Review*, 2009, *99* (1), 112–45.

Morris, Carl N, "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of the American statistical Association*, 1983, *78* (381), 47–55.

Mullainathan, Sendhil and Jann Spiess, "Machine Learning: An Applied Econometric Approach," *Journal of Economic Perspectives*, 2017, *31* (2), 87–106.

Naritomi, Joana, "Consumers as Tax Auditors," *American Economic Review*, 2019, *109* (9), 3031–3072.

OECD, "Revenue Statistics in Africa – Democratic Republic of the Congo," *Policy Report*, 2020.

Okunogbe, Oyebola, "State Capacity and Taxation: The Role of Detection and Enforcement Capacity on Property Tax Compliance in Liberia," *Working Paper*, 2021.

Olson, Mancur, "Collective action," in "The invisible hand," Springer, 1989, pp. 61–69.

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

Rasul, Imran and Daniel Rogger, "Management of Bureaucrats and Public Service Delivery: Evidence from the Nigerian Civil Service," *The Economic Journal*, 2018, *128* (608), 413–446.

Rotemberg, Julio J., "Human Relations in the Workplace," *Journal of Political Economy*, 1994, *102* (4), 684–717.

Rubin, Donald B., "The Bayesian Bootstrap," *Annals of Statistics*, 1981, *9* (1), 130–134.

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

Thakur, Ashutosh, "Matching in the Civil Service: A Market Design Approach to Public Administration and Development," *Working Paper*, 2020.

Topkis, Donald M., *Supermodularity and Complementarity*, Princeton University Press, 1998.

Weigel, Jonathan L, "The Participation Dividend of Taxation: How Citizens in Congo Engage More with the State When it Tries to Tax Them," *The Quarterly Journal of Economics*, 2020, *135* (4), 1849–1903.

Wong, Francis, "Mad as Hell: Property Taxes and Financial Distress," *Working Paper*, 2020.

Xu, Guo, "The Costs of Patronage: Evidence from the British Empire," *American Economic Review*, 2018, *108* (11), 3170–98.

Supplementary Data and Appendix
For Online Publication

A1 Additional Tables and Figures

Table A1: Components of the Tax Campaign and Its Evaluation

Activity	Actor	Timing	Observations	Neighborhoods
Tax Campaign				
Property register	Collectors	May-Dec 2018	19,600	180
Tax collection	Collectors	May-Dec 2018	19,600	180
Evaluation				
Baseline citizen survey	Enumerators	Jul-Dec 2017	1,404	180
Midline citizen survey	Enumerators	Jun 2018-Feb 2019	16,346	180
Baseline collector survey	Enumerators	April-May 2018	34	N/A

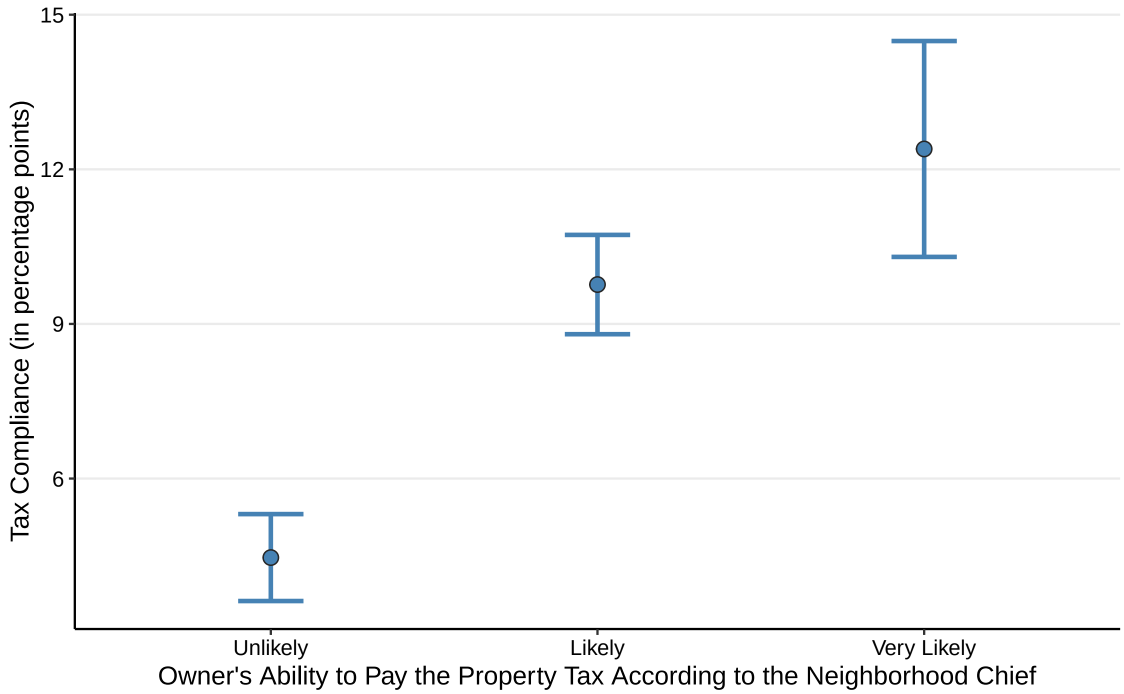
Notes: This table reports the actors, the timing, the number of observations (properties) and the number of clusters (neighborhoods) associated with each tax campaign activity. The property register has more observations per neighborhood than the midline survey because the former includes information on all compounds, including (exempt) government buildings, churches, and empty lots, while the midline survey was only conducted with privately owned plots liable for the property tax. The primary tax outcomes result from merging official property tax records with data from the property register. The mechanics of the tax campaign and data sources are discussed, respectively, in Sections 2 and 4.

Table A2: Balance

	Sample (1)	Observations (2)	Mean (L-L pairs) (3)	L-H pairs (4)	H-H pairs (5)
Panel A: Property Characteristics					
Distance to State Buildings (in km)	Registration	19,354	0.829	-0.009	0.165
				(0.107)	(0.125)
Distance to Health Institutions (in km)	Registration	19,354	0.349	0.014	-0.008
				(0.036)	(0.035)
Distance to Education Institutions (in km)	Registration	19,354	0.356	0.059*	-0.003
				(0.033)	(0.029)
Distance to Roads (in km)	Registration	18,849	0.442	-0.028	-0.058
				(0.061)	(0.066)
Distance to Eroded Areas (in km)	Registration	18,849	0.123	0.001	-0.019
				(0.015)	(0.018)
Walls Quality	Midline	16,131	1.123	0.054	0.024
				(0.036)	(0.038)
Roof Quality	Midline	16,346	0.976	-0.017**	-0.009
				(0.008)	(0.011)
Fence Quality	Midline	14,857	1.362	0.054	-0.055
				(0.078)	(0.099)
Property value (in USD)	Registration	19,587	1171.490	387.369	-29.377
				(321.349)	(314.303)
F Statistic, p-value				1.417, 0.186	1.423, 0.192
Panel B: Property Owner Characteristics					
Gender	Midline	9,396	0.804	0.005	0.004
				(0.016)	(0.018)
Age	Midline	8,270	51.789	0.676	-0.359
				(0.859)	(1.048)
Employed Indicator	Midline	10,295	0.789	0.018	0.007
				(0.018)	(0.021)
Salaried Indicator	Midline	10,295	0.269	-0.006	0.003
				(0.016)	(0.016)
Work for Government Indicator	Midline	10,295	0.164	-0.05	0.010
				(0.013)	(0.015)
Relative Work for Government Indicator	Midline	11,448	0.224	0.008	0.037*
				(0.017)	(0.021)
F Statistic, p-value				1.046, 0.398	0.405, 0.874
Panel C: Property Owner Characteristics					
Main Tribe Indicator	Baseline	1,404	0.722	0.056*	0.006
				(0.032)	(0.039)
Years of Education	Baseline	1,399	10.714	-0.040	-0.111
				(0.356)	(0.414)
Has Electricity	Baseline	1,404	0.108	0.041**	0.051*
				(0.021)	(0.026)
Log Monthly Income (in CF)	Baseline	1,245	10.999	0.031	0.101
				(0.080)	(0.083)
Trust Chief	Baseline	1,399	3.128	0.020	-0.080
				(0.090)	(0.104)
Trust National Government	Baseline	1,342	2.651	-0.181*	-0.126
				(0.097)	(0.110)
Trust Provincial Government	Baseline	1,348	2.503	-0.146	-0.040
				(0.104)	(0.121)
Trust Tax Ministry	Baseline	1,337	2.405	-0.075	-0.090
				(0.093)	(0.123)
F Statistic, p-value				1.299, 0.249	1.619, 0.132
Panel D: Neighborhood Characteristics					
Tax Compliance in 2016	Baseline	180	0.061	-0.011	0.013
				(0.017)	(0.025)
Tax Revenue Per Property Owner in 2016	Baseline	180	170.711	98.057	518.404
				(159.501)	(487.404)
Affected by Conflict in 2017	Baseline	180	0.000	0.031*	0.053
				(0.018)	(0.037)
F Statistic, p-value				0.511, 0.676	1.079, 0.359
Panel E: Attrition					
Registration to Midline	Registration	19,587	0.149	0.024	0.014
				(0.064)	(0.064)

Notes: This table reports coefficients from balance tests conducted by regressing baseline and midline characteristics of properties (Panel A), property owners (Panels B and C), and neighborhoods (Panel D) on an indicator for the type of the collector pair (low-high or LH, high-high or HH, with low-low or LL as the omitted category). Panel E shows differences in attrition from registration to midline surveying. Standard errors are clustered at the neighborhood level ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). All balance checks are conducted in the primary analysis sample of 180 neighborhoods, which excludes the logistics pilot, pure control, and local taxation neighborhoods in Balan et al. (2022) and exempted properties. The results are summarized in Section 3.2. The variables are described in detail in Section A14.

Figure A1: Neighborhood Chief Estimates of Household Type v. Tax Compliance



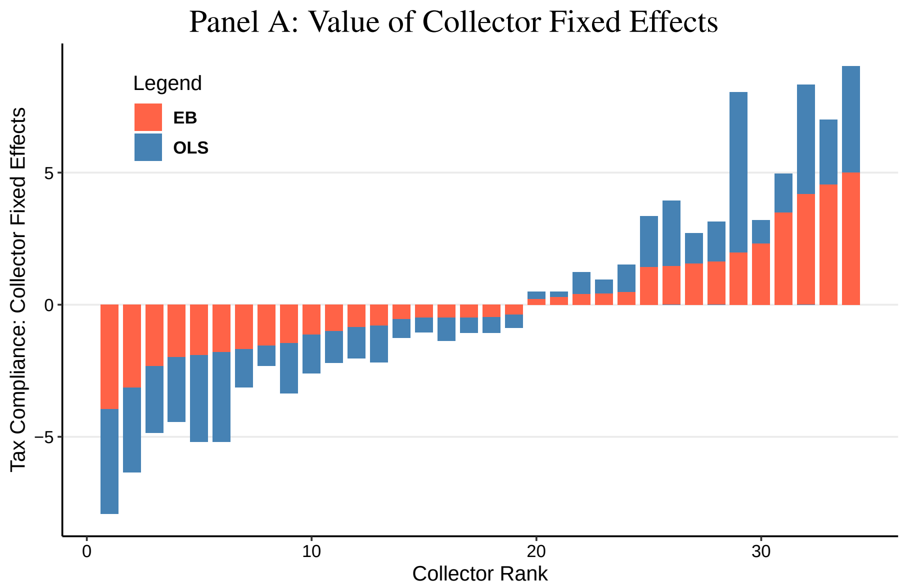
Notes: This figure shows property tax compliance by owner's ability to pay the property tax according to the neighborhood chief. Neighborhood chiefs report whether each property owner is "unlikely," "likely," or "very likely" to be able to pay the property tax. The sample comes from the 80 randomly assigned neighborhoods in the analysis sample. We discuss these results in Section 6.1.

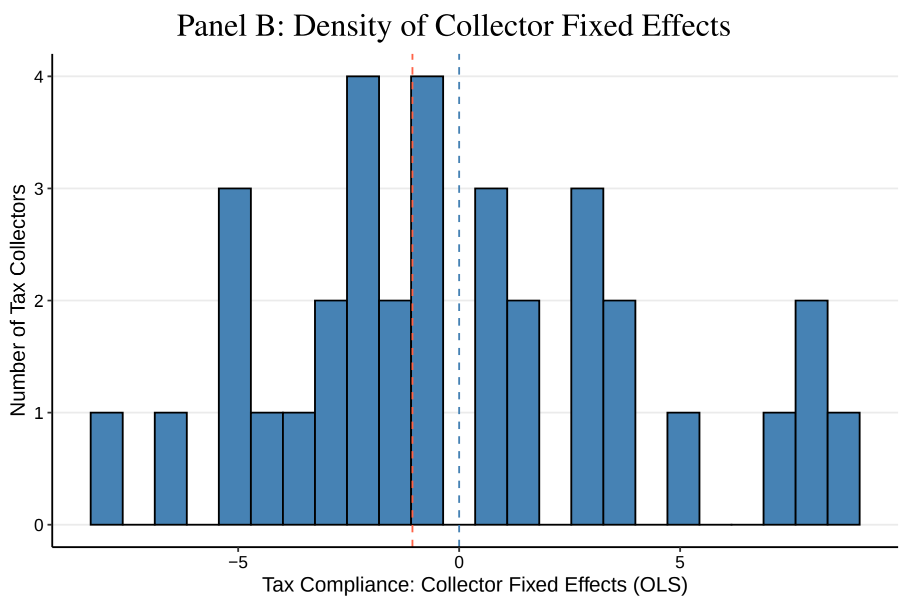
Table A3: Correlates of high-type households

	Coef. (1)	SE (2)	p-value (3)	Mean (4)	Observations (5)	Sample (6)
Panel A: Property Characteristics						
Distance to State Buildings (in km)	0.003	0.014	0.819	0.832	6,903	Registration
Distance to Health Institutions (in km)	0.011*	0.007	0.090	0.402	6,903	Registration
Distance to Education Institutions (in km)	-0.002	0.006	0.750	0.425	6,903	Registration
Distance to Roads (in km)	-0.004	0.011	0.706	0.429	6,901	Registration
Distance to Eroded Areas (in km)	-0.001	0.003	0.774	0.120	6,901	Registration
Walls Quality	0.009	0.005	0.106	0.965	5,737	Midline
Roof Quality	0.034***	0.010	0.000	1.147	5,737	Midline
Fence Quality	0.000	0.016	0.992	1.374	5,177	Midline
Property value (in USD)	276.721***	59.648	0.000	1325.137	6,903	Registration
Panel B: Property Owner Characteristics						
Employed Indicator	0.061***	0.015	0.000	0.800	3,681	Midline
Salaried Indicator	0.061***	0.015	0.000	0.253	3,681	Midline
Work for Government Indicator	0.026**	0.013	0.047	0.163	3,681	Midline
Relative Work for Government Indicator	0.039***	0.014	0.006	0.241	4,103	Midline
Panel C: Property Owner Characteristics						
Gender	-0.036	0.046	0.430	1.367	542	Baseline
Age	-2.624*	1.515	0.084	47.674	542	Baseline
Main Tribe Indicator	0.033	0.041	0.426	0.765	542	Baseline
Years of Education	0.620	0.405	0.127	10.496	542	Baseline
Has Electricity	0.051*	0.029	0.080	0.130	542	Baseline
Log Monthly Income (in CF)	0.154	0.251	0.538	10.621	540	Baseline
Trust Chief	-0.056	0.095	0.555	3.216	540	Baseline
Trust National Government	0.055	0.122	0.649	2.524	526	Baseline
Trust Provincial Government	0.030	0.120	0.806	2.426	525	Baseline
Trust Tax Ministry	-0.068	0.117	0.564	2.320	516	Baseline

Notes: This table reports the relationship between household type (low or high) and property or property owner's characteristics. More specifically, we regress each property or property owner's characteristic on an indicator for the household being high type. Columns 1–6 report the regression coefficient, robust standard errors and the associated p-values ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$), mean of the characteristic, number of non-missing observations, and the survey the data comes from (registration, midline or baseline). The characteristics are described in detail in Section A14. We discuss these results in Section 6.1.

Figure A2: Collector Fixed Effects Estimates





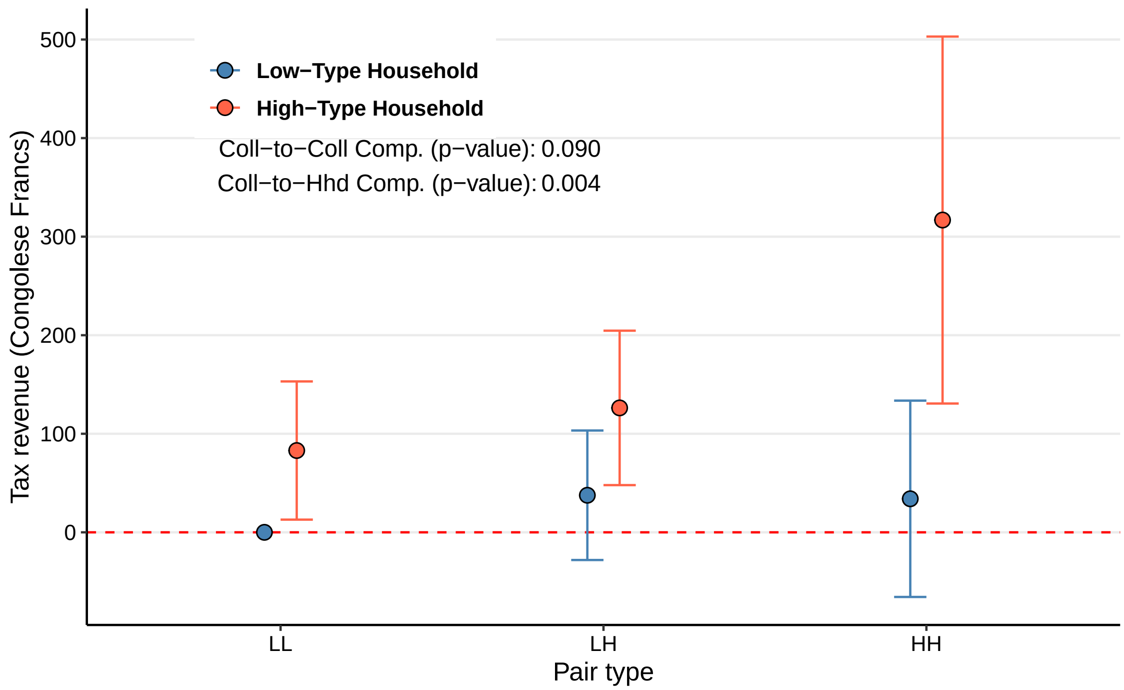
Notes: This figure shows the value and density of the tax collector fixed effects. Panel A reports the value of each collector's fixed effect (y-axis). Collector fixed effects ($\hat{\alpha}_c$) are obtained by running the OLS regression (6) with tax compliance as the outcome and are shown in blue. We also report the Empirical Bayes estimate ($\hat{\alpha}_c^{EB}$) for each collector in red. Collectors are ranked in terms of their Empirical Bayes estimate (x-axis). Panel B reports a histogram of the collector fixed effects obtained by running the OLS regression (6) with tax compliance as the outcome. A blue dotted line represents the average collector fixed effect value, while a red one represents the median value. We discuss these results in Section 6.

Table A4: Correlates of high-type collectors

	Coef. (1)	SE (2)	p-value (3)	Mean (4)	Observations (6)	Sample (7)
Panel A: Demographics						
Female	0.000	0.083	1.000	0.059	34	Baseline
Age	4.342	2.713	0.120	30.424	33	Baseline
Main Tribe	0.176	0.140	0.215	0.206	34	Baseline
Level of Education	0.507**	0.197	0.015	3.636	33	Baseline
Math Score	0.853**	0.337	0.017	-0.091	33	Baseline
Literacy (Tshiluba)	0.449	0.312	0.160	0.054	33	Baseline
Literacy (French)	0.303	0.308	0.334	0.067	33	Baseline
Monthly Income	61.388*	32.635	0.069	109.844	33	Baseline
Possessions	0.684	0.417	0.111	1.727	33	Baseline
Works Other Job	-0.040	0.169	0.813	0.667	33	Baseline
Born in Kananga	-0.154	0.177	0.389	0.545	33	Baseline
Panel B: Trust in the Government						
Trust Nat. Gov.	0.059	0.337	0.863	2.971	34	Baseline
Trust Prov. Gov.	0.235	0.306	0.448	3.000	34	Baseline
Trust Tax Min.	0.294	0.256	0.258	3.500	34	Baseline
Index	0.247	0.273	0.372	0.128	34	Baseline
Panel C: Perceived Performance of Government						
Prov. Gov. Capacity	-0.294*	0.164	0.082	0.382	34	Baseline
Prov. Gov. Responsiveness	0.000	0.310	1.000	1.765	34	Baseline
Prov. Gov. Performance	0.412	0.449	0.366	4.559	34	Baseline
Prov. Gov. Use of Funds	-0.056	0.093	0.553	0.665	33	Baseline
Index	-0.169	0.347	0.628	0.135	34	Baseline
Panel D: Government Connections						
Job through Connections	0.036	0.168	0.833	0.267	30	Baseline
Relative work for Prov. Gov.	-0.257*	0.149	0.093	0.242	33	Baseline
Relative work for Tax Ministry	-0.136	0.153	0.381	0.242	33	Baseline
Index	-0.422	0.344	0.229	-0.022	33	Baseline
Panel E: Tax Morale						
Taxes are Important	0.294*	0.158	0.073	2.794	34	Baseline
Work of Tax Min. is Important	0.000	0.173	1.000	3.765	34	Baseline
Paid Taxes in the Past	-0.083	0.223	0.713	0.381	21	Baseline
Index	0.220	0.287	0.449	0.094	34	Baseline
Panel F: Redistributive Preferences						
Imp. of Progressive Taxes	0.176	0.169	0.304	1.618	34	Baseline
Imp. of Progressive Prop. Taxes	-0.118	0.158	0.463	1.176	34	Baseline
Imp. to Tax Employed	0.353	0.248	0.164	3.353	34	Baseline
Imp. to Tax Owners	0.294	0.343	0.398	3.088	34	Baseline
Imp. to Tax Owners w. title	0.235	0.185	0.212	3.353	34	Baseline
Index	0.371	0.364	0.315	-0.294	34	Baseline

Notes: This table reports the relationship between characteristics and the type (low or high) of the tax collector. More specifically, we regress each collector's characteristic on an indicator for the collector being high type. Columns 1–6 report the regression coefficient, robust standard errors and the associated p-values ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$), mean of the characteristic among collectors, and number of non-missing observations. The variables come from a baseline surveys with tax collectors described in Section 4. We discuss these results in Section 6.2.

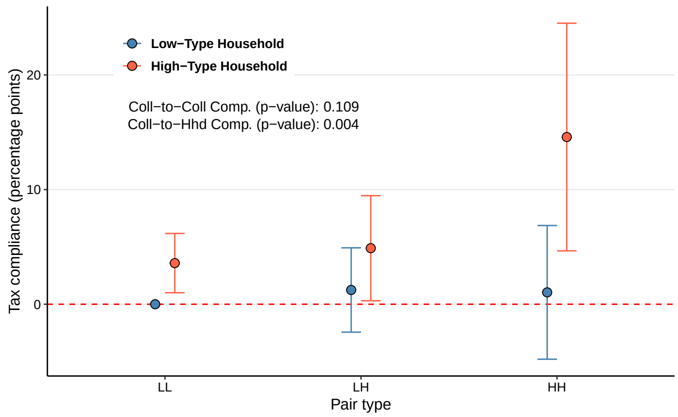
Figure A3: Tax Revenue By Collector and Household Types



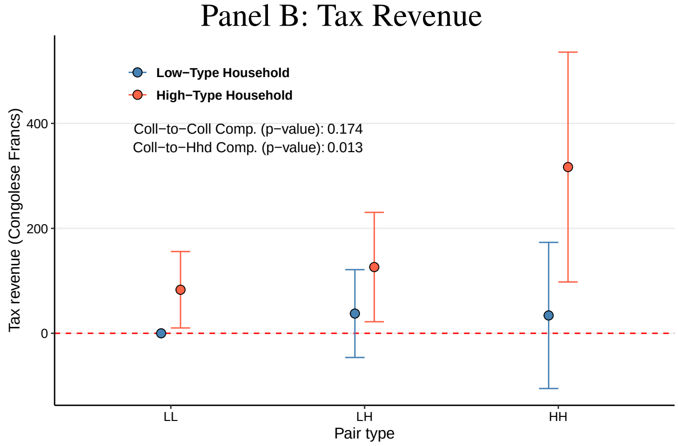
Notes: This figure shows the estimates of the average tax revenue per property owner (in Congolese Francs) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the tax revenue per property owner for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from equation (7) with tax revenue per owner as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. The figure also reports p-values associated with tests for the average tax compliance function exhibiting increasing differences in collector type and in collector and household type. We report the p-value associated with a test that tax revenue, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: [$Y(H,H,h)$-$Y(L,H,h)$] - [$Y(H,L,h)$-$Y(L,L,h)$] > 0 against H_0: [$Y(H,H,h)$-$Y(L,H,h)$] - [$Y(H,L,h)$-$Y(L,L,h)$] ≤ 0) and increasing differences in collector and household type (we test H_1: [$Y(H,H,h)$-$Y(L,L,h)$] - [$Y(H,H,l)$-$Y(L,L,l)$] > 0 against H_0: [$Y(H,H,h)$-$Y(L,L,h)$] - [$Y(H,H,l)$-$Y(L,L,l)$] ≤ 0). We discuss these results in Section 7.1.

Figure A4: Tax Compliance and Tax Revenue By Collector and Household Types — Bootstrapped Standard Errors

Panel A: Tax Compliance



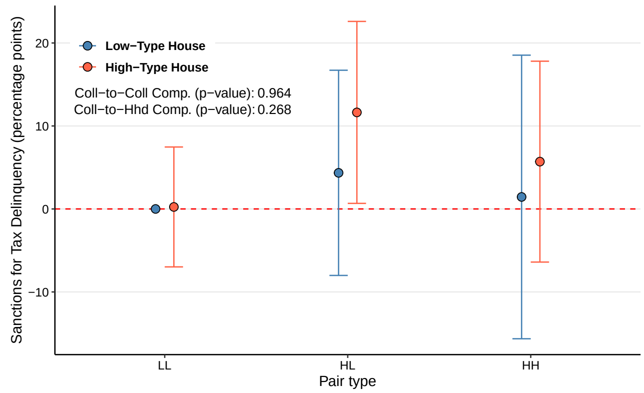
Panel B: Tax Revenue



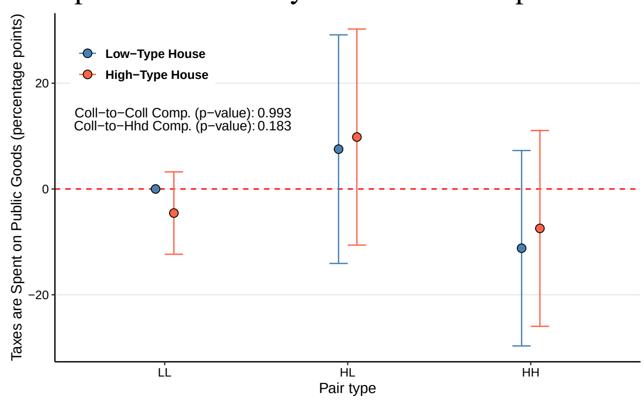
Notes: This figure shows the estimates of the average tax compliance (Panel A) and tax revenue per owner (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures tax compliance probability (Panel A) and tax revenue per owner (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax compliance (Panel A) or tax revenue (Panel B) as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates corresponding to clustered standard errors that use Bayesian bootstrap re-sampling (100 samples) at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.1.

Figure A5: Citizens' Perception of Enforcement and Use of Tax Revenue by Collector and Household Types

Panel A: Self-Reported Probability of Sanctions for Delinquency

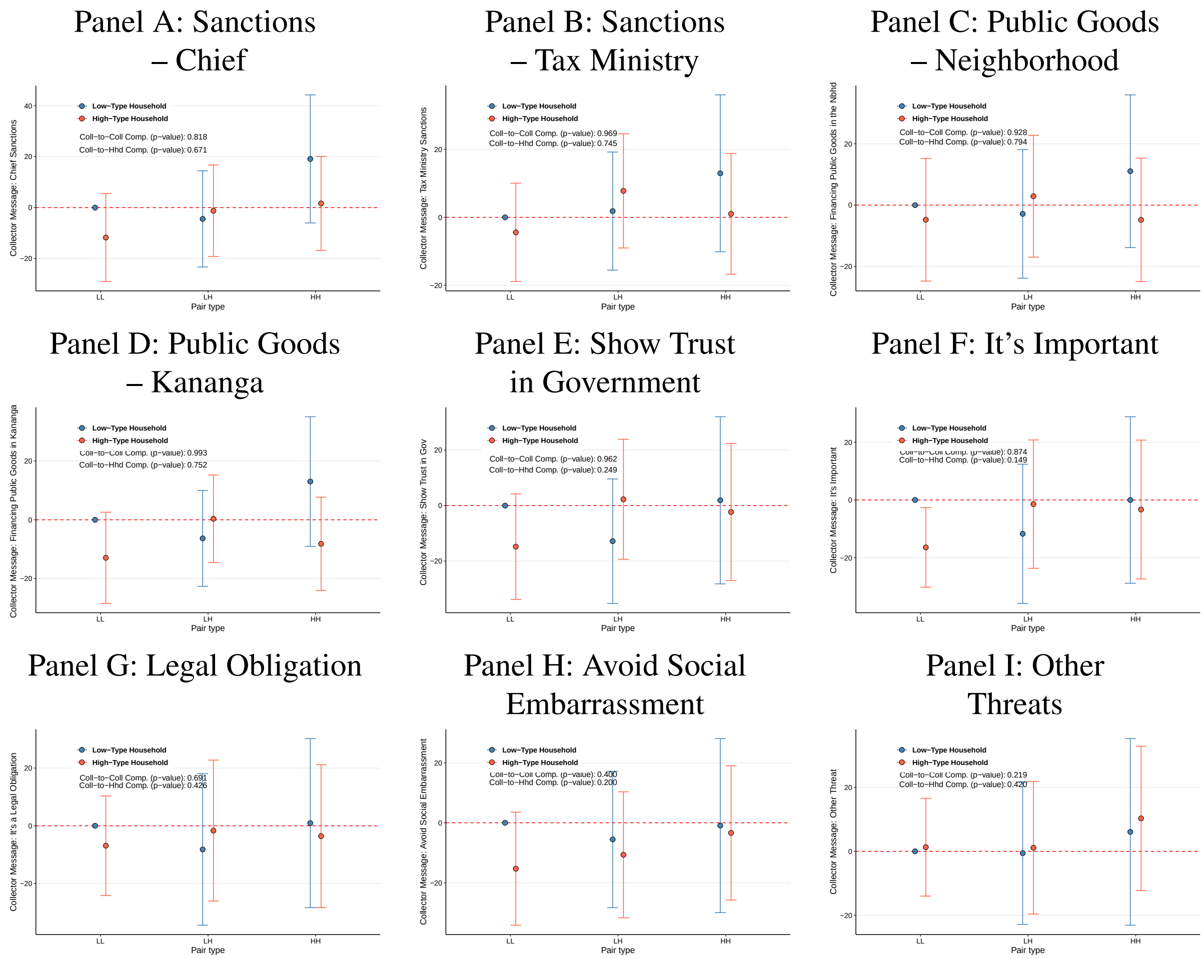


Panel B: Self-Reported Probability that Taxes are Spent on Public Goods



Notes: This figure shows the estimates of the average perception of enforcement and spending of tax revenues on public goods measured when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the perceived probability of sanctions for tax delinquency (Panel A) and the perceived probability that tax revenues are spent on public goods (Panel B) measured in the midline survey and for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from equation (7) with perception of enforcement or that tax revenues are spent on public goods as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. Each outcome is multiplied by 100 so the coefficients can be interpreted as percentage point changes. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.2.

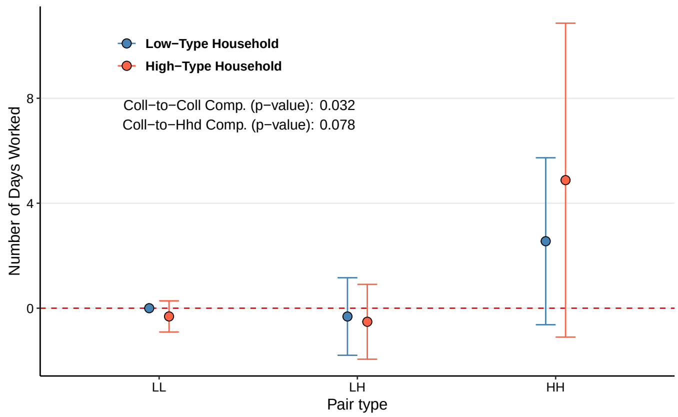
Figure A6: Collectors' Strategies by Collector and Household Types



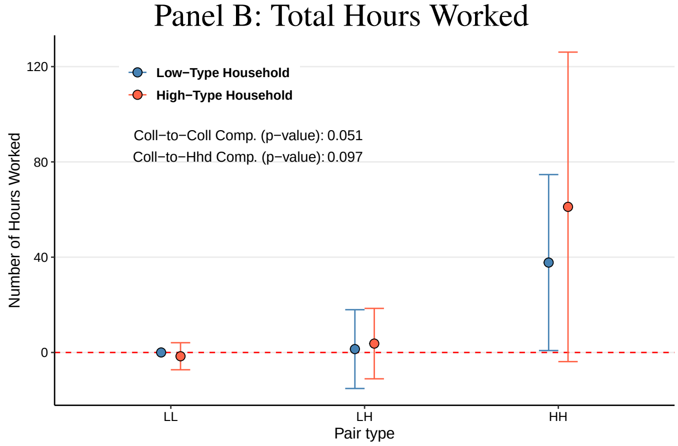
Notes: This figure shows the estimates of the different possible messages used by collectors when soliciting payment when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the messages used by collectors when demanding payment measured in the endline survey and for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from equation (7) with the collectors' message as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. Each outcome is an indicator for whether the collector used the message, multiplied by 100 so the coefficients can be interpreted as percentage point changes. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.2.

Figure A7: Days and Hours Collectors Worked by Collector and Household Types

Panel A: Distinct Days Worked



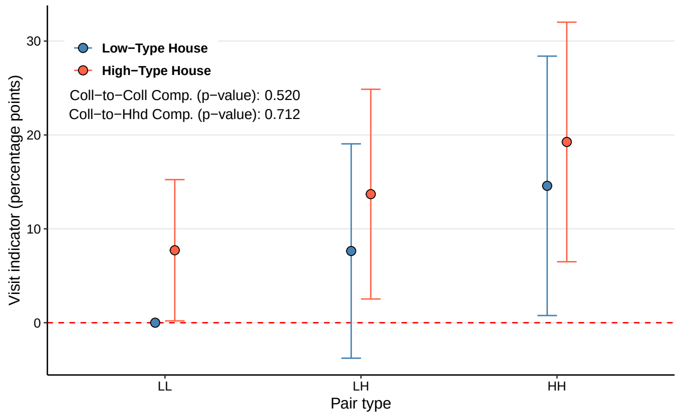
Panel B: Total Hours Worked



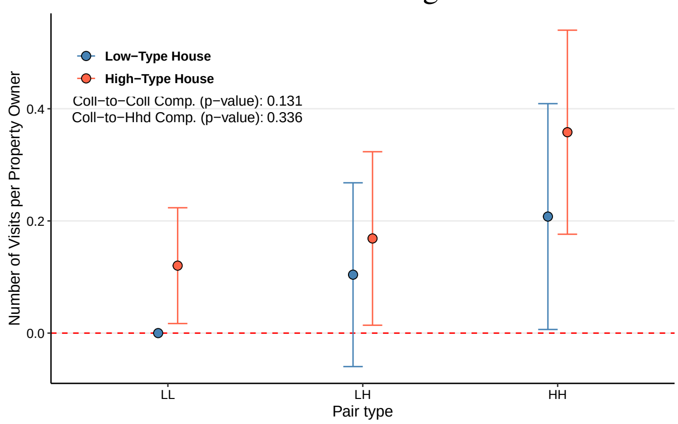
Notes: This figure shows the estimates of distinct days worked by the tax collectors (Panel A) and the total number of hours worked by the tax collectors (Panel B) for different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis uses the dated chalk marks midline survey data and the tax receipt data to captures numbers of days worked (Panel A) and number of hours worked (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax visits as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: [$Y(H,H,h)$-$Y(L,H,h)$] - [$Y(H,L,h)$-$Y(L,L,h)$] > 0 against H_0: [$Y(H,H,h)$-$Y(L,H,h)$] - [$Y(H,L,h)$-$Y(L,L,h)$] ≤ 0) and increasing differences in collector and household type (we test H_1: [$Y(H,H,h)$-$Y(L,L,h)$] - [$Y(H,H,l)$-$Y(L,L,l)$] > 0 against H_0: [$Y(H,H,h)$-$Y(L,L,h)$] - [$Y(H,H,l)$-$Y(L,L,l)$] ≤ 0). We discuss these results in Section 7.2.

Figure A8: Tax Visits by Collector and Household Types

Panel A: Post-Registration Visit Indicator



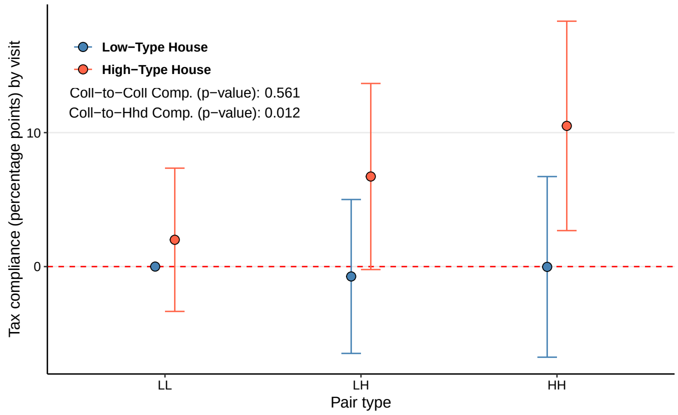
Panel B: Number of Post-Registration Visits



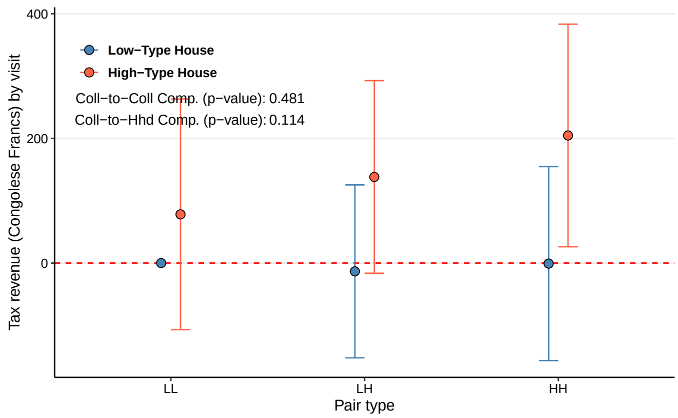
Notes: This figure shows the estimates of post-registration extensive margin visits (Panel A) and intensive margin number of visits (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures extensive margin tax visits (Panel A) and intensive margin number of tax visits (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax visits as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.2.

Figure A9: Productivity of Tax Visits by Collector and Household Types

Panel A: Tax Compliance per Post-Registration Visit



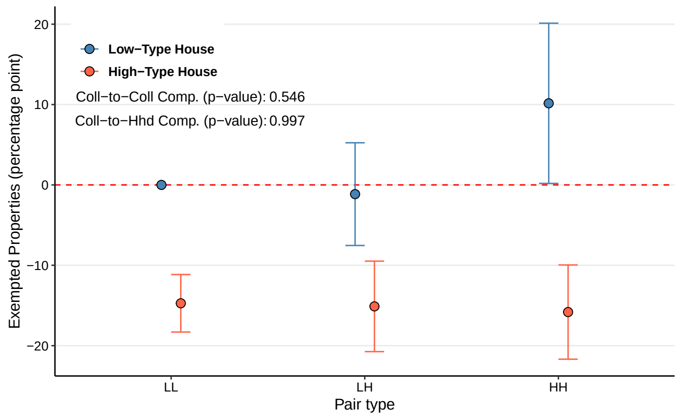
Panel B: Tax Revenue per Post-Registration Visit



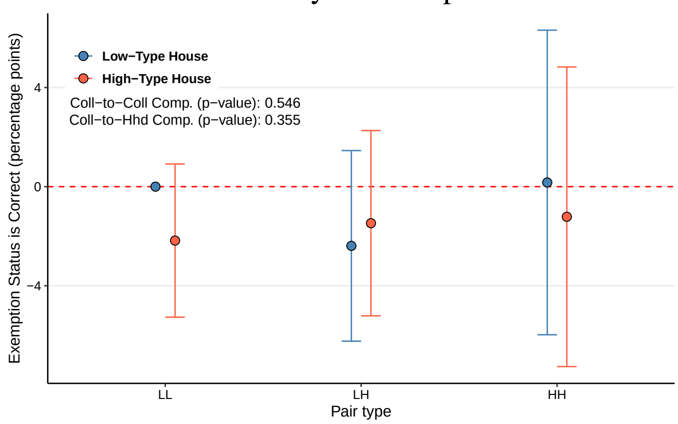
Notes: This figure shows the estimates of tax compliance per post-registration visits (Panel A) and tax revenue per post-registration visits (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures tax compliance (in percentage points) per tax visit (Panel A) tax revenue (in Congolese Frances) per tax visit (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax compliance or revenue per visit as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.2.

Figure A10: Exemption by Collector and Household Types

Panel A: Exemption Status



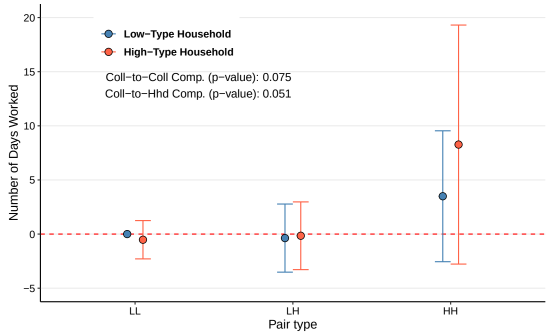
Panel B: Accuracy of Exemption Status



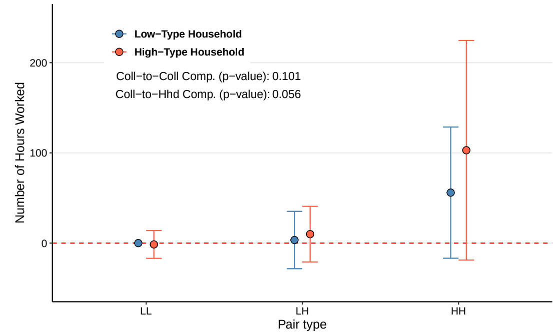
Notes: This figure shows the estimates of the property's tax exemption status at registration (Panel A) and whether this exemption status was deemed accurate by the enumerator during the registration survey (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the exemption status of the houshold (Panel A) and whether this exemption status was judged accurate by the enumerator (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax exemption status (Panel A) or the accuracy of this exemption status (Panel B) as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 7.2.

Figure A11: Days and Hours Collectors Worked by Collector Types, Household Types, and Employment Rates

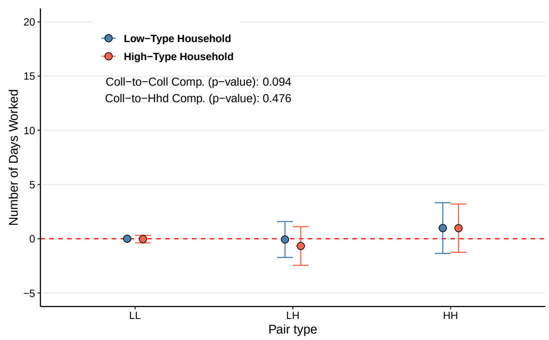
Panel A: Distinct Days Collectors Worked Above Median Employment Rate Nbhd



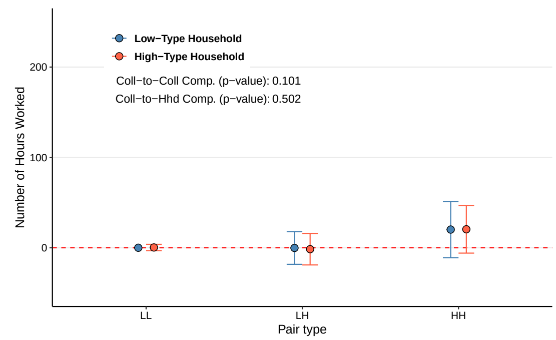
Panel B: Total Hours Collectors Worked Above Median Employment Rate Nbhd



Panel C: Distinct Days Collectors Worked Below Median Employment Rate Nbhd

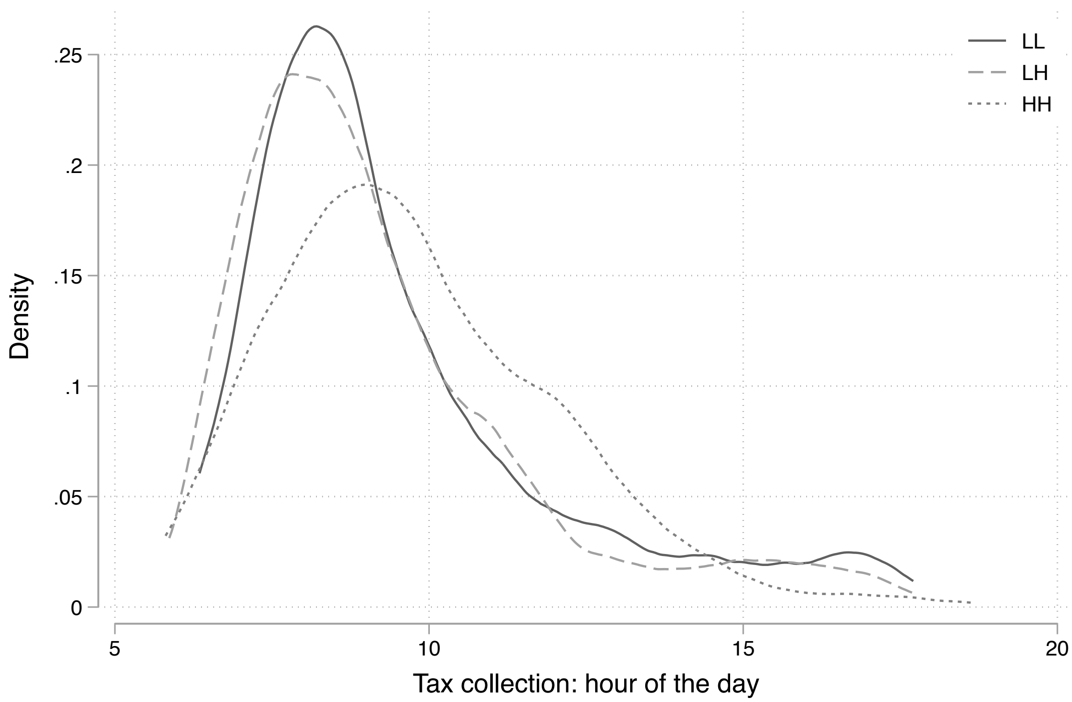


Panel D: Total Hours Collectors Worked Below Median Employment Rate Nbhd



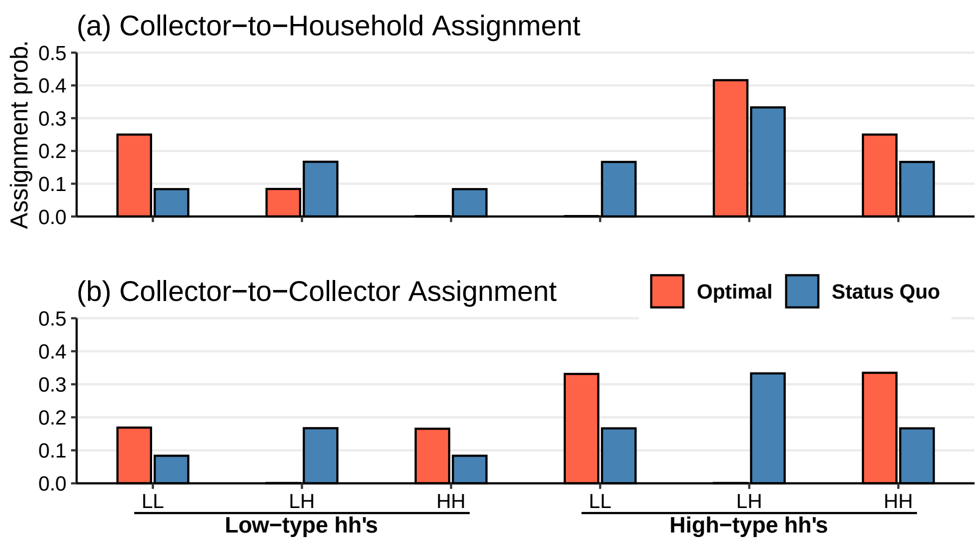
Notes: This figure shows the estimates of distinct days worked by the tax collectors (Panel A and C) and the total number of hours worked by the tax collectors (Panel B and D) for different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The estimation is reported for neighborhoods characterized by an above median level of employment (Panel A and B) and a below median level of employment (Panel C and D). The x-axis shows the three different types of collectors' pair: LL, LH, HH. The y-axis uses the dated chalk marks midline survey data and the tax receipt data tax to captures numbers of days worked (Panel A) and number of hours worked (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from Equation (7) with tax visits as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: [$Y(H,H,h)$-$Y(L,H,h)$] - [$Y(H,L,h)$-$Y(L,L,h)$] > 0 against H_0: [$Y(H,H,h)$-$Y(L,H,h)$] - [$Y(H,L,h)$-$Y(L,L,h)$] ≤ 0) and increasing differences in collector and household type (we test H_1: [$Y(H,H,h)$-$Y(L,L,h)$] - [$Y(H,H,l)$-$Y(L,L,l)$] > 0 against H_0: [$Y(H,H,h)$-$Y(L,L,h)$] - [$Y(H,H,l)$-$Y(L,L,l)$] ≤ 0). We discuss these results in Section 7.2.

Figure A12: Time of Tax Collection by Collector Types



Notes: This figure shows the distribution of tax collection time within the day for different types of collector pairs (low-low or LL, low-high or LH, high-high or HH). Information on the precise date and time (including hour, minute, second) at which each tax collection took place comes from the tax receipt data. We discuss these results in Section 7.2.

Figure A13: Collector-to-Household and Collector-to-Collector Optimal Assignments



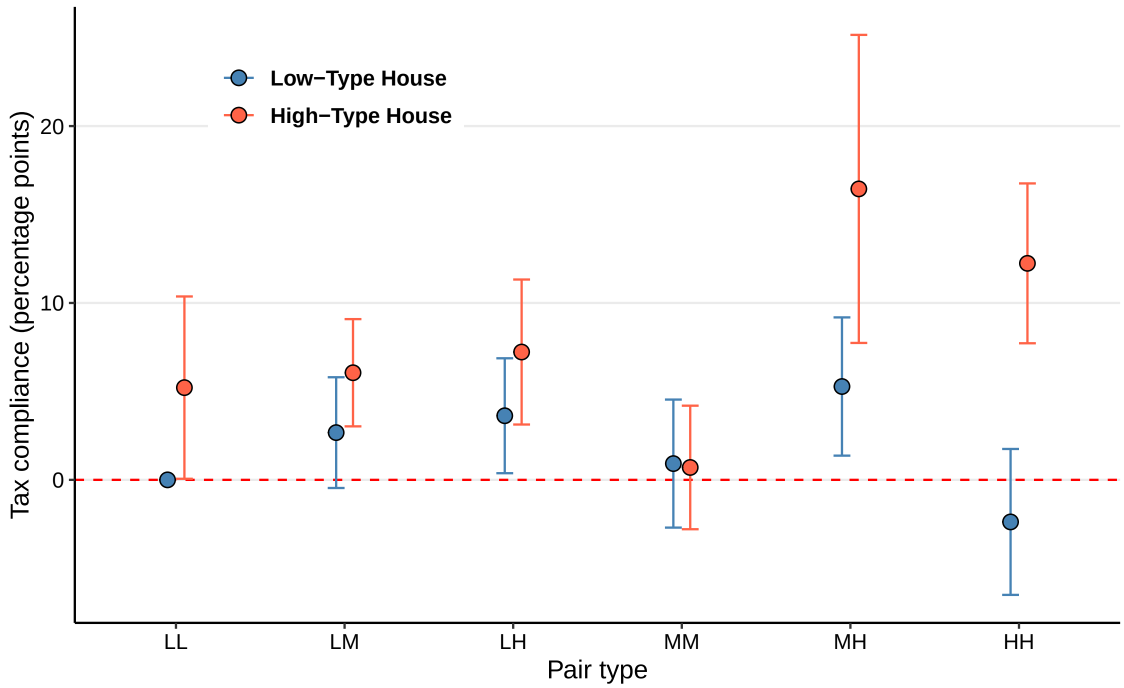
Notes: This figure shows the assignment function from two alternative counterfactual optimal assignment mechanisms in comparison to the status quo assignment. Panel A shows the collector-to-household-only counterfactual optimal assignment. Panel B shows the collector-to-collector-only counterfactual optimal assignment. In both graphs, each bar represents the probability of each match type under the optimal (red) and status quo (blue) assignment functions. The first 6 bars show the assignment functions with matches involving low-type households. The 6 subsequent bars show the assignment functions with matches involving high-type households. We discuss these results in Section 8.1.

Table A5: Effects of the Optimal Assignment: Compliance and Revenues – Standard Vs Bootstrapped Standard Errors

	Household Types: Household Propensity to Pay			
	Collector Types: Fixed Effects Model			
	Standard Errors: Clustered at Neighborhood-Level		Standard Errors: Bayesian Bootstrap	
	Tax Compliance (in percentage points) (1)	Tax Revenue (in Congolese Francs) (2)	Tax Compliance (in percentage points) (3)	Tax Revenue (in Congolese Francs) (4)
Optimal Assignment	2.941**	54.471*	2.941*	54.471
	(1.239)	(30.52)	(1.682)	(37.872)
	[0.024]	[0.074]	[0.080]	[0.150]
Collector-to-Collector Only	1.294	21.444	1.294	21.444
	(0.947)	(21.675)	(1.308)	(30.373)
	[0.172]	[0.322]	[0.323]	[0.480]
Collector-to-Household Only	0.837***	17.156**	0.837**	17.156*
	(0.312)	(8.520)	(0.384)	(9.929)
	[0.007]	[0.044]	[0.029]	[0.084]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,732	11,732
Observations (Analysis Sample)	6,904	6,904	6,904	6,904

Notes: This table shows the impact of the counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1 and 3 show results for probability of compliance, i.e., a dummy indicating whether households paid the property tax (multiplied by 100). The point estimates should be interpreted as percentage point changes. Columns 2 and 4 show results for average tax revenue per household in Congolese Francs. All columns present results when collector types are estimated using a fixed effects model as described in Section 6.2. Each row shows counterfactual results for a different optimal assignment. The first row presents results when optimizing on both the collector-to-household and the collector-to-collector dimension. The second and third rows show results when only optimizing the collector-to-household and the collector-to-collector dimension of the assignment, respectively. We report conventional clustered standard errors at the neighborhood level in Columns 1 and 2. In Columns 3 and 4, we instead report standard errors from Bayesian bootstrap re-sampling at the neighborhood level (100 samples) in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax compliance (Columns 1 and 3) and tax revenue (Columns 2 and 4) is reported at the bottom of the table. We also report the size of the holdout and analysis sample. We discuss these results in Section 6.4 and 8.1.

Figure A14: Tax Compliance By Collector and Household Types– Three Types of Collectors



Notes: This figure shows the estimates of average tax compliance when assigned to different types of collector pairs (low-low or LL, low-medium or LM, low-high or LH, medium-medium or MM, medium-high or MH, high-high or HH) by households' type (low or high). The x-axis shows the six different types of collector pairs: LL, LM, LH, MM, MH, HH. The y-axis captures tax compliance probability for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from equation (7) with tax compliance as the outcome and when including eleven dummies: (H, H, h), (H, H, l), (M, H, h), (M, H, l), (M, M, h), (M, M, l), (L, H, h), (L, H, l), (L, M, h), (L, M, l), and (L, L, h) (the excluded category is (L, L, l)) reflecting matches of households of type $V = l, h$ and collectors of type $A = L, M, H$. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We discuss these results in Section 8.2.

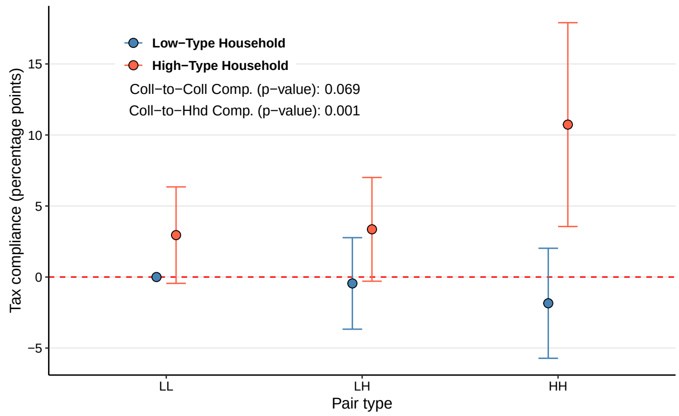
Table A6: Effects of the Optimal Assignment on Tax Compliance and Revenues – Three Types of Collectors

	Household Types: Household Propensity to Pay			
	Collector Types: Fixed Effects Model		Collector Types: Coll. Chars. Model	
	Tax Compliance (in percentage points) (1)	Tax Revenue (in Congolese Francs) (2)	Tax Compliance (in percentage points) (3)	Tax Revenue (in Congolese Francs) (4)
Optimal Assignment	4.411**	62.212	3.296	49.675
	(2.062)	(48.797)	(2.135)	(44.713)
	[0.032]	[0.202]	[0.123]	[0.267]
Collector-to-Collector Only	3.105**	73.921*	1.592	36.288
	(1.542)	(39.767)	(1.741)	(37.677)
	[0.044]	[0.063]	[0.360]	[0.335]
Collector-to-Household Only	1.345***	38.887***	1.271***	30.219***
	(0.335)	(9.731)	(0.354)	(8.498)
	[0.000]	[0.000]	[0.000]	[0.000]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,732	11,732
Observations (Analysis Sample)	6,904	6,904	6,904	6,904

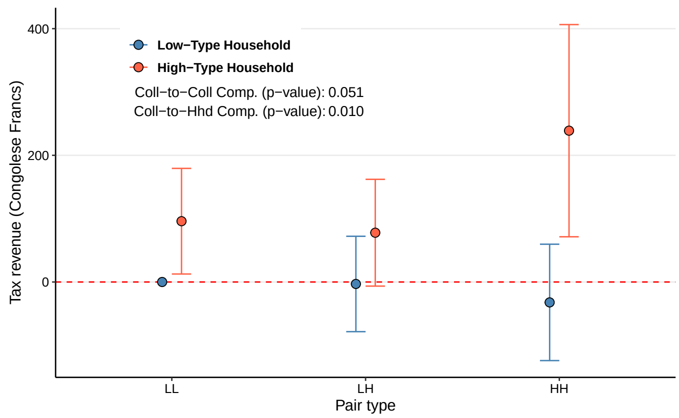
Notes: This table shows the impact of the counterfactual optimal assignment policy with three types of tax collectors (low or L, medium or M, high or H), relative to the status quo (random) assignment. Columns 1 and 3 show results for probability of compliance, i.e., a dummy indicating whether households paid the property taxes (multiplied by 100). The point estimates should be interpreted as percentage point changes. Columns 2 and 4 show results for average tax revenue per household in Congolese Francs. Columns 1–2 present results when collectors' types are estimated using a fixed effects model as described in Section 6.2. Columns 3–4 show results when collectors' types are estimated from tax collectors' characteristics as described in Section 8.2. Each row represents a counterfactual for a different optimal assignment. The first row presents results when optimizing on both the collector-to-household and the collector-to-collector dimension. The second and third rows show results when only optimizing the collector-to-collector and the collector-to-household dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax compliance (Columns 1 and 3) and tax revenue (Columns 2 and 4) is reported at the bottom of the table. We also report the size of the holdout and analysis sample. We discuss these results in Section 8.2.

Figure A15: Tax Compliance and Revenue By Collector and Household Types – Collectors' Type: Collector Characteristics Model

Panel A: Tax Compliance



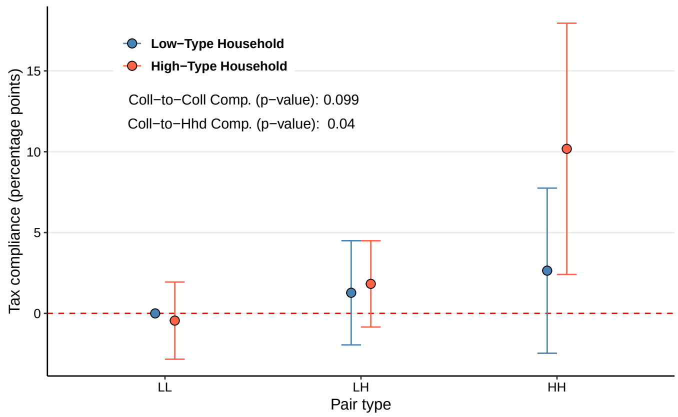
Panel B: Tax Revenue



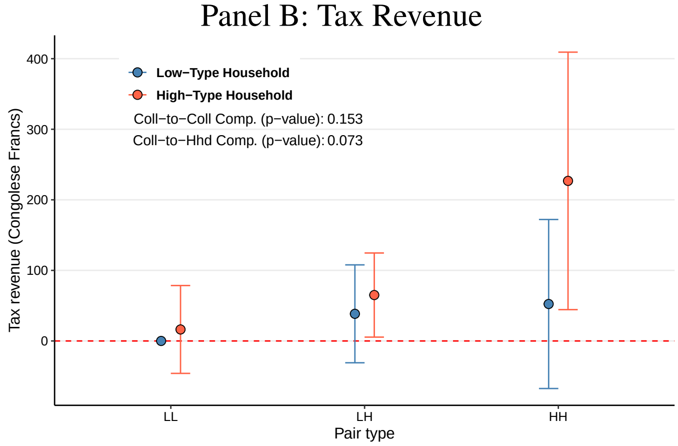
Notes: This figure shows the estimates of average tax compliance when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). Collectors' types are estimated from tax collectors' characteristics as described in Section 8.2. The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the tax compliance probability (Panel A) or tax revenue (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from equation (7) with tax compliance as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 8.2.

Figure A16: Tax Compliance and Revenue By Collector and Household Types – Households' Type: Households Characteristics Model

Panel A: Tax Compliance

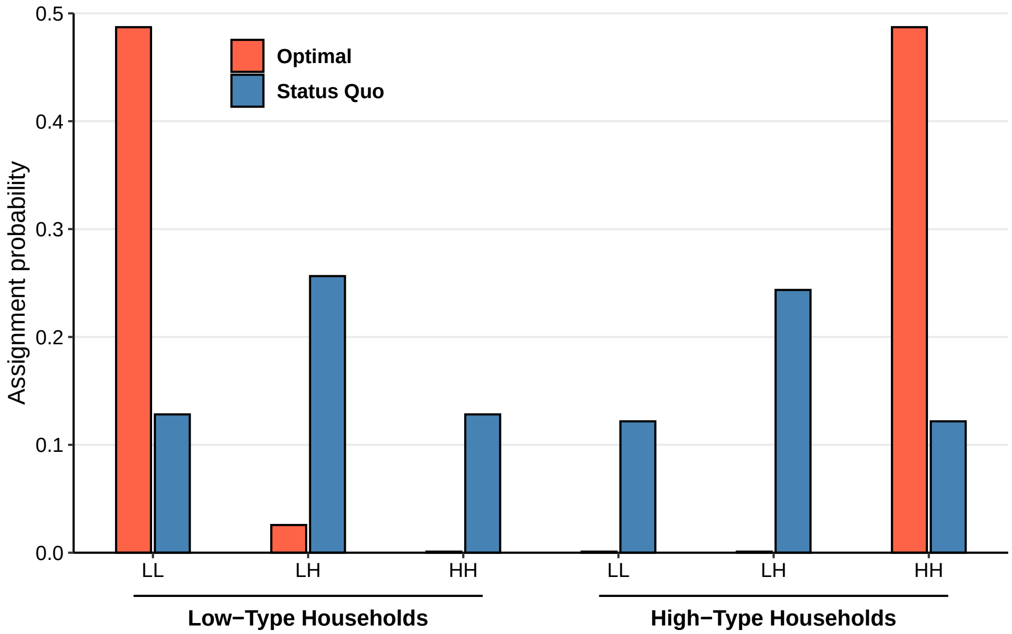


Panel B: Tax Revenue



Notes: This figure shows the estimates of average tax compliance when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). Collectors' types are estimated from the fixed effects model described in Section 6.2 and household types are estimated using household characteristics as described in Section 8.2. The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the tax compliance probability (Panel A) or tax revenue (Panel B) for different types of collector pairs and households. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The point estimates are estimated from equation (7) with tax compliance as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test that each outcome, denoted Y, exhibits increasing differences in collector type for high-type households (we test H_1: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,H,h)]$ - $[Y(H,L,h)$-$Y(L,L,h)] \leq 0$) and increasing differences in collector and household type (we test H_1: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] > 0$ against H_0: $[Y(H,H,h)$-$Y(L,L,h)]$ - $[Y(H,H,l)$-$Y(L,L,l)] \leq 0$). We discuss these results in Section 8.2.

Figure A17: Optimal Vs. Status Quo Assignments – Households' Type: Households Characteristics Model



Notes: This figure shows the optimal and the status quo assignment functions. Each bar represents the probability of each match type under the optimal (red) and status quo (blue) assignment functions. The first 6 bars show the assignment functions with matches involving low-type households. The 6 subsequent bars show the assignment functions with matches involving high-type households. We discuss these results in Section 8.2.

Table A7: Effects of the Optimal Assignment on Compliance and Revenues – Household Types: Households Characteristics Model

	Collector Types: Fixed Effects Model			
	Household Types: Household Propensity to Pay		Household Types: Household Chars. Model	
	Tax Compliance (in percentage points) (1)	Tax Revenue (in Congolese Francs) (2)	Tax Compliance (in percentage points) (3)	Tax Revenue (in Congolese Francs) (4)
Optimal Assignment	2.941**	54.471*	2.759*	50.417
	(1.239)	(30.52)	(1.504)	(34.836)
	[0.024]	[0.074]	[0.067]	[0.148]
Collector-to-Collector Only	1.294	21.444	0.773	11.085
	(0.947)	(21.675)	(0.770)	(17.251)
	[0.172]	[0.322]	[0.315]	[0.520]
Collector-to-Household Only	0.837***	17.156**	1.000*	19.828
	(0.312)	(8.520)	(0.572)	(13.622)
	[0.007]	[0.044]	[0.080]	[0.146]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,732	11,732
Observations (Analysis Sample)	6,904	6,904	7,866	7,866

Notes: This table shows the impact of the counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1 and 3 show results for probability of compliance, i.e., a dummy indicating whether households paid the property taxes (multiplied by 100). The point estimates should be interpreted as percentage point changes. Columns 2 and 4 show results for average tax revenue per household in Congolese Francs. All columns present results when collectors' types are estimated using a fixed effects model as described in Section 6.2. In Columns 1–2, household types are defined using chiefs' estimates of household type as described in Section 6.1. The results are therefore identical to Columns 1–2 of Table 1. In Columns 3–4, household types are estimated using household characteristics as described in Section 8.2. Each row represents a counterfactual for a different optimal assignment. The first row presents results when optimizing on both the collector-to-household and the collector-to-collector dimension. The second and third rows show results when only optimizing the collector-to-collector and the collector-to-household dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax compliance (Columns 1 and 3) and tax revenue (Columns 2 and 4) is reported at the bottom of the table. We also report the size of the holdout and analysis sample. We discuss these results in Section 8.1 and 8.2.

Table A8: Effects of the Optimal Assignment on Tax Compliance and Revenue – Objective: Tax Revenue Maximization

	Household Types: Household Propensity to Pay			
	Objective: Tax Revenue Maximization		Objective: Tax Revenue Net of Bribes Maximization	
	Tax Revenue (in Congolese Francs) (1)	Bribe Payments (in Congolese Francs) (2)	Tax Revenue (in Congolese Francs) (3)	Bribe Payments (in Congolese Francs) (4)
Optimal Assignment	61.014**	14.902	37.256	-0.404
	(26.179)	(12.447)	(29.925)	(4.783)
	[0.020]	[0.231]	[0.213]	[0.933]
Collector-to-Collector Only	36.530*	5.734	38.225*	4.197
	(21.871)	(7.101)	(23.195)	(5.747)
	[0.095]	[0.419]	[0.099]	[0.465]
Collector-to-Household Only	15.631*	2.206	18.669*	5.596**
	(8.208)	(3.188)	(10.138)	(2.757)
	[0.057]	[0.489]	[0.066]	[0.042]
Mean	206.213	30.431	206.213	30.431
Observations (Holdout Sample)	11,732	11,732	7,694	7,694
Observations (Analysis Sample)	6,904	4,691	6,904	4,691

Notes: This table shows the impact of the counterfactual optimal assignment policy, in the case where the government aims at maximizing tax revenue or tax revenue net of bribes, relative to the status quo (random) assignment. Columns 1 and 3 show results for average tax revenue per household in Congolese Francs. Columns 2 and 4 show results for average bribe payments per household in Congolese Francs, drawn from midline surveys. All columns present results when collector types are estimated using a fixed effects model as described in Section 6.2 and household types are defined using chiefs' estimates of household type as described in Section 6.1. Each row represents a counterfactual for a different optimal assignment. The first row presents results when optimizing on both the collector-to-household and the collector-to-collector dimension. The second and third rows show results when only optimizing the collector-to-collector and the collector-to-household dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax revenue (Columns 1 and 3) and bribe amount (Columns 2 and 4) is reported at the bottom of the table. We also report the size of the holdout and analysis sample. We discuss these results in Section 8.2.

Table A9: Effects of the Neighborhood-Level Optimal Assignment: Compliance and Revenues

	Neighborhood Type: Share of High-Type Households		Neighborhood Type: Number of High-Type Households	
	Tax Compliance (in percentage points) (1)	Tax Revenue (in Congolese Francs) (2)	Tax Compliance (in percentage points) (3)	Tax Revenue (in Congolese Francs) (4)
Optimal Assignment	1.764*	30.667	2.906**	56.181*
	(1.023)	(23.572)	(1.472)	(34.232)
	[0.085]	[0.193]	[0.048]	[0.100]
Collector-to-Collector Only	1.159	18.606	2.802*	54.250
	(0.915)	(20.901)	(1.465)	(33.994)
	[0.205]	[0.373]	[0.056]	[0.111]
Collector-to-Household Only	0.260***	5.315**	1.408***	30.146**
	(0.099)	(2.531)	(0.532)	(12.749)
	[0.009]	[0.036]	[0.008]	[0.018]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,732	11,732
Observations (Analysis Sample)	6,904	6,904	6,904	6,904

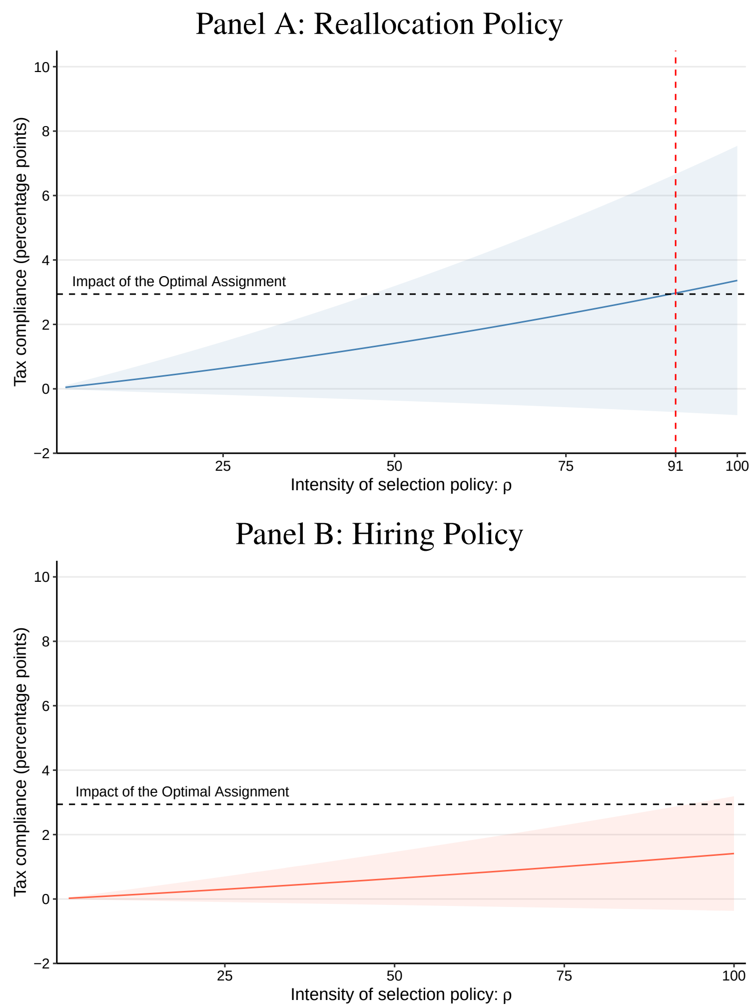
Notes: This table shows the impact of the neighborhood-level counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1–2 assume that the government defines neighborhoods type based on the share of high and low type households. Columns 3–4 instead assume that the government defines neighborhood type based on the number of high and low type households. The coefficients in Columns 1 and 3 show the impact on tax compliance, i.e., a dummy indicating whether households paid the property taxes (multiplied by 100). The point estimates should be interpreted as percentage point changes. Columns 2 and 4 show results for average tax revenue per household in Congolese Francs. All the results use collector types estimated using a fixed effects model as described in Section 6.2 and property types are estimated as described in Section 6.1. Each row represents a counterfactual for a different optimal assignment. The first row presents results when optimizing on both the collector-to-household and the collector-to-collector dimension. The second and third rows show results when only optimizing the collector-to-collector and the collector-to-household dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax compliance (Columns 1 and 3) and tax revenue (Columns 2 and 4) is reported at the bottom of the table. We also report the size of the holdout and analysis sample. We discuss these results in Section 8.2.

Table A10: Effects of the Optimal Assignment on Tax Compliance and Revenue – Robustness: Inference on Winners

	Objective: Compliance Maximization		Objective: Revenue Maximization	
	Tax Compliance (in percentage points)	Tax Revenue (in Congolese Francs)	Tax Compliance (in percentage points)	Tax Revenue (in Congolese Francs)
	(1)	(2)	(3)	(4)
Benchmark Estimator	2.941	54.471	3.172	61.014
	[0.394–5.488]	[-5.361–114.302]	[0.773–5.570]	[9.703–112.325]
Conditional Estimator	2.897	51.229	3.160	60.554
	[0.311–5.027]	[-18.562–103.222]	[0.890–5.138]	[10.653–103.063]
Hybrid Estimator	2.890	51.296	3.162	60.592
	[0.324–5.053]	[-16.452–104.095]	[0.884–5.163]	[10.560–103.629]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,732	11,732
Observations (Analysis Sample)	6,904	6,904	6,904	6,904

Notes: This table provides estimates and 90% confidence intervals for the impact of the counterfactual optimal policy after accounting for possible over-fitting concerns associated with the "winner's curse" problem (Andrews et al., 2021). We adapt Andrews et al. (2021) to our context, a non-discrete optimal assignment policy space, in several steps. First, the solution must lie at the intersection of three hyper-planes defined by the two linearly independent constraints in Problem 1 and the requirement that the distribution probabilities sum up to 1. Second, the Fundamental Theorem of Linear Programming (Dantzig, 1951) — which states that if an optimal solution exists, there exists an optimal solution consisting of extreme points on the policy space — allows us to select three points in this 3 dimensional space. We focus on the three solutions in the (finite) set of extreme points that are linearly independent and that yield the highest value when applied to the objective function. Row 1 provides our baseline estimates from Table 1 and Table A8. Rows 2 and 3 provide the conditional and hybrid estimators suggested by Andrews et al. (2021). Columns 1-2 examine the case in which the government seeks to maximize tax compliance, while Columns 3-4 examines the revenue maximization case. The average tax compliance (Columns 1 and 3) and tax revenue (Columns 2 and 4) is reported at the bottom of the table. We also report the size of the holdout and analysis sample. We discuss these results in Section 8.2.

Figure A18: Effects of Selection Policies when Collector Types are Estimated using Collectors' Characteristics



Notes: This figure shows the impact of the selection policies on the probability of tax compliance (y-axis). Selection policies involve reassigning $\rho\%$ (x-axis) of the assignments that a low-ability collector would receive under the status quo assignment to other collectors. Panel A shows the estimated effects of the *reallocation policy*, where the workload is reassigned to existing high-ability collectors in the sample. Panel B shows the estimated effects of the *hiring policy*, where the workload is reassigned to newly hired collectors with types drawn uniformly from {L,H}. In both Panels, collector types are estimated from tax collectors' characteristics as described in Section 8.2. The shaded areas represent 95% confidence intervals. The dashed horizontal line indicates the counterfactual impact of the optimal assignment policy on tax compliance when collector types are estimated from tax collectors' characteristics as reported in Column 3 of Table 1. We discuss these results in Section 9.1.

Table A11: Effect of Collectors' Wage Increases

	Tax Compliance	Tax Revenue	Visit Indicator	Nb of Visits	Bribe Indicator	Bribe Amount
log. Wage	0.037**	54.126**	0.046	0.104**	0.010	9.281
	(0.015)	(25.113)	(0.030)	(0.049)	(0.007)	(8.017)
Mean	0.074	153.609	.415	0.546	0.016	1288.265
Elasticity	0.492	0.352	0.110	0.190	0.643	0.461
Observations	18,775	18,775	12,525	12,383	12,544	196
Tax Rate FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table examines treatment effects of the collectors' piece-rate wage on tax compliance, tax revenues, tax visits, and bribe payments. It reports the results of regressions of the log of the piece-rate wage on tax compliance (Columns 1), tax revenue (Columns 2), a post-registration visit indicator (Column 3), the number of post-registration visits (Column 4), an indicator for any bribe payment (Column 5), and the amount of bribe paid (Column 6). We discuss these results in Section 9.2.

Table A12: Effect of Enforcement Messages

	Tax Compliance			Tax Revenue (in CF)		
	(1)	(2)	(3)	(4)	(5)	(6)
Central Enforcement	0.014	0.016*		32.837*	36.510**	
	(0.009)	(0.009)		(18.610)	(18.453)	
Local Enforcement	0.014	0.016*		31.244*	35.545*	
	(0.009)	(0.009)		(18.723)	(18.783)	
Pooled Enforcement			0.016**			36.038**
			(0.007)			(15.589)
Observations	2665	2665	2665	2665	2665	2665
Mean	0.029	0.029	0.029	57.671	57.671	57.671
FE: neighborhood	No	Yes	Yes	No	Yes	Yes

Notes: This table examines treatment effects of randomized tax letter enforcement messages on compliance and revenues. It reports estimates from a regression of tax compliance (Columns 1–3) and tax revenue (Columns 4–6) on treatment dummies for households assigned to enforcement messages on tax letters distributed during property registration. Bergeron et al. (2021) describe these tax letters and the message randomization. The excluded category is the control message in all regressions. Columns 2–3 and 5–6 introduce randomization stratum (neighborhood) fixed effects. Columns 3 and 6 pool households assigned to the *central enforcement* message and the *local enforcement* message. The data are restricted to the sample of 2,665 properties subject to randomized messages on tax letters, which were introduced toward the end of the tax campaign. We discuss these results in Section 9.2.

A2 Properties of the Optimal Assignment Function

A2.1 Uniqueness

The optimal assignment problem is a linear program. As a consequence its solutions are constrained to be in a convex set, implying that it has at least one solution (Luenberger, 1984). However, there might be more than one solution to the optimal assignment problem.[1] We follow Bhattacharya (2009) and assume uniqueness of the optimal assignment.

Assumption 1. *There exists a unique f^* that solves the Optimal Assignment Problem*

A2.2 Asymptotic Distribution Properties

The importance of the uniqueness assumption lies in the asymptotic properties of the optimal assignment and the ARE estimator (Bhattacharya, 2009). Two key results apply under the uniqueness assumption. First, our estimator is consistent for the optimal assignment function (f^* in Problem 1). Second, our estimator of the impact of the optimal assignment ARE is consistent.

These results are obtained if β identifies the average compliance function up to a constant. This can be obtained by assuming that the assignment is conditionally exogenous:

Assumption 2. $Y_h(c_1, c_2) \perp D_h(c_1, c_2) | X_{h,c_1,c_2,t}$

Where $D_h(c_1, c_2)$ is an indicator for match h, c_1, c_2 and $X_{h,c_1,c_2,t}$ is a vector of observable household and collector characteristics and time dummies. Assumption 2 requires that, conditional on observable characteristics, the status quo assignment is independent of potential compliance $Y_h(c_1, c_2)$.[2] In general matching problems, this assumption is enough to show that the ARE is identified (Graham et al., 2020b). Empirical evidence consistent with Assumption 2 are shown in Table A2 and described in Section 3.

Proposition 1 summarises the main properties of our key estimators.

Proposition 1. *Assume that* $\sqrt{n}\left(\widehat{\beta} - \beta\right) \xrightarrow{d} \mathcal{N}(0, \Sigma)$ *and Assumptions 1–2 hold. Then:*

1. $\widehat{f}^*$ *is consistent to* f^*.
2. $\widehat{ARE}$ *is consistent to* ARE.
3. $\sqrt{n}\left(\widehat{ARE} - ARE\right) \xrightarrow{d} \mathcal{N}(0, (f^* - f^{SQ})'\Sigma(f^* - f^{SQ}))$

[1] For example, if Y is separable in a_1, a_2, and v, all feasible assignment functions yield the same average compliance, and the solution is not unique.

[2] If the assignment were to depend on some unobservable characteristics, we would not be able to identify the expected compliance for counterfactual matches (i.e., those we do not observe in the data). This is critical given that the optimal assignment function requires consistently estimating the expected output for pairs of collectors and households that we do not observe in the data conditional exclusively on their observable types.

The third result states that the sampling error of $\widehat{f}^*$ is asymptotically irrelevant for the estimation of ARE, which relies on $\widehat{f}^* \xrightarrow{p} f^*$ at a faster rate than $\sqrt{n}$ (Bhattacharya, 2009).
Proof:

1. It is exactly the same as proof of Bhattacharya (2009)'s Proposition 1.

2. We denote vectors in bold and scalars in normal font. $ARE = \boldsymbol{Y}(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}})$. Under Assumptions 2 and 3, $\boldsymbol{\beta} + k\mathbf{1} = \boldsymbol{Y}$, with k a constant and $\mathbf{1}$ a vector of 1's. Thus,

$$\begin{aligned} ARE &= \boldsymbol{Y}(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}}) \\ &= (\boldsymbol{\beta} + k\mathbf{1})(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}}) \\ &= \boldsymbol{\beta}(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}}) + k\mathbf{1}\boldsymbol{f}^* - k\mathbf{1}\boldsymbol{f}^{SQ} \end{aligned}$$

Since $\boldsymbol{f}^*$ and $\boldsymbol{f}^{SQ}$ are probability mass functions, they sum to 1 and $k\mathbf{1}(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}}) = 0$. Thus, $ARE = \boldsymbol{\beta}(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}})$. As a result, showing that $\widehat{ARE} \xrightarrow{p} ARE$ is equivalent to showing that

$$\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{f}}^* - \boldsymbol{f}^{\boldsymbol{SQ}}) \xrightarrow{p} \boldsymbol{\beta}(\boldsymbol{f}^* - \boldsymbol{f}^{\boldsymbol{SQ}})$$

which results from $\widehat{\boldsymbol{f}}^* \xrightarrow{p} \boldsymbol{f}^*$ (Proposition 1), $\widehat{\boldsymbol{\beta}}$ converging in probability to $\boldsymbol{\beta}$ (by assumption), and the fact that the limit of the multiplications of two objects is the multiplication of the limit (in probability) of these two objects.

3. The proof is a particular case (assuming uniqueness of the solution of Problem 1) of Bhattacharya (2009). We show the proof for this simpler case and we drop the bold notation for vectors since there is no ambiguity here and by definition

$$\sqrt{n}\left(\widehat{ARE} - ARE\right) = \sqrt{n}\left(\widehat{\beta}\widehat{f}^* - \beta f^*\right) - \sqrt{n}\left(\widehat{\beta}\widehat{f}^{SQ} - \beta f^{SQ}\right)$$

The first term can be written as

$$\sqrt{n}\left(\widehat{\beta}\widehat{f}^* - \beta f^*\right) = f^*\sqrt{n}\left(\widehat{\beta} - \beta\right) 1_{\left[\widehat{f}^* = f^*\right]} + \sqrt{n}\left(\widehat{\beta} - \beta\right)\widehat{f}^* 1_{\left[\widehat{f}^* \neq f^*\right]} + \sqrt{n}\beta\left(\widehat{f}^* - f^*\right) 1_{\left[\widehat{f}^* \neq f^*\right]}$$

where $\sqrt{n}\left(\widehat{\beta} - \beta\right)\widehat{f}^* 1_{\left[\widehat{f}^* \neq f^*\right]}$, is $o_p(1)$ (i.e., converges in probability to zero) since $\widehat{f}^*$ is bounded (it is a probability mass function), and $\left(\widehat{\beta} - \beta\right)\widehat{f}^*$ and $\sqrt{n}1_{\left[\widehat{f}^* \neq f^*\right]}$ are $o_p(1)$ (see Corollary 1 in Bhattacharya (2009)). Similarly, $\sqrt{n}\beta\left(\widehat{f}^* - f^*\right) 1_{\left[\widehat{f}^* \neq f^*\right]}$ is also $o_p(1)$ since $\widehat{f}^* - f^*$ is bounded (both are probability mass functions), β is not a random vector (and is finite), and $\beta\left(\widehat{f}^* - f^*\right)$ and $\sqrt{n}1_{\left[\widehat{f}^* \neq f^*\right]}$ are $o_p(1)$ (see Corollary 1 in Bhattacharya (2009)). Ignoring $o_p(1)$ terms, we thus have

$$\sqrt{n}\left(\widehat{\beta}\widehat{f}^* - \beta f^*\right) = f^*\sqrt{n}\left(\widehat{\beta} - \beta\right) 1_{\left[\widehat{f}^* = f^*\right]}$$

By Item 1 of Proposition 1, $1_{\left[\widehat{f}^* = f^*\right]}$ converges in probability to 1 and can be ignored when deriving the asymptotic distribution. Therefore, $\sqrt{n}\left(\widehat{\beta}\widehat{f}^{*}-\beta f^{*}\right)\xrightarrow{d}\mathcal{N}(0,(f^{*})'\Sigma f^{*})$.

The second term can be written as

$$\sqrt{n}\left(\widehat{\beta}f^{SQ}-\beta f^{SQ}\right)=f^{SQ}\sqrt{n}\left(\widehat{\beta}-\beta\right)$$

and by definition $\sqrt{n}\left(\widehat{\beta}-\beta\right)\xrightarrow{d}\mathcal{N}(0,\Sigma)$, so $\sqrt{n}\left(\widehat{\beta}f^{SQ}-\beta f^{SQ}\right)\xrightarrow{d}\mathcal{N}(0,(f^{SQ})'\Sigma f^{SQ})$.

Combining these two results, we have $\sqrt{n}\left(\widehat{ARE}-ARE\right)=\sqrt{n}\left(\widehat{\beta}\widehat{f}^{*}-\beta f^{*}\right)-\sqrt{n}\left(\widehat{\beta}\widehat{f}^{SQ}-\beta f^{SQ}\right)$, so $\sqrt{n}\left(\widehat{ARE}-ARE\right)\xrightarrow{d}\mathcal{N}(0,(f^{*}-f^{SQ})'\Sigma(f^{*}-f^{SQ}))\quad\square$.

A3 Optimal Number of Collector Types

We rely on unsupervised machine learning methods to shed light on the optimal number of collector types in our context. Specifically, we apply several clustering validation methods to identify the optimal number of collector types for k-means clustering (Lloyd, 1982). We use these methods in a collector-level dataset containing the average (*i*) tax compliance (i.e., the fraction of owners who paid the property tax), (*ii*) tax revenue (i.e., the average amount of property taxes paid per owner), (*iii*) extensive margin tax visits (i.e., the fraction of property owners visited), and (*iv*) intensive margin tax visits (i.e., the number of tax visits per property owner). For each collector, measures (*i*)–(*iv*) are computed across all neighborhoods assigned to a collector during the 2018 property tax campaign.

One of the most popular cluster optimization methods is the "elbow method" (Thorndike, 1953). It involves running k-means clustering and calculating the sum of squared errors (SSE) for a range of values of k. The SSE can be defined as:

$$SSE=\sum_{j=1}^{k}\sum_{i\in C_j}(x(i)-\bar{x}_j)^2$$

where $x(i)-\bar{x}_i$ is the distance between point i and $\bar{x}_i$, the predicted center of point i's cluster, C_j. The SSE measures the sum of the squared distances between each observation and the predicted cluster center. The optimal number of clusters is given by the "elbow" of the relationship between the SSE and the number of clusters k. The optimal number of collector types — i.e., the "elbow" for the collector-level data — appears to be equal to three for tax compliance and revenue (Figure A19, Panel A1) and two for extensive and intensive margin tax visits (Figure A19, Panel A2).

The "silhouette method" (Rousseeuw, 1987) involves running k-means clustering and calculating the silhouette coefficient for a range of values of k. The silhouette coefficient of observation i is defined as:

$$S(i)=\frac{b(i)-a(i)}{max\{a(i),b(i)\}}$$

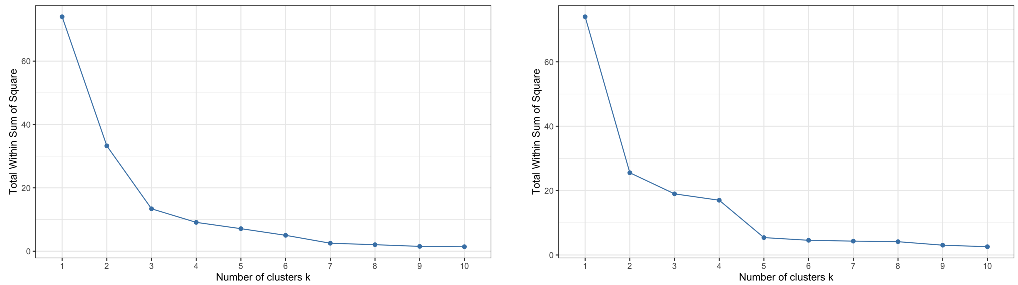
where $b(i)$ is the smallest average distance of point i to all points in any cluster and $a(i)$ is the average distance of i from all other points in its cluster. The silhouette value of point i measures how similar point i is to its own cluster relative to other clusters. The silhouette coefficient of the dataset is the average of the silhouette coefficient of the individual points in the data. The optimal number of clusters according to the silhouette method is then given by the global maximum of the silhouette coefficient. According to the silhouette method, the optimal number of collector types is equal to two when focusing on tax compliance and revenue (Figure A19, Panel B1) or intensive and extensive margin tax visits (Figure A19, Panel B2).

Additionally, we compute the optimal number of clusters for 30 indices that aim at identifying the optimal number of clusters for k-means clustering using the *NbClust* R package. Charrad et al. (2014) provides the list of 30 indices used by the *NbClust* package. Across these indices, the number of collector types that appears to be most frequently optimal is two for both for tax (Figure A19, Panel C1) and visits (Figure A19, Panel C2) outcomes.

Figure A19: Optimal Number of Collector Types

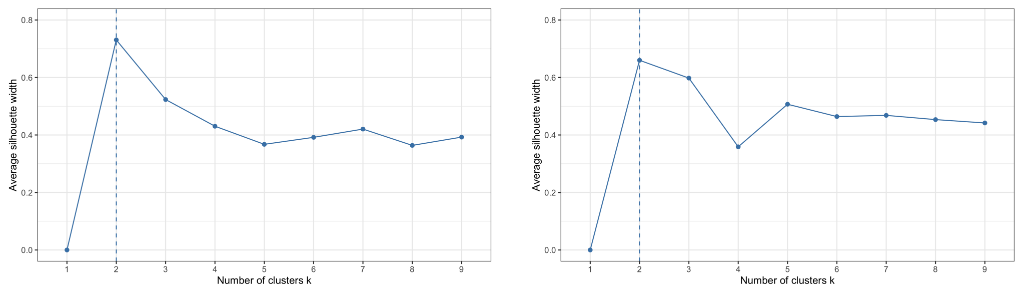
Panel A: Elbow Method

A1: Tax Outcomes A2: Visit Outcomes



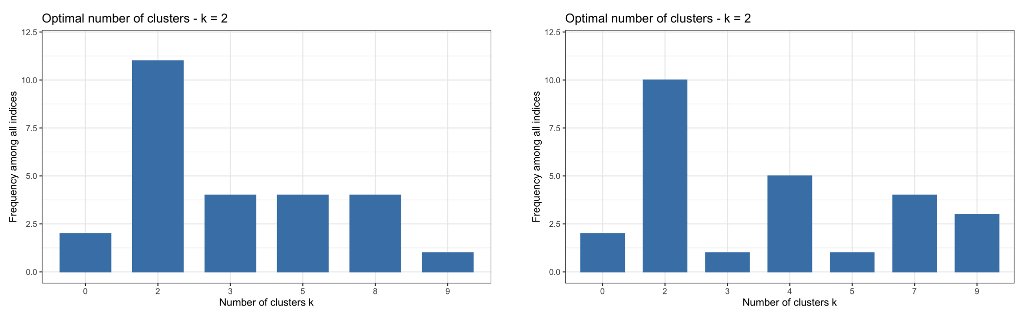
Panel B: Silhouette Method

B1: Tax Outcomes B2: Visit Outcomes



Panel C: 30 Methods

C1: Tax Outcomes C2: Visit Outcomes



Notes: This figure reports the optimal number of collector types for k-means clustering using several methods. Panel A presents the results of the "elbow method" by reporting the sum of squared errors (y-axis) for different number of types (x-axis). Panel B presents the results of the "silhouette method" by showing the silhouette coefficient (y-axis) for different number of types (x-axis). Panel C presents the optimal number of types according to the 30 methods used by the *NbClust* package (Charrad et al., 2014). It also specifies the optimal number of types that arose with the highest frequency across the 30 methods. Panels A1, B1, and C1 present results when considering average tax compliance and revenue by collector as the outcome. Panels A2, B2, and C2 present results when considering extensive and intensive margin average tax visits as the outcome. We discuss these results in Section 6.2.

A4 Monte Carlo Simulations

This section uses Monte Carlo simulations to evaluate to what extent potential misspecifications of the econometric model in equation (6) affect the accuracy of the categorization of tax collectors as high or low type when using the approach described in Section 6.2.

For each Monte Carlo simulation, we report results for 1,000 simulated datasets. Each simulated dataset mimics our tax data structure. First, we assume that 34 tax collectors are involved in the tax campaign. We also assume that we know each collector's "true type" and that there 17 low-type and 17 high-type collectors. Second, we assume that tax collectors are randomly assigned to a new teammate and randomly assigned to work in two neighborhoods every month. We also assume that this monthly reassignment happens over six months (the duration of the tax campaign), thus resulting in 204 neighborhoods per simulated dataset. We evaluate several families of distributions to simulate neighborhood-level tax compliance from: Normal, Uniform, Exponential, Logistic, Beta, Log-Normal, and Gamma. According to distribution tests (see Figure A20 for the Cullen and Frey (1999) plot, also known as the Pearson plot), the distribution of tax compliance at the neighborhood level might belong to the Beta, Gamma, or Log-Normal distribution, and we perform separate Monte Carlo simulations for these three families of distribution.

We first consider simulations that assume that the average tax compliance function exhibits non-linearities in collector type (i.e., the econometric model in equation (6) is misspecified). In our context the average compliance function exhibits complementarities in collector types (Figure 1) and we can therefore draw from the distribution — within a given family distributions (Beta, Gamma, Log-Normal) — that fits the neighborhood-level tax compliance data best for each type of collector pair (*L-L*, *L-H*, *H-H*), which we identify using moment matching estimation. We then construct a household-level simulated dataset by assuming that each neighborhood comprises 127 households (the average number of households per neighborhood in the data) and by creating an individual-level tax compliance indicator such that for each neighborhood, the average tax compliance matches the compliance in the neighborhood-level simulation.

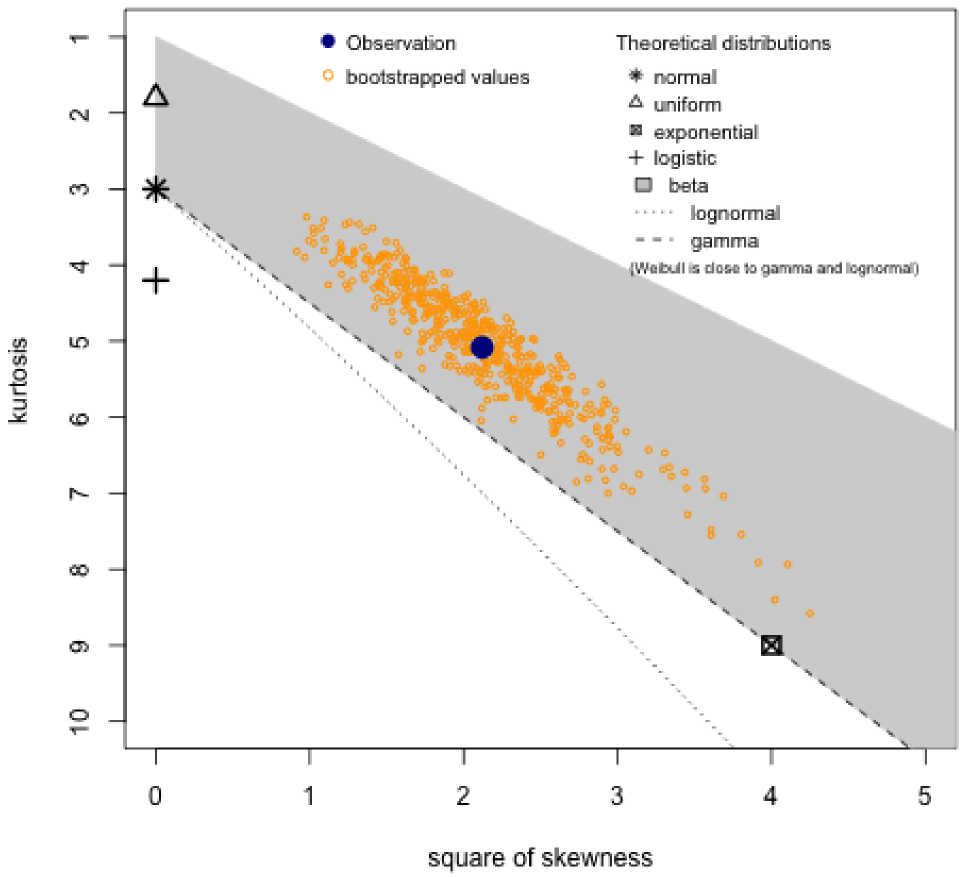
We then turn to simulations that assume that the average tax compliance function is linear in collector type (i.e., the econometric model in equation (6) is not misspecified). For *L-L* and *H-H* collector pairs, we simulate the neighborhood-level tax compliance from the distribution — within a given family of distribution (Beta, Gamma, Log-Normal) — that fits the neighborhood-level tax compliance data best for each type of collector pair (*L-L*, *L-H*, *H-H*), which we identify using moment matching estimation. To obtain linearity in collector type, we simulate the neighborhood-level tax compliance for *L-H* collector by identifying and sampling from the distribution — within a given family of distribution (Beta, Gamma, Log-Normal) — that has a mean equal to the mid-point between the average tax compliance for *L-L* and *H-H* pairs and a variance equal to the sample variance in tax compliance across *L-H* pairs. We then follow the procedure described above to construct a household-level simulated dataset from the neighborhood-level simulated dataset.

For each simulation, we estimate the $\hat{\alpha}_c$ coefficients in equation (6) and estimate the type of each collector using the rank of the $\hat{\alpha}_c$ coefficients, denoted $r_c = rank(\hat{\alpha_c})/N_c$. The

17 collectors characterized by $r_c < 0.5$ (i.e., ranked below median) are deemed low-type, while the 17 collectors characterized by $r_c > 0.5$ (i.e., ranked above median) are deemed high-type. For each collector we can then compare their estimated type with their true type. More specifically, we report the percentage of collectors that are misclassified (i.e., whose estimated type differs from their true type) across the 1,000 simulated datasets constituting each Monte Carlo simulation.

We present the results in Table A13. Column 1 shows the percentage of collector misclassifications when assuming that the average tax compliance function exhibits complementarities in collector type (i.e., the econometric model in equation (6) is misspecified) when simulating data for each type of collector pairs using Beta (Panel A), Gamma (Panel B), or Log-Normal (Panel C) distributions. We find that the percentage of collectors that are misclassified is 18.771% when simulating data using Beta distributions, 18.518% when using Gamma distributions, and 18.040% when using Log-Normal distributions. Column 2 reports the percentage of collector misclassifications when we assume that the average tax compliance function is linear in collector type (i.e., the econometric model in equation (6) is not misspecified). We find that the percentage of collectors that are misclassified is 18.106% when simulating data for each type of collector pair using Beta distributions, 18.059% when using Gamma distributions, and 17.507% when using Log-Normal distributions. The results show that the percentage of collectors that are misclassified is always higher when the average tax compliance function exhibits complementarities in collector type than when it is linear. However, the corresponding difference in the percentage of collectors that are misclassified, reported in Column 3 of Table A13, is small: 0.665% when simulating data using Beta distributions, 0.459% when using Gamma distributions, and 0.533% when using Log-Normal distributions. The Monte Carlo simulation results thus provide reassuring evidence that potential mispecifications of the econometric model in equation (6) are unlikely to have large effects on the estimated collector types and on the estimated average tax compliance function and impact of implementing the optimal assignment.

Figure A20: Distribution Fitting – Skewness and Kurtosis Plot



Notes: This figure reports a skewness-kurtosis graph (Cullen and Frey, 1999) to help choose the distribution that fits the neighborhood-level tax compliance data best. The skewness and kurtosis of the neighborhood-level tax compliance data are represented by the blue circle while the orange hollow circles represent the skewness and kurtosis of 500 bootstrap samples of the neighborhood-level tax compliance data, drawn with replacement. The figure also shows the skewness and kurtosis of several candidate distributions: normal, uniform, exponential, logistic, beta, lognormal, gamma. Some of the distributions (uniform, normal, logistic, and exponential) have only one possible value for the skewness and kurtosis, while others (lognormal, gamma, and beta) have areas of possible values, presented as lines or areas. We discuss the results in Section 6.2.

Table A13: Monte Carlo Simulations

Outcome:	Percentage of Type misclassification (i.e., estimated type differs from true type)		
	Average tax compliance function exhibits complementarities in collector type (equation (6) is misspecified)	Average tax compliance function is linear in collector type (equation (6) is not misspecified)	Difference
	(1)	(2)	(3)
Panel A: Beta Distribution Fit			
Misclassification (in%)	18.771 %	18.106%	0.665%
	[18.356 %–19.186%]	[17.697%–18.515%]	
Panel B: Gamma Distribution Fit			
Misclassification (in%)	18.518%	18.059%	0.459%
	[18.105%–18.931%]	[17.650%–18.468%]	
Panel C: Log-Normal Distribution Fit			
Misclassification (in%)	18.040%	17.507%	0.533%
	[17.605–18.475]	[17.077–17.937]	

Notes: This table reports the Monte Carlo simulations results. For each Monte Carlo simulation, we report results for 1,000 simulated datasets. Each simulated dataset mimics our tax data: 34 tax collectors are randomly assigned to a new teammate and work in two randomly chosen neighborhoods every month throughout the six-month tax campaign. We assume that we know the collectors' "true type": 17 collectors are low-types and 17 are high-types. We simulate the neighborhood-level tax compliance data from several families of distributions: Beta distributions (Panel A), Gamma distributions (Panel B), and Log-Normal distributions (Panel C). Column 1 assumes that the average tax compliance function is non-linear in collector type (i.e., equation (6) is misspecified). In our context, the average tax compliance function exhibits complementarities in collector types and we can therefore draw from the distribution — within a given family of distribution (Beta, Gamma, Log-Normal) — that fits the neighborhood-level tax compliance data best for each type of collector pair (*L-L*, *L-H*, *H-H*), which we identify using moment matching estimation. Column 2 assumes that the average tax compliance function is linear in collector type (i.e., equation (6) is not misspecified). For *L-L* and *H-H* collector pairs, we simulated the neighborhood-level tax compliance from the distribution — within a given family of distribution (Beta, Gamma, Log-Normal) — that fits the neighborhood-level tax compliance data best for each type of collector pair. To obtain linearity in collector type, we simulate the neighborhood-level tax compliance for *L-H* collector pairs by identifying and sampling from the distribution — within a given family of distribution (Beta, Gamma, Log-Normal) — that has a mean equal to the mid-point between the average tax compliance for *L-L* and *H-H* pairs and a variance equal to the sample variance for *L-H* pairs. For each simulation, we estimate the $\hat{\alpha}_c$ coefficients using equation (6) and obtain the estimated collector types using their rank $r_c = rank(\hat{\alpha_c}/N_c)$. Collectors with $r_c > 0.5$ are deemed high-type, while collectors with $r_c < 0.5$ are deemed low-type. We report the percentage of collector that are missclassified, i.e., the percentage of tax collectors whose estimated type differs from their true type across the 1,000 simulations. We also report the corresponding 95 percent confidence intervals. Finally, Column 3 reports the difference in the percentage of misclassification between Columns 1 and Column 2. We discuss these results in Section 6.2.

A5 Estimation of the Average Tax Compliance Function

When estimating the average compliance function using Equation (7), the coefficients of interest are the $\beta(a_1, a_2, v)$ coefficients. Absent the campaign month dummies, these coefficients define the average tax compliance function $Y(a_1, a_2, v)$. When campaign month dummies are included, $\beta(a_1, a_2, v)$ should be interpreted as a convex combination of $Y(a_1, a_2, v, t) - Y(L, L, l, t)$, where $Y(.)$ is a function of the campaign month t (Abadie and Cattaneo, 2018).[3] To avoid this complication in the notation, we make the additional assumption that the average compliance function is separable in campaign month.

Assumption 3. *The average compliance function* $Y(a_1, a_2, v, t) = Y(a_1, a_2, v) + \lambda(t)$, *where the latter term is an arbitrary function of time.*

A6 Additional Mechanism Tests

This section builds on the discussion of skill and effort mechanisms in Section 7.2 by exploring several additional possible mechanisms that could explain that the average compliance function exhibits complementarities in collector and collector-household type.

Homophily. A possible explanation for complementarities in collector types is performance gains due *H-H* pairs' homophily. Tax collection could for example be enhanced for *H-H* pairs if high-type collectors have an easier time communicating due to they shared background. For homophily to explain complementarities in collector type, we would need to observe that (*i*) similarity between collectors in certain traits is associated with higher tax compliance, and (*ii*) benefits from homophily are more pronounced among *H-H* pairs.[4]

Regarding (*i*), we find relatively few traits for which similarity between tax collectors is associated with higher tax compliance. The only trait where homophily is associated with higher compliance is redistributive preferences (Table A14).[5] Turning to (*ii*), we find little evidence that the relationship between collector similarity and tax collection is more pronounced for *H-H* pairs (relative to *L-H* and *L-L* pairs). Similarity in redistributive preferences, or other traits do not appear to differentially boost compliance for *H-H* pairs (Table A15).[6] Overall, these results suggest that homophily is unlikely to explain the complementarity in collector and collector-household type documented in Section 7.1.

Social Incentives. A related but distinct explanation for complementarities in collector type stems from social incentives: i.e., being paired with a friend or person from the same social network might boost effort and lead to higher tax compliance differentially among high-type collectors (Ashraf and Bandiera, 2018). Social incentives could generate complementarities in collector type if pairing friends together in *H-H* pairs triggers "contagious

[3] Since the vector of coefficients $\boldsymbol{\beta}$ is only identified up to a constant , we define $\beta(L, L, l) = 0$.

[4] We restrict our analysis to high-type households since complementarities in collector types are only present among high-type households (Figure 1).

[5] By contrast, similarity in traits typically associated with homophily — gender, age, and education (Lang, 1986) — are not associated with higher team performance (Table A14, Panel A).

[6] The only exception is gender, for which similarity between teammates is correlated with larger increases in compliance for *H-H* pairs. However, less than 6% of collectors are female and thus the gains to gender similarity in collection are unlikely to explain the complementarities in collector type we observe.

enthusiasm," while pairing friends together in *H-L* or *L-L* pairs triggers an averaging of productivity (conformity) or even generates "contagious malaise" (Bandiera et al., 2010).[7]

Although we do not directly observe social links, we examine several proxies, including whether collectors live in the same neighborhood of Kananga,[8] started collecting taxes in the same campaign month,[9] or share religious denomination.[10] There is marginally significant evidence that *H-H* pairs conduct more tax visits when the collectors are from the same neighborhood but this does not translate into higher compliance (Table A16, Columns 1–2). Being in the same cohort appears to differentially suppress effort for *L-L* (marginally significant), but no clear differences emerge between *H-L* and *H-H* pairs (Columns 3–4). Finally, there is some evidence that church links boost effort and compliance among *H-L* pairs compared to *L-L* pairs, but this does not appear to be the case among *H-H* pairs (Columns 5–6).[11] Thus, while social incentives might matter for tax collection, they are unlikely to explain the complementarities in collector and collector-household type documented in Section 7.1.

Differential exemptions. Another potential explanation is that *L-L* and *L-H* pairs exempt more properties, which then translates into lower levels of tax payments. To investigate this issue, we add exempted properties to the data and estimate tax exemption status by collector and household type (Figure A10). Tax exemption does not appear to exhibit increasing differences in collector or collector-household type and is thus unlikely to explain the complementarities shown in section 7.1.

[7] Again, because complementarities in collector types are only present among high-type households (Figure 1), we restrict our analysis to high-type households.

[8] Which we proxy by the distance between the location of the collectors' homes.

[9] Most collectors began at the start of the tax campaign, but some joined in later months.

[10] Churches are an important nexus of social activity in Kananga, and while we do not observe the precise church in which collectors pray, we do know their religious denomination (e.g., Catholic, Protestant, Pentecostal, etc.).

[11] As we show in Table A14, for other potential proxies for social links (age, tribe, education, and income), similarity in these traits is not associated with higher tax collection performance for *H-H* collector pairs relative to *L-H* and *L-L* pairs when assigned to high-type households (Table A15).

Table A14: Tax Compliance by Similarity in Collector Characteristics

Outcome: Tax Compliance	Col. Similarity Coef. (1)	SE (2)	p-value (3)	Mean Char. (4)	Obs. (5)
Panel A: Demographics					
Female	0.010	0.007	0.163	0.068	4,598
Age	0.014	0.017	0.434	30.527	4,480
Main Tribe	-0.020	0.017	0.251	0.223	4,598
Years of Education	-0.013	0.014	0.362	3.622	4,480
Math Score	0.009	0.011	0.405	-0.111	4,480
Literacy (Tshiluba)	-0.038**	0.016	0.023	0.018	4,480
Literacy (French)	0.005	0.017	0.750	-0.004	4,480
Monthly Income	-0.005	0.018	0.796	172.640	4,598
Possessions	0.002	0.011	0.888	1.731	4,480
Born in Kananga	-0.003	0.013	0.808	0.560	4,598
Panel B: Trust in the Government					
Trust Nat. Gov.	-0.008	0.012	0.514	2.895	4,598
Trust Prov. Gov.	0.007	0.009	0.450	2.920	4,598
Trust Tax Min.	-0.011	0.015	0.487	3.486	4,598
Index	0.009	0.014	0.499	0.065	4,598
Panel C: Perceived Performance of Government					
Prov. Gov. Capacity	0.001	0.013	0.925	0.414	4,598
Prov. Gov. Responsiveness	0.015	0.017	0.389	1.614	4,598
Prov. Gov. Performance	0.000	0.009	0.970	4.476	4,598
Prov. Gov. use of Funds	0.013	0.019	0.499	614.686	4,598
Index	-0.005	0.010	0.618	0.063	4,598
Panel D: Government Connections					
Job through Connections	-0.033***	0.012	0.007	0.285	3,934
Relative work for Prov. Gov.	0.003	0.010	0.237	0.006	4,598
Relative work for Tax Ministry	-0.010	0.013	0.467	0.285	4,598
Index	-0.012	0.013	0.329	0.034	4,480
Panel E: Tax Morale					
Taxes are Important	0.007	0.022	0.745	2.806	4,598
Work of Tax Min. is Important	0.005	0.016	0.757	3.796	4,598
Paid Taxes in the Past	0.002	0.010	0.868	2.095	4,598
Index	0.004	0.017	0.835	0.124	4,598
Panel F: Redistributive Preferences					
Imp. of Progressive Taxes	0.016	0.011	0.167	1.622	4,598
Imp. of Progressive Prop. Taxes	0.021**	0.008	1.179	0.004	4,598
Imp. to Tax Employed	-0.005	0.014	0.696	3.316	4,598
Imp. to Tax Owners	0.010	0.016	0.552	3.099	4,598
Imp. to Tax Owners w. title	-0.011	0.010	0.290	3.334	4,598
Index	0.032***	0.009	0.000	-0.292	4,598

Notes: This table reports the relationship between tax compliance and similarity in individual collectors' characteristics. We regress an indicator for tax compliance on the absolute value of a standardized measure of the difference between each collectors' characteristic reverse-coded to be increasing in similarity, controlling for the value of each individual collector's characteristic within the team. The sample used is only high-type households in the analysis sample. We focus on high-type households since complementarities in collector types are only present among high-type households (Figure 1). Columns 1–3 report the correlation coefficient, standard errors (clustered at the neighborhood level) and the corresponding p-values on the similarity measure ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). Columns 4–5 reports the mean collector characteristics (the average within teams) and number of non-missing observations, respectively. Monthly income (Panel A) is in 1000's of Congolese Francs. We focus on high-type households since complementarities in collector types are only present among high-type households (Figure 1). The variables come from surveys with tax collectors described in Section 4. We discuss these results in Section 7.2.

Table A15: Tax Compliance by Pair Type and Proxies for Social Links By Collector Types

	Measure of Similarity in Collector Characteristics								
Outcome: Tax Compliance	Female (1)	Age (2)	Main Tribe (3)	Born in Kananga (4)	Years Edu. (5)	Mon. Income (6)	Govt Conn. Index (7)	Possess. (8)	Redist. Views Index (9)
Similarity X *H-H* Pair (I)	0.085***	-0.075	-0.057*	0.022	0.034	-0.022	-0.064**	-0.032	-0.002
	(0.015)	(0.055)	(0.034)	(0.033)	(0.039)	(0.037)	(0.025)	(0.038)	(0.034)
Similarity X *L-H* Pair (II)	0.037***	-0.021	-0.026	0.021	0.005	0.015	0.002	0.003	0.014
	(0.010)	(0.019)	(0.016)	(0.020)	(0.017)	(0.019)	(0.015)	(0.014)	(0.015)
Similarity (III)	-0.019**	0.019	0.007	-0.010	-0.006	-0.026**	0.003	-0.012*	0.010
	(0.007)	(0.015)	(0.010)	(0.014)	(0.009)	(0.009)	(0.007)	(0.007)	(0.008)
H-H Pair	0.121**	0.093*	0.122***	0.117**	0.110***	0.100**	0.117***	0.145***	0.118**
	(0.036)	(0.050)	(0.034)	(0.036)	(0.027)	(0.033)	(0.030)	(0.042)	(0.042)
L-H Pair	0.017	0.020	0.013	0.004	0.017	0.014	0.011	0.017	0.007
	(0.017)	(0.019)	(0.017)	(0.021)	(0.020)	(0.017)	(0.018)	(0.017)	(0.018)
p-value Test: (I)=(II)	0.002	0.325	0.370	0.981	0.476	0.342	0.019	0.387	0.636
p-value Test: (I)=(III)	<0.001	0.124	0.096	0.441	0.333	0.925	0.022	0.630	.746
L-L Pair Mean	0.072	0.072	0.072	0.072	0.072	0.072	0.072	0.072	0.072
Observations	4,598	4,480	4,598	4,598	4,480	4,598	4,480	4,480	4,598

Notes: This table reports the relationship between tax compliance and similarity in individual collectors' characteristics interacted with pair type. We regress an indicator for tax compliance on pair types interacted with the absolute value of a standardized measure of the difference between collectors' characteristics, reverse-coded to be increasing in similarity, for proxies of social links. Column titles list the measure of similarity used as a regressor and in interaction terms with pair type indicators. All regressions cluster standard errors at the neighborhood level ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The sample used is only high-type households in the analysis sample. We focus on high-type households since complementarities in collector types are only present among high-type households (Figure 1). Test (I)=(II) reports the p-value from the test that correlation coefficients for Similarity X *H-H* Pair and Similarity X *L-H* Pair are equal. Test (I)=(III) reports the p-value from the test that correlation coefficients for Similarity X *H-H* Pair and Similarity are equal. The *L-L* Pair Mean reports average tax compliance within neighborhoods assigned *L-L* pairs. The variables come from surveys with tax collectors described in Section 4. We discuss these results in Section 7.2.

Table A16: Social Incentives: Collector Home Location, Cohort, and Church by Collector Type

	Measure of Similarity in Collector Characteristics					
	Collector Homes (proximity)		Collector Cohort (same)		Collector Church (same)	
	Compliance (1)	Visited (2)	Compliance (3)	Visited (4)	Compliance (5)	Visited (6)
Similarity X *H-H* Pair (I)	0.023	0.072*	0.073	0.198	0.075	0.068
	(0.028)	(0.042)	(0.088)	(0.158)	(0.108)	(0.206)
Similarity X *L-H* Pair (II)	0.014	0.027	-0.003	0.139	0.134**	0.266**
	(0.010)	(0.038)	(0.043)	(0.106)	(0.043)	(0.082)
Similarity (III)	-0.010	-0.012	0.001	-0.136*	-0.055***	-0.086
	(0.008)	(0.029)	(0.028)	(0.081)	(0.015)	(0.055)
H-H Pair	-0.038	-0.413	0.073	0.083	0.112**	0.133**
	(0.230)	(0.314)	(0.073)	(0.140)	(0.045)	(0.064)
L-H Pair	-0.069	-0.141	0.013	0.031	-0.011	-0.006
	(0.066)	(0.265)	(0.020)	(0.068)	(0.018)	(0.068)
p-value Test: (I)=(II)	0.754	0.247	0.400	0.700	0.607	0.343
p-value Test: (I)=(III)	0.282	0.208	0.475	0.118	0.249	0.500
L-L Pair Mean	0.072	0.357	0.072	0.357	0.072	0.357
Observations	3,415	2,261	4,598	3,116	4,598	3,116

Notes: This table examines if social links among collectors are differentially associated with performance among high-type collectors and high-type households. It considers three proxies for social links: the distance between collectors' home locations in kilometers (Columns 1–2); whether collectors began working on the campaign in the same month (Columns 3–4); and whether collectors belong to the same church (Columns 5–6). In each column, we regress the outcome — tax compliance or visits — on pair types interacted with these measures of social links. The outcome is tax compliance in odd columns and receipt of post-registration visits from collectors in even columns. All regressions cluster standard errors at the neighborhood level ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The sample used is only high-type households in the analysis sample. We focus on high-type households since complementarities in collector types are only present among high-type households (Figure 1). Test (I)=(II) reports the p-value from the test that correlation coefficients for Similarity X *H-H* Pair and Similarity X *L-H* Pair are equal. Test (I)=(III) reports the p-value from the test that correlation coefficients for Similarity X *H-H* Pair and Similarity are equal. The *L-L* Pair Mean reports average tax compliance within neighborhoods assigned *L-L* pairs. The variables come from surveys with tax collectors described in Section 4. We discuss these results in Section 7.2.

A7 Number of Collector Types and Impact of the Optimal Assignment

This section uses our theoretical framework to explore the relationship between the number of collector types and the impact of implementing the optimal assignment policy, as summarized in Section 8.2.[12] We prove that increasing the number of collector types from K to $2K$ would magnify the effect on tax compliance from implementing the optimal assignment.

Setup

For analytical tractability, we assume a continuum of collectors of mass 1. We assume that a collector m has ability q_m, distributed uniformly in $[0,1]$. Collectors work in pairs and the expected tax compliance of a household assigned to collectors m and n is given by:

$$y : [0,1] \times [0,1] \to [0,1] \tag{13}$$

We refer to the expected tax compliance when assigned to tax collectors of ability q_m and q_n as the production function. The production function is symmetrical ($y(q_m, q_n) = y(q_n, q_m)\ \forall q_m, q_n$) and increasing in each of its components

Because we lack the data to estimate the production function for each level of tax collector ability, we divide the collectors into a finite number of types. We find the optimal assignment and analyze its effects on tax compliance (relative to the status quo assignment) using these collector types.

We first define a partition of collectors into K types as the list of k sets given by:

$$Q_k^K = [\frac{k-1}{K}, \frac{k}{K}] \quad \forall k \in \{1, ..., K\} \tag{14}$$

We then define the production function over types as

$$Y^K(Q_m^K, Q_n^K) = K^2 \int_{m-1/K}^{m/K} \int_{n-1/K}^{n/K} y(q_m, q_n) dq_m dq_n \tag{15}$$

An assignment function $m(Q_m^K, Q_n^K)$ is the probability that a pair of type (Q_m^K, Q_n^K) is assigned to a household. It has the following properties:

1. Symmetry: $m(Q_m^K, Q_n^K) = m(Q_n^K, Q_m^K)$.
2. $0 \leq m(Q_m^K, Q_n^K) \leq 1\ \forall Q_m^K, Q_n^K$.
3. $\sum_{i=1}^{K} m(Q_i^K, Q_m^K) + m(Q_m^K, Q_i^K) = 2/K.\ \forall Q_m^K$
4. $\sum_{i=1}^{K} \sum_{j=1}^{K} m(Q_i^K, Q_j^K) = 1$.

[12]We ignore the assignment of collectors to households since it is irrelevant for our argument.

When collectors are partitioned into K types, the optimal assignment function solves the problem

$$m^*(K) = \arg\max_m \sum_{i=1}^{K} \sum_{j=1}^{K} m(Q_i^K, Q_j^K) \tag{16}$$

and the expected tax compliance under the optimal assignment if we partition the collectors into K types is given by

$$Y^*(K) = \sum_{i=1}^{K} \sum_{j=1}^{K} Y^K(Q_i^K, Q_j^K) \cdot m^*(Q_i^K, Q_j^K) \tag{17}$$

Results

We would like to prove that the expected tax compliance under the optimal assignment is (weakly) increasing in the size of the type partition K. Our setup allows us to prove the case stated in the following proposition:

Proposition. If we double the number of types, the expected tax compliance under the optimal assignment is weakly increasing, i.e., for every $K > 0$, $Y^*(2K) \geq Y^*(K)$.

Proof

Consider the optimal assignment function with K types, $m^*(K)$. We show that we can find an assignment function when partitioning the type space into $2K$ types that yields the same expected tax compliance to the optimal assignment function $m^*(K)$.

First, we define the function $g : \mathcal{N} \to \mathcal{N}$ such that

$$g(n) = (n+1)//2$$

where $//$ is the integer division operator.

We then consider the following assignment function with $2K$ types:

$$m(Q_i^{2K}, Q_j^{2K}) = \frac{1}{4} m^*(Q_{g(i)}^K, Q_{g(j)}^K) \quad \forall 1 \leq i, j \leq 2K$$

The assignment function m thus finds the type associated with each of the types Q_i^{2K} and Q_j^{2K} if we had partitioned the type space into K instead of $2K$ types and attributes the same probability as the optimal assignment function would.

First, it is straightforward to show that m satisfies the properties of an assignment function. Below we demonstrate this for Property 4 above. The proofs for Properties 1–3 are also straightforward.

To see that Property 4 is satisfied, note that:

$$
\begin{aligned}
&\sum_{i=1}^{2K}\sum_{j=1}^{2K} m(Q_i^{2K}, Q_j^{2K}) \\
&= \sum_{a=1}^{K}\sum_{b=1}^{K} m(Q_{2a-1}^{2K}, Q_{2b-1}^{2K}) + m(Q_{2a-1}^{2K}, Q_{2b}^{2K}) + m(Q_{2a}^{2K}, Q_{2b-1}^{2K}) + m(Q_{2a}^{2K}, Q_{2b}^{2K}) \\
&= 4 \cdot \sum_{a=1}^{K}\sum_{b=1}^{K} \frac{1}{4} \cdot m^*(Q_a^K, Q_b^K) \\
&= \sum_{a=1}^{K}\sum_{b=1}^{K} \cdot m^*(Q_a^K, Q_b^K) = 1 \quad \square
\end{aligned}
$$

Second, we show that m yields the same expected tax compliance as m^*:

$$
\begin{aligned}
&\sum_{i=1}^{2K}\sum_{j=1}^{2K} m(Q_i^{2K}, Q_j^{2K}) Y^{2K}(Q_i^{2K}, Q_j^{2K}) \\
&= \sum_{a=1}^{K}\sum_{b=1}^{K} m(Q_{2a-1}^{2K}, Q_{2b-1}^{2K}) Y^{2K}(Q_{2a-1}^{2K}, Q_{2b-1}^{2K}) + m(Q_{2a-1}^{2K}, Q_{2b}^{2K}) Y^{2K}(Q_{2a-1}^{2K}, Q_{2b}^{2K}) + \\
&\quad m(Q_{2a}^{2K}, Q_{2b-1}^{2K}) Y^{2K}(Q_{2a}^{2K}, Q_{2b-1}^{2K}) + m(Q_{2a}^{2K}, Q_{2b}^{2K}) Y^{2K}(Q_{2a}^{2K}, Q_{2b}^{2K}) \\
&= \sum_{a=1}^{K}\sum_{b=1}^{K} m^*(Q_a^K, Q_b^K) \Big[Y^{2K}(Q_{2a-1}^{2K}, Q_{2b-1}^{2K}) + Y^{2K}(Q_{2a-1}^{2K}, Q_{2b}^{2K}) + Y^{2K}(Q_{2a}^{2K}, Q_{2b-1}^{2K}) \\
&\qquad + Y^{2K}(Q_{2a}^{2K}, Q_{2b}^{2K}) \\
&= \sum_{a=1}^{K}\sum_{b=1}^{K} m^*(Q_a^K, Q_b^K) Y^K(Q_a^K, Q_b^K) = Y^*(K)
\end{aligned}
$$

where the last line is obtained by linearity of the integral.

In sum, we show that if we partition the type space in $2K$ types, we can find an assignment function that yields the same expected tax compliance as the best assignment function if we partition the type space into K types. Since this is not necessarily the optimal assignment in $2K$ space, we conclude that expected tax compliance $Y^*(2K)$ is at least as high as $Y^*(K)$. $\square$.

A8 Using Bonhomme (2021) "Heterogeneity, Sorting, and Complementarity" Methodology

In this section, we estimate a nonlinear model for tax collector team production using the econometric framework proposed by Bonhomme (2021). This method allows us to document heterogeneity in collector performance and complementarities between tax collectors when only team-level tax compliance is observed. Additionally, it does not involve estimating collector type in a first step and is thus unaffected by potential misspecifications in the estimation of tax collector type (cf. Section 6.2). It does not rely on a split-sample approach and is therefore more powered to estimate the average tax compliance function and the optimal assignment with a higher number of tax collector types (cf. Section 8.2).

In order to estimate the nonlinear model in the presence of unobserved worker heterogeneity in a team setting, we follow Bonhomme (2021) and rely on a finite mixture model, where the distribution of the discrete tax collector type is modeled using a random-effects approach.[13] To estimate the nonlinear random-effects model we again follow Bonhomme (2021) by using a mean-field variational method.[14]

Figures A21 and A22 document the patterns of heterogeneity and complementarity in the nonlinear model estimated for 2–6 collector types. For conciseness, we only comment on the results for 2 and 3 collector types (Figure A21). The type proportions are $46.67\%, 53.33\%$ for two types and $43.11\%, 31.99\%, 24.89\%$ for three types,[15] and the matrices below show the average tax compliance distribution (in percent) with two and three collector types and confirm that tax collectors have heterogeneous productivity levels:

$$\begin{pmatrix} 5.86 & 7.33 \\ 7.33 & 15.22 \end{pmatrix}$$

$$\begin{pmatrix} 7.84 & 5.69 & 7.95 \\ 5.69 & 4.64 & 14.26 \\ 7.95 & 14.26 & 29.28 \end{pmatrix}$$

The implications of the estimated model for heterogeneity and complementarity are shown graphically for two and three collector types in Figure A21. Panel B suggests the presence of complementarity, since with two (three) collector types the return of two high types working together is 1.21 (2.76) standard deviations above the return of two low types working together. To assess how complementarities might affect the allocation of collectors, Panel C reports the allocation that maximizes total tax compliance while keeping the marginal distribution of collector types the same as under the status quo assignment. The results confirm that the positive assortative assignment by collector type would be optimal. Figure A22 shows similar results for 4–6 collector types, although the results are noisier due to the small number of observations for each collector match type (Figure A22, Panel A).[16]

Table A17 reports the decomposition of the tax compliance variance for 2–6 collector types. In the nonlinear case, the total tax compliance variance (Table A17, row 1) has four

[13] We use random effects because the estimates of collector types are not sufficiently precise to follow a grouped fixed-effects approach as in Bonhomme et al. (2019).

[14] The presence of unobserved collector heterogeneity in a network of collector pairs makes the estimation of a nonlinear random-effects model challenging since the same collector may participate in multiple pairs, and pairs contain multiple collectors. Variational estimators are widely used in networks and other complex data settings (Bishop, 2006; Blei et al., 2017).

[15] The type proportions are 36.18%, 20.40%, 21.33%, 22.09% for four types, 30.76%, 21.13%, 5.56%, 23.08%, 19.47% for five types, and 11,69%, 22.14%, 21.70%, 8.32%, 14.28%,21.87% for six types.

[16] With five (six) collector types, each type combination is estimated from 0.01–0.09 (0.01–0.07) of the total sample, i.e. from 2 to 16 (2 to 13) neighborhoods (Figure A22, Panel A2–A3).

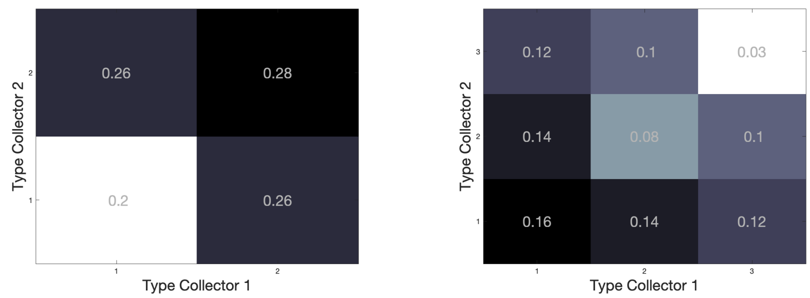
components, again following Bonhomme (2021). The "heterogeneity" component (Table A17, row 2) reflects the variation in collector effects on output. The "sorting" component (Table A17, row 3) reflects the variance contribution due to team composition not being random. As described in Section 3, tax collectors are randomly assigned to teams monthly and the sorting component is therefore equal to zero in our context. The "nonlinearities" component (Table A17, row 4) reflects interaction effects between tax collectors, above and beyond the additive effects of tax collector types. The remainder of the variance is attributed to other factors (Table A17, row 5).

The results presented in Table A17 show that heterogeneity explains a large fraction of the total variance: between 20.86% and 33.80% depending on the number of collector types. Similarly, nonlinearities explain a substantial fraction of the variance in tax compliance: between 3.14% and 21.69% depending on the number of collector types. This is considerably higher than the share of the variance explained by complementarities in other contexts analyzed by Bonhomme (2021) such as complementarities between economic researchers (0.84%–2.41%) or inventors (3.42%–7.85%).

Figure A21: Nonlinear Model Estimates and Optimal Allocation of Tax Collectors

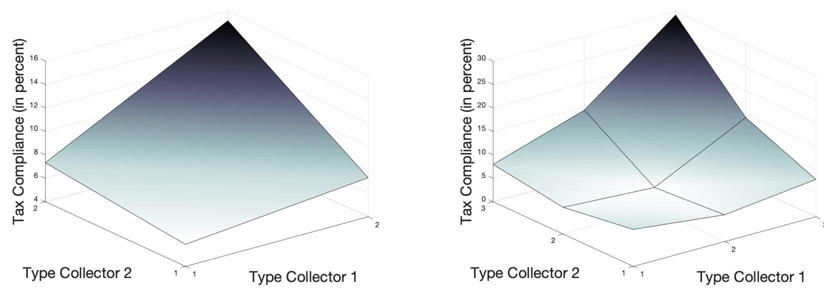
Panel A: Number of Neighborhoods by Collector Type

A1. Nb of Coll. Types = 2 A2. Nb of Coll. Types = 3



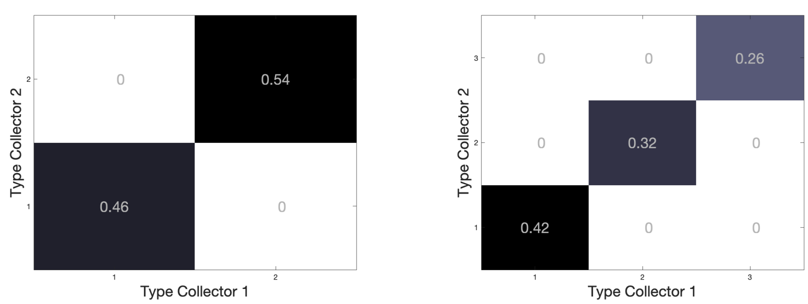
Panel B: Heterogeneity and Complementarity

B1. Nb of Coll. Types = 2 B2. Nb of Coll. Types = 3



Panel C: Optimal Assignment

C1. Nb of Coll. Types = 2 C2. Nb of Coll. Types = 3

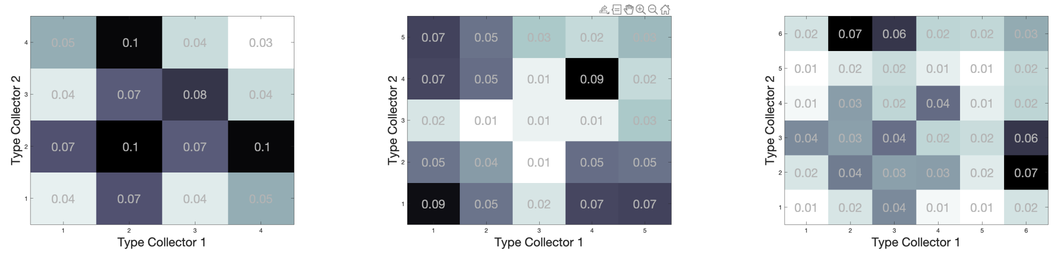


Notes: This figure reports estimates of the finite mixture model estimated using random effects following Bonhomme (2021). Panel A shows the proportion of each combination of types in the data. Panel B shows the average tax compliance (in percent) for different combinations of collector types. Panel C shows the proportion of collector types in the optimal allocation. The first figure of each panel (A1, B1, C1) shows the results with two collector types. The second figure of each panel (A2, B2, C2) shows the results with three collector types. We discuss these results in Section 8.2.

Figure A22: Nonlinear Model Estimates and Optimal Allocation of Tax Collectors

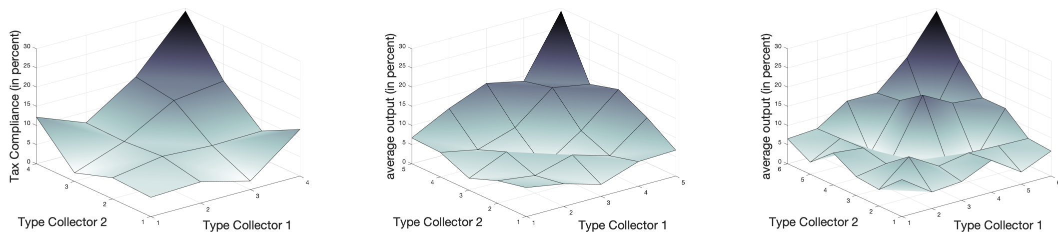
Panel A: Number of Neighborhoods by Collector Type

A1. Nb of Coll. Types $= 4$ A2. Nb of Coll. Types $= 5$ A3. Nb of Coll. Types $= 6$



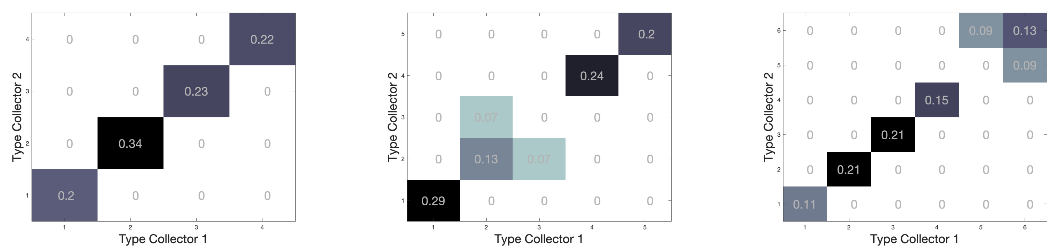
Panel B: Heterogeneity and Complementarity

B1. Nb of Coll. Types $= 4$ B2. Nb of Coll. Types $= 5$ B3. Nb of Coll. Types $= 6$



Panel C: Optimal Assignment

C1. Nb of Coll. Types $= 4$ C2. Nb of Coll. Types $= 5$ C3. Nb of Coll. Types $= 6$



Notes: This figure reports estimates of the finite mixture model estimated using random effects following Bonhomme (2021). Panel A shows the proportion of each combination of types in the data. Panel B shows the average tax compliance (in percent) for different combinations of collector types. Panel C shows the proportion of collector types in the optimal allocation. The first figure of each panel (A1, B1, C1) reports the results with four collector types. The second figure of each panel (A2, B2, C2) reports the results with five collector types. The third figure of each panel (A3, B3, C3) reports the results with five collector types. We discuss these results in Section 8.2.

Table A17: Variance Decomposition with Nonlinear Production Function in Tax Collector Type

	Collector Types $= 2$		Collector Types$= 3$		Collector Types $= 4$		Collector Types $= 5$		Collector Types $= 6$	
	Variance	%	Variance	%	Variance	%	Variance	%	Variance	%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Total variance	58.88		58.88		58.88		58.88		58.88	
Heterogeneity	12.29	20.86%	17.49	29.71%	17.97	30.52%	19.90	33.80%	18.77	31.87%
Sorting	0	0%	0	0%	0	0%	0	0%	0	0%
Nonlinearities	1.85	3.14%	6.32	10.73%	9.21	15.65%	8.51	14.45%	12.77	21.69%
Other factors	44.75	76.00%	35.07	59.56%	31.70	53.83%	30.48	51.76%	27.34	46.44%

Notes: This table reports estimates of variance components from a nonlinear model in collector type estimated following Bonhomme (2021). We report results for different number of collector types: two (Columns 1–2), three (Columns 3–4), four (Columns 5–6), five (Columns 7–8), six (Columns 9–10). Columns 1, 3, 5, 7, and 9 report the total variance (row 1) and its components: heterogeneity (row 2), sorting, which is always equal to zero given the random assignment of collectors to teammates (row 3), nonlinearities (row 4), and other components (row 5). Columns 2, 4, 6, 8, and 10 report the percentage of the total variance represented by each component. We discuss these results in Section 6.2 and 8.2.

A9 Neighborhood-Level Optimal Assignment

The neighborhood-level optimal assignment f^* can then be defined as:

$$f^* \equiv \arg\max_f \sum_{n \in N} \sum_{a_1, a_2 \in \{L,H\}^2} f(a_1, a_2, n) \overline{Y}_n(a_1, a_2)$$

$$\sum_{a_1, a_2 \in \{L,H\}^2} f(a_1, a_2, n) = 1 \qquad \forall n \in N$$

$$\sum_{n \in N} \left[2f(a, a, n) + \sum_{a' \neq a} \left(f(a', a, n) + f(a, a', n) \right) \right] = N^{nbh} \qquad \forall a \in \{L, H\}$$

As in Problem 1, the objective function is the expected tax compliance under assignment f, but we now consider the average tax compliance over all neighborhoods N:

$$\overline{Y}_n(a_1, a_2) = \frac{N_n(l)\widehat{\beta}(a_1, a_2, l) + N_n(h)\widehat{\beta}(a_1, a_2, h)}{N_n(l) + N_n(h)}$$

with $N_n(l)$ and $N_n(h)$ the number of low- and high-type households in neighborhood n.[17]

The constraints are analogous to those in Problem 1. The first constraint imposes that all neighborhoods are assigned to one pair of collector (i.e., the probability that a neighborhood is assigned to one pair of collectors equals one.) The second constraint imposes that tax collectors of each type are assigned to the same number of neighborhoods as under the status quo assignment.

[17] We exclude 6 neighborhoods with less than 10 observations from the analysis.

An alternative possible objective function to the expected tax compliance $\overline{Y}_n(a_1, a_2)$ is the expected number of tax payers, $N_n\overline{Y}_n(a_1, a_2)$:

$$N_n\overline{Y}_n(a_1, a_2) = N_n(l)\widehat{\beta}(a_1, a_2, l) + N_n(h)\widehat{\beta}(a_1, a_2, h)$$

which would imply the following objective function:

$$\sum_{n \in N} \sum_{a_1, a_2 \in \{L,H\}^2} N_n f(a_1, a_2, n)\overline{Y}_n(a_1, a_2)$$

and would allow the government to assign high-type pairs to neighborhoods with a large number of households, increasing the number of households assigned to high-type collectors in comparison to the status quo assignment.

Whether the outcome of interest is average compliance, $\overline{Y}_n(a_1, a_2)$, or the expected number of tax payers, $N_n\overline{Y}_n(a_1, a_2)$, the impact of the optimal assignment function, relative to the status quo assignment, is given by

$$\sum_{n \in N} \sum_{a_1, a_2 \in \{L,H\}^2} N_n\overline{Y}_n(a_1, a_2) \left[f^*(a_1, a_2, n) - f^{SQ}(a_1, a_2, n) \right]$$

where $f^{SQ}(a_1, a_2, n) = 1/4$ for all $a_1, a_2 \in \{L, H\}^2$.

A10 Endogenous Responses to Implementing the Optimal Policy

Throughout the analysis, we assume that the average tax compliance function would be unaffected by changes in the assignment function.[18] This assumption is essential for the implementation of the optimal policy to have the effects documented in Sections 8.1-8.2. To see this, let's assume that the average compliance function depends on the assignment function f and is denoted $Y(a_1, a_2, v_h, f)$. Unless $Y(a_1, a_2, v_h, f^{SQ}) = Y(a_1, a_2, v_h, f^*)$, the tax compliance achieved under the status quo assignment function f^{SQ}

$$\sum_{a_1, a_2, v_h} f^*(a_1, a_2, v_h)Y(a_1, a_2, v_h, f^{SQ})$$

would differ from the tax compliance achieved under f^*

$$\sum_{a_1, a_2, v_h} f^*(a_1, a_2, v_h)Y(a_1, a_2, v_h, f^*)$$

In our context, changes in the assignment function could affect the average tax compliance function through changes in collectors' effort or in their opportunities for learning. We explore both possibilities below.

[18] This assumption is known as the stable unit treatment value assumption (SUTVA) in the impact evaluation literature.

A10.1 Endogenous Effort Provision

A10.1.1 Endogenous Effort due to Time Constraints

A first source of concern is that changing collectors' assignment could impact effort levels by match types, which would impact the average tax compliance function and result in $Y(a_1, a_2, v_h, f^{SQ}) \neq Y(a_1, a_2, v_h, f^*)$. Endogenous effort could affect the average tax compliance function if collectors target high-type households for tax visits and are time-constrained, i.e., are unable to do all the tax visits that would have a positive return during the month-long campaign period. Under these conditions, implementing the optimal assignment could lead to lower visit levels and lower tax compliance for (H, H, h) match types than observed under the random assignment.

To see this, consider the simplified case where there are four households in Kananga, two low-types (v^L) and two high-types (v^H). Additionally assume that there are two collector teams, a low-type team (a^{L-L}) and a high-type team (a^{H-H}), each assigned to two households. Finally, assume that collector teams are time-constrained and can only visit one of the two households they are assigned to. We assume that the probability of household h paying the property tax is $\Pr(y_h = 1) = e_{p,h} v_h a^p$, where $e_{p,h}$ indicates whether collector pair p visited household h after registration. Under the status quo assignment, each collector pair is assigned to a low-type and a high-type household. Since $v^H > v^L$, both collectors choose to visit the high-type household.[19] Tax compliance under the status quo assignment would thus be $v^H a^{H-H} + v^H a^{L-L}$. Under the optimal assignment, high-type households would be assigned to the high-type team and low-type households would be assigned to the low-type team because $a^{H-H} > a^{L-L}$. Due to time-constraints, the high-type team would only visit one of the high-type households and the low-type team would only visit one of the low-type households and tax compliance would be $v^H a^{H-H} + v^L a^{L-L}$, which is strictly lower than the compliance achieved under the status quo assignment since $v^H > v^L$. By contrast, if collectors were not time-constrained, compliance under the optimal assignment would be $2v^H a^{H-H} + 2v^L a^{L-L}$ which would be strictly higher than the compliance achieved under the status quo assignment $(v^H + v^L)a^{H-H} + (v^H + v^L)a^{L-L}$ since $v^H > v^L$ and $a^{H-H} > a^{L-L}$.

We first investigate if in our context tax collectors target high-type households for tax visits. Examining heterogeneity in post-registration collector visits by household type, we do find evidence that collectors target high-type households for tax visits (Figure A8).[20] We then investigate whether tax collectors are time-constrained in our context. We first examine the distribution of tax payments over the month-long tax collection period in each neighborhood.[21] If collectors were time-constrained, the marginal value of an additional visit should be larger than its marginal cost at the end of the month and we would expect a steady stream of tax payments until the end of the tax collection period. However, the data reveal that

[19] For example if they face financial or promotion incentives based on performance.

[20] This is especially the case for *L-L* teams, which are 8 percentage points more likely to visit high- than low-type households ($p = 0.045$). By contrast, *H-H* teams are 5 percentage points more likely to visit high- than low-type households ($p = 0.17$).

[21] The month-long collection periods were staggered throughout the experiment and did not systematically coincide with calendar months.

tax payments across neighborhoods are on average close to zero on the last few days of the tax collection period (Figure A23, Panel A), suggesting that the marginal value of visits at the end of the tax collection period is on average very small.[22] Second, if collectors were time-constrained, they should visit a lower fraction of households when assigned to a larger neighborhood.[23] However, we find no significant relationship between neighborhood size and proportion of households visited (Figure A23 Panel B).[24] Taken together, these results suggest that changes in collector effort by match type resulting from tax collectors' time constraints are unlikely to result in changes in the average tax compliance function when the assignment function changes.

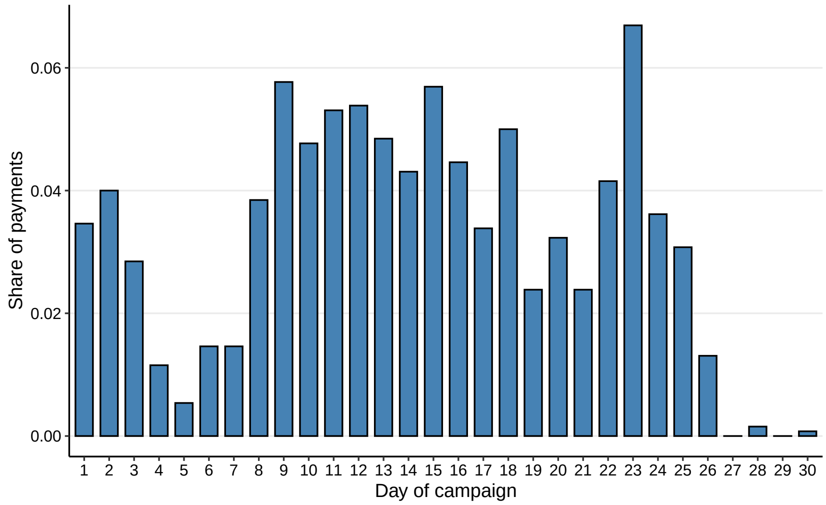
[22]This is unlikely to be explained by collector fatigue given that their activity jumps sharply immediately following the assignment to new neighborhoods in the next campaign month.

[23]Collector characteristics are orthogonal to neighborhood size due to the random assignment of collectors to neighborhoods.

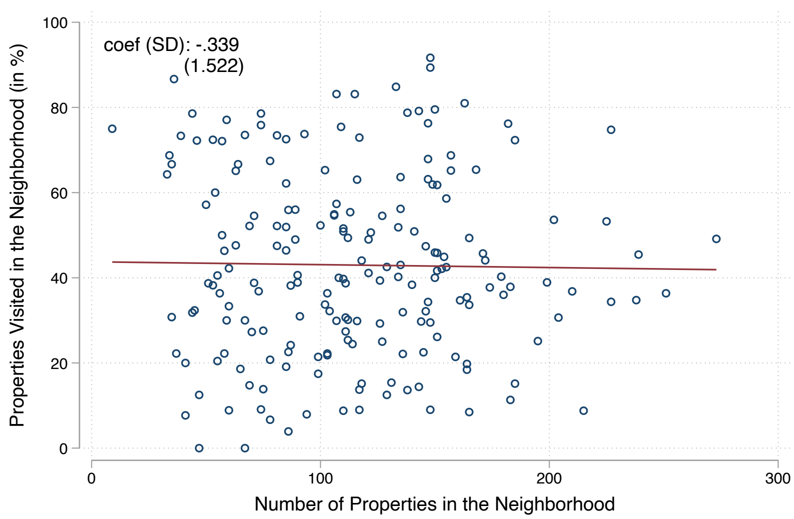
[24]A one standard deviation increase in the number of households (50 households) in a neighborhood has a small and insignificant effect on the likelihood of being visited (0.3 percentage points, $p = 0.8$).

Figure A23: Tax Collectors' Time Constraints

Panel A: Distribution of Tax Payments over Time



Panel B: Visits as a Function of Neighborhood Size



Notes: This figure investigates whether tax collectors experienced various forms of time constraints when collecting taxes in Kananga. Panel A shows the distribution of tax payments across the days of the month-long tax collection period across all neighborhoods. Day 1 corresponds to the first day of the month-long tax collection period across all neighborhoods and day 30 to the last day of the month-long tax collection period across all neighborhoods. The month-long collection periods were staggered throughout the experiment and did not systematically coincide with calendar months. Panel B shows the relationship between the size of the neighborhoods (i.e., the number of properties) and the fraction of households visited by the tax collectors in the neighborhood. Panel B, also reports the coefficient and robust standard errors of a neighborhood-level regression of the percentage of properties visited after registration on the standardized number of properties. We discuss these results in Section 8.3

A10.1.2 Endogenous Effort due to Demoralization

Endogenous effort could also affect the average tax compliance function if assigning low-type collectors to low-type teammates and low-type households — as in the optimal assignment — demoralizes them and lead to lower effort and tax compliance for (L, L, l) match types than observed under random assignment.[25]

We explore this possibility by analyzing whether the exogenous variation in collectors' assignments to low-type teammates and households during the 2018 campaign affected collectors' motivation measured at endline. Drawing on the psychology literature (Tremblay et al., 2009), the endline collector survey asked to what extent collectors were motivated in their work by (*i*) extrinsic motivation (i.e., due to financial compensation), (*ii*) intrinsic motivation (i.e., due to the fulfilling nature of the job), (*iii*) introjection (i.e., due to a positive self-image from the work), or (*iv*) goal orientation (i.e., due to the social importance of the work). We compute standardized indices for each motivation type based on the corresponding set of questions. We then estimate the correlation of collectors' endline motivation with the share of low-type teammates they were assigned to during the tax campaign (Table A18) and the share of low-type households they were assigned to during the campaign (Table A19). While we find that low-type collectors exhibited lower levels of motivation at endline (Table A18 and A19, Column 1), there is no evidence that being exogenously exposed to a higher fraction of low-type teammates or low-type households during the campaign undermined collectors' motivation (Table A18 and A19, Column 2) for low- or high-type collectors (Table A18 and A19, Column 3).[26] Overall, these results run counter to the low-type collector demoralization story.

We also investigate a more extreme form of demoralization, namely the possibility that low-type tax collectors could drop out of the campaign entirely under the optimal assignment (Table A20). We find no evidence that low-type collectors are more likely to drop out (Column 1) or that being exogenously exposed to a higher fraction of low-type teammates or low-type households during the campaign is associated with a higher probability of dropping out (Columns 2 and 4) for low- or high-type collectors (Columns 3 and 5).[27] Thus, according to available evidence, it appears unlikely that the assignment of low-type collectors to low-types teammates or households under the optimal assignment would trigger demoralization and reduce low-type collector pairs' effort levels compared to the random assignment.

Nonetheless, for completeness, we examine how the effect of the optimal policy would vary if low-type collectors were to become so demoralized under the optimal assignment that they drop out from the tax campaign. Specifically, we assume that low-type collectors cease their work on the tax campaign immediately (on day 1) and permanently, thereby contributing zero revenue to the state. Figure A24 investigates tax compliance under the optimal assign-

[25] While we assume that collectors' financial incentives (piece-rate performance-based wages) would remain the same under the optimal assignment, it is possible that low-type collectors anticipate lower group productivity under the optimal assignment, which could lower their motivation.

[26] If anything, low-type collectors' motivation levels appear to have been *less* impacted than high-type collectors by assignment to low-type teammates and households (Table A18 and A19, Column 3).

[27] If anything, Column 3 of Table A20 suggests that low-type collectors are less likely to drop out from the tax campaign than high-type collectors when assigned to low-type teammates.

ment when a fraction of low-type collectors drop out relative to tax compliance under the status quo assignment. As expected, a higher fraction of low-type collectors dropping out is associated with a lower effect of implementing the optimal assignment on tax compliance. That said, the estimated effect remains positive and significant at the 5% level for dropout rates below 25% and at the 10% level for dropout rates below 50%.[28] Thus, our results suggest that the optimal assignment would outperform the status quo even for high rates of collector dropout. As a benchmark, only three tax collectors in our sample (8.82%) did not complete the full 2018 tax campaign.[29]

[28]For dropout rates above 50%, the estimated impact of the optimal assignment is still positive but not statistically different from zero at conventional significance levels.

[29]Moreover, Figure A24 assumes that low-type collectors drop out on day 1 before they collect any revenue. However, in practice low-type collectors would likely work for a few months before becoming demoralized and dropping out. For the three collectors in our sample who dropped out of the 2018 tax campaign, two worked for two months and one worked for four months. If we assume that low-type collectors would work for a few months before dropping out, then Figure A19 underestimates the effect of the optimal assignment policy when a fraction of low-type collectors drop out.

Table A18: Collector Motivation by Teammates Type

	(1)	(2)	(3)
Panel A: Extrinsic Motivation			
Coll. Low-Type	-1.207***		-1.668**
	(0.275)		(0.562)
Frac. Low-Type Teammates		-0.214	-0.201
		(0.555)	(0.584)
Coll. Low-Type X Frac. Low-Type Teammates			0.873
			(0.998)
Panel B: Intrinsic Motivation			
Coll. Low-Type	-0.892**		-1.571**
	(0.311)		(0.661)
Frac. Low-Type Teammates		-0.318	-0.617
		(0.561)	(0.601)
Coll. Low-Type X Frac. Low-Type Teammates			1.335
			(1.182)
Panel C: Introjection			
Coll. Low-Type	-0.787**		-1.041
	(0.319)		(0.803)
Frac. Low-Type Teammates		-0.172	-0.126
		(0.558)	(0.767)
Coll. Low-Type X Frac. Low-Type Teammates			0.483
			(1.293)
Panel D: Goal Orientation			
Coll. Low-Type	-0.714**		-1.520*
	(0.325)		(0.757)
Frac. Low-Type Teammates		0.096	-0.333
		(0.528)	(0.498)
Coll. Low-Type X Frac. Low-Type Teammates			1.522
			(1.247)
Observations	34	34	34

Notes: This table shows the impact of each collector's own type (Column 1), of their teammates' types (Column 2), and their interaction (Column 3) on endline measures of collectors' extrinsic motivation (Panel A), intrinsic motivation (Panel B), introjection (Panel C), and goal orientation (Panel D) in collecting taxes during the 2018 property tax campaign. Each outcome variable is a standardized index for each motivation type. Column 1 reports the effect of collector's own type on motivation by regressing motivation on an indicator for the collector being low-type. Column 2 reports the effect of collectors' teammates type on motivation by regressing the motivation outcomes on the fraction of each collector's teammates that were low-type during the tax campaign. Column 3 studies heterogeneity by collector type in the effect of their teammates' type on motivation. It regresses the motivation outcome on collector type, the fraction of each collector's teammates that were low-type during the tax campaign, and the interaction of both variables. We report robust standard errors ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The sample size is reported at the bottom of the table. We discuss these results in Section 8.3.

Table A19: Collector Motivation By Household Assignment Type

	(1)	(2)	(3)
Panel A: Extrinsic Motivation			
Coll. Low-Type	-1.207***		-1.353
	(0.275)		(0.974)
Frac. Low-Type Households		-2.029	-2.716*
		(1.842)	(1.571)
Coll. Low-Type X Frac. Low-Type Households			0.365
			(3.106)
Panel B: Intrinsic Motivation			
Coll. Low-Type	-0.892**		-0.716
	(0.311)		(1.052)
Frac. Low-Type Households		-1.690	-1.810
		(1.436)	(1.703)
Coll. Low-Type X Frac. Low-Type Households			-0.630
			(3.300)
Panel C: Introjection			
Coll. Low-Type	-0.787**		-1.050
	(0.319)		(1.076)
Frac. Low-Type Households		-2.250	-2.915**
		(1.404)	(1.227)
Coll. Low-Type X Frac. Low-Type Households			0.731
			(3.478)
Panel D: Goal Orientation			
Coll. Low-Type	-0.714**		-0.921
	(0.325)		(1.204)
Frac. Low-Type Households		-1.313	-1.881
		(1.600)	(1.114)
Coll. Low-Type X Frac. Low-Type Households			0.589
			(4.006)
Observations	34	34	34

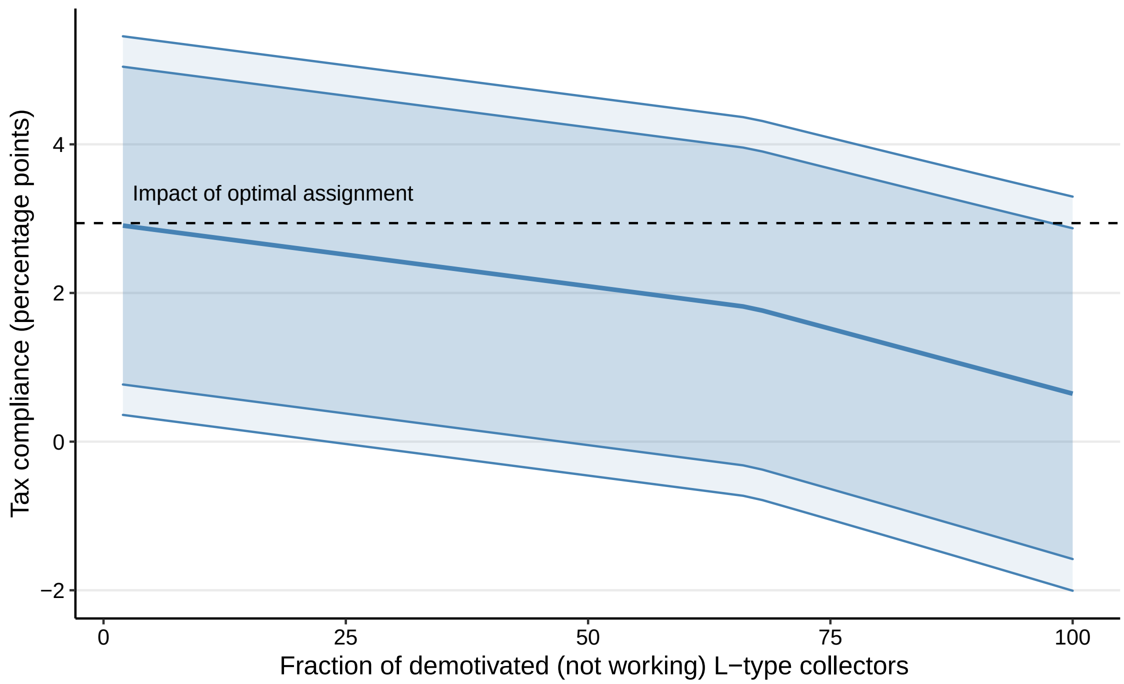
Notes: This table shows the impact of each collector's own type (Column 1), of the household type they were assigned to (Column 2), and their interaction (Column 3) on endline measures of collectors' extrinsic motivation (Panel A), intrinsic motivation (Panel B), introjection (Panel C), and goal orientation (Panel D) in collecting taxes during the 2018 property tax campaign. Each outcome variable is a standardized index for each motivation type. Column 1 reports the effect of collectors' own type on motivation by regressing motivation on an indicator for the collector being low-type. Column 2 reports the effect of the household type they collected from on motivation by regressing the motivation outcomes on the fraction of each collector's assignment that were low-type households during the tax campaign. Column 3 studies heterogeneity by collector type in the effect of the household type they collected from on motivation. It regresses the motivation outcome on collector type, the fraction of each collector's assignment that were low-type households during the tax campaign, and the interaction of both variables. We report robust standard errors ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The sample size is reported at the bottom of the table. We discuss these results in Section 8.3.

Table A20: Collector Dropout By Teammate Type and Household Assignment Type

	(1)	(2)	(3)	(4)	(5)
Coll. Low-Type	0.059		0.590**		0.108
	(0.100)		(0.279)		(0.405)
Frac. Low-Type Teammates		-0.006	0.431		
		(0.293)	(0.345)		
Coll. Low-Type X Frac. Low-Type Teammates			-1.037**		
			(0.502)		
Frac. Low-Type Households				-0.725	-0.626
				(0.480)	(0.681)
Coll. Low-Type X Frac. Low-Type Households					-0.180
					(0.980)
Observations	34	34	34	34	34
Mean	0.088	0.088	0.088	0.088	0.088

Notes: This table shows the impact of each collector's own type (Column 1), of their teammates' types (Column 2), of the interaction between collectors' own type and their teammates' types (Column 3), of the household type they were assigned to (Column 4), and the interaction between collectors' own type and the household type they were assigned to (Column 5) on an indicator for not completing the entire property tax campaign (i.e., "dropping out"). We report robust standard errors ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The sample size is reported at the bottom of the table. We discuss these results in Section 8.3.

Figure A24: Effects of the Optimal Assignment when a fraction of Low-Type Collectors Drop Out



Notes: This figure shows the potential impact of low-type tax collectors dropping out of the tax campaign (x-axis) on the effect of the optimal assignment on tax compliance relative to the status quo assignment (y-axis). We assume that collectors who drop out of the tax campaign stop working immediately and entirely (they collect no property taxes) and are not replaced by any other tax collector. Collector types are estimated using a fixed effects model described in Section 6.2. The shaded areas in dark blue represent the 90% confidence interval while the one in light blue represents the 95% confidence interval. Standard errors use bootstrap re-sampling (100 samples) at the neighborhood level. The dashed black horizontal line indicates the impact of the optimal assignment policy on tax compliance with no low-type collector dropout and when collector types are estimated using a fixed effects model, corresponding to the impact reported in Column 1 of Table 1. The kink represents the point in which all low-type households are exhausted and then high-type households are matched to $L-L$ teams. We discuss these results in Section 8.3.

A10.2 Endogenous Learning Dynamics

A10.2.1 Learning from Teammates.

Endogenous learning could also affect the average tax compliance function if collectors learn tax collection skills from their teammates — e.g., high-type collectors might increase their teammates' performance by sharing skills and knowledge useful for tax collection, such as techniques for convincing households to pay — and low-type collectors learn more tax collection skills than high-type collectors when assigned to high-type teammates.[30] Under these conditions, positive assortative matching on the collector-collector dimension would lead to low levels of learning from teammates among low-type collectors, which might lead us to overestimate the impact of the optimal assignment.

To investigate this possibility, we exploit the random assignment of collectors into different pairs over the course of the tax campaign. Specifically, we first estimate whether past assignment to a high-type teammate affects tax collectors' subsequent performance by estimating the following equation:[31]

$$y_{h,n,t} = \delta \cdot \mathrm{E}_{c_1(n),c_2(n),t} + \lambda_t + \varepsilon_{h,n,t} \tag{18}$$

where h, n, and t index household, neighborhood, and tax campaign month, respectively. $y_{h,n,t}$ is the tax compliance decision of household h, and $\mathrm{E}_{c_1(n),c_2(n),t}$ captures collector $c_1(n)$ and $c_2(n)$'s exposure to high-type collectors prior to campaign month t. λ_t are campaign month fixed effects. Standard errors are clustered at the neighborhood level. The coefficient of interest is δ, which captures whether the productivity of collector pairs in campaign month t is affected by past exposure to high-type teammates.

We use several measures of past exposure to high-type teammates. The first measure captures collector c's exposure to high-type teammates during past campaign month l. Formally, it is defined by:

$$\text{Exposure}_{c,t}(l) = \sum_{c' \in C} 1_{[a_{c'}=H]} \cdot 1_{[m_c(t-l)=c']} \tag{19}$$

where $1_{[c'=m_c(t-l)]}$ is an indicator for tax collectors c' and c being teammates in tax campaign month $t-l$ and $1_{[a_{c'}=H]}$ is an indicator for collector c' being high-type. A second measure examines cumulative exposure to high-type teammates in all campaign months prior to month t. Formally, it is defined as:

$$\text{Exposure}_{c,t} = \frac{1}{t-t_c^0} \sum_{l=1}^{t-t_c^0} \text{Exposure}_{c,t}(l) \tag{20}$$

where t_c^0 is the first time period of tax collection for collector c. For ease of interpretation, we standardize this measure and the estimates should be interpreted as the effect of a one standard deviation change in cumulative past exposure to high-type teammates.

We use these measures to estimate the OLS regression specifications given by Equation

[30] Learning tax collection skills might be more pronounced when paired with a high-type teammate than a low-type one because they have more skills to transfer or because they are viewed as higher prestige individuals and thus their partners are more attentive to them (e.g., Bursztyn et al., 2014).

[31] One challenge when studying skill transmission is that we do not separately observe the contribution of each collector to the team's output, but rather observe tax compliance at the team level. As a consequence, we cannot directly test whether collector c's average tax compliance increases when assigned to a high-type collector during the campaign months when both collectors work together. Instead, we can test whether the teams collector c is a part of in subsequent periods are characterized by higher compliance after c was assigned to a high-type teammate.

(18). This equation relies on measuring exposure to high-type collectors prior to campaign month t, $\text{E}_{c_1(n),c_2(n),t}$, which is defined by one of the following two equations:

$$\text{E}_{c_1(n),c_2(n),t}(l) = \text{Exposure}_{c_1(n),t}(l) + \text{Exposure}_{c_2(n),t}(l) \tag{21}$$

$$\text{E}_{c_1(n),c_2(n),t} = \text{Exposure}_{c_1(n),t} + \text{Exposure}_{c_2(n),t} \tag{22}$$

depending on whether past exposure to high-type teammates is defined using $\text{Exposure}_{c,t}(l)$ or $\text{Exposure}_{c,t}$.[32] We estimate collector types in the holdout sample, and we estimate equation (18) in the analysis sample, described in Section 3.[33]

We find evidence of learning from high-type teammates (Table A21, Columns 1–3 and 6–8). A one standard deviation increase in cumulative past exposure to high-type teammates increases subsequent tax compliance by 3.53 percentage points $(p = 0.03)$ (Column 1) and tax revenue by 83.02 CF $(p = 0.02)$ (Column 6). Similarly, being assigned to a high-type teammate during the previous tax campaign month increases subsequent tax compliance by 2.34 percentage points $(p = 0.15)$ (Column 2) and tax revenue by 50.56 CF $(p = 0.18)$ (Column 7). The results are weaker for the effect of being assigned to a high-type teammate in an earlier campaign month (Columns 3 and 8).

These results suggest that collectors learn tax collection skills from high-type teammates. However, for learning from teammates to impact the average tax compliance function and the impact of the optimal assignment, it would have to affect collectors of different types differently. To see this, consider the expected tax compliance of household h in campaign month t when assigned to collectors of type a_1 and a_2:

$$\mathbb{E}\left[y_{ht}|a_1, a_2\right] = m(a_1, a_2) + \left[l(a_1) + l(a_2)\right] \tag{23}$$

where $m(a_1, a_2)$ is the expected effect on compliance of an assignment to collectors of type a_1 and a_2 absent any learning. The additional effect of learning is captured by $l(a_1) + l(a_2)$, where $l(a)$ is the *expected* impact of what collector a has learned prior to campaign month t on tax compliance in month t, y_{ht}. The expectation is taken over the teammates collector a is assigned to under assignment function f.[34] We define the learning function of a collector of

[32] Most, but not all, collectors started working in the first month of the tax campaign. When campaign month t is the first period of tax collection for collector c_1, we calculate $\text{E}_{c_1(n),c_2(n),t}(l)$ as $2 \times \text{Exposure}_{c_2(n),t}(l)$ and vice-versa for collector c_2. When campaign month t is the first period of tax collection for both collectors, we exclude the observation from the regression. As a consequence the data from the first period of tax collection are excluded from the estimation of Equations (18) and (25).

[33] When estimating learning from teammates, we might overestimate the ability of collector c's past teammates when c is high-type. We would then mechanically find that past assignment to high-type teammates is associated with high tax compliance.

[34] Because we are now considering dynamics, this assignment function also depends on tax campaign month t. However, we restrict the assignment function to be identical at every t. For the particular type of average tax compliance in Equation (23), this restriction is harmless, since accounting for dynamics cannot improve over a static assignment.

type a as

$$l(a) = \sum_{a' \in A} g(a') f(a'|a) \tag{24}$$

where $g(a')$ is the effect on tax compliance of being assigned to a teammate of type a' in collection month $t-1$. The likelihood that a type-a collector is assigned to a type-a' collector is $f(a'|a)$ where f the assignment function. $l(a)$ is the expected impact on collector type a of learning from a collector type a' in the previous period. If learning takes the form described in Equations (23) and (24), then Proposition 2 states that learning does not affect the difference in average compliance under two assignment functions that keep the composition of the workforce constant.

Proposition 2. *Assume that $\mathbb{E}[y_{ht}|a_1, a_2]$ takes the form defined in Equations (23) and (24). Consider two assignment functions $f^1(a_1, a_2)$ and $f^2(a_1, a_2)$ such that the marginal distributions of type $f^1(a) = f^2(a)$. Then the difference in average tax compliance under the two assignment functions is given by*

$$\sum_{a_1, a_2 \in A^2} m(a_1, a_2) \left(f^1(a_1, a_2) - f^2(a_1, a_2) \right)$$

Proof:

The average tax compliance for the assignment function f is given by

$$\mathbb{E}[y_{ht}|f] = \sum_{a_1, a_2 \in A^2} f(a_1, a_2) m(a_1, a_2) \; + \sum_{a_1, a_2 \in A^2} f(a_1, a_2)[l(a_1) + l(a_2)]$$

where

$$\begin{aligned}
\sum_{a_1, a_2 \in A^2} f(a_1, a_2) l(a_1) &= \sum_{a_1 \in A} f(a_1) l(a_1) \\
&= \sum_{a_1 \in A} f(a_1) \sum_{a' \in A} g(a') f(a'|a_1) \\
&= \sum_{a_1 \in A} \sum_{a' \in A} g(a') f(a'|a_1) f(a_1) \\
&= \sum_{a' \in A} \sum_{a_1 \in A} g(a') f(a_1, a') \\
&= \sum_{a' \in A} g(a') f(a')
\end{aligned}$$

and as a result

$$\mathbb{E}[y_{ht}|f] = \sum_{a_1, a_2 \in A^2} f(a_1, a_2) m(a_1, a_2) \; + 2 \sum_{a' \in A} g(a') f(a')$$

The difference in average tax compliance between assignment functions f^1 and f^2 is

$$\mathbb{E}[y_{ht}|f_1] - \mathbb{E}[y_{ht}|f_2] = \sum_{a_1,a_2 \in A^2} f^1(a_1,a_2)m(a_1,a_2) \; - f^2(a_1,a_2)m(a_1,a_2)$$

since $2\sum_{a' \in A} g(a')f^1(a') = 2\sum_{a' \in A} g(a')f^2(a')$ when $f^1(a') = f^2(a') \; \forall a'$. $\square$

One scenario where Proposition 2 would not hold is if learning depends on collector type.[35] In particular, if low-type collectors were better learners than high-type collectors (e.g., because they have more to learn), then the results presented in Section 8 would *overestimate* the true effect of optimal matching by ignoring learning effects. Conversely, if high-type collectors were the better learners (e.g., because they are more open to learning from their peers), our results would *underestimate* the true effect of optimal matching by ignoring learning effects.

We provide evidence on whether learning from high-type teammates is more pronounced for low-type or high-type collectors by estimating the following equation:

$$\begin{aligned} y_{h,n,t} = \gamma_1 \mathrm{E}_{c_1(n),c_2(n),t} \cdot HH_{c_1(n),c_2(n)} + \gamma_2 \mathrm{E}_{c_1(n),c_2(n),t} \cdot LH_{c_1(n),c_2(n)} + \\ \delta \mathrm{E}_{c_1(n),c_2(n),t} + \omega_1 HH_{c_1(n),c_2(n)} + \omega_2 LH_{c_1(n),c_2(n)} + \lambda_t + \varepsilon_{h,n,t} \end{aligned} \tag{25}$$

which interacts past exposure to high-type teammates, $\mathrm{E}_{c_1(n)c_2(n)t}$, with indicators for *H-H* and *H-L* collector teams, $HH_{c_1(n),c_2(n)}$ and $LH_{c_1(n),c_2(n)}$, controlling for whether the team is *H-H* or *H-L*. Throughout the analysis, *L-L* teams are the comparison group. The coefficients of interests are γ_1 and γ_2, capturing the additional learning accrued to *H-H* and *H-L* teams (relative to *L-L* teams), respectively.

The results presented in Table A21 do not show evidence that low-type collectors are better than high-type collectors in terms of learning tax collection skills when exposed to high-type teammates in past tax campaign months. If anything, we find suggestive evidence of more pronounced learning among high-type collectors, i.e., $\gamma_1 > 0$, across measures of past exposure to high-type teammates. As mentioned above, if high-type collectors are better at learning from high-type teammates than high-type collectors, our results would *underestimate* the true effect of optimal matching by ignoring learning effects. However, the γ_1 coefficients reported in Table A21 are not statistically significant at conventional levels, making the results only suggestive.

[35] Additionally, Proposition 2 would not hold if learning is not separable, i.e. if $[l(a_1) + l(a_2)]$ is replaced by $l(a_1, a_2)$ in Equation (23). We cannot directly test whether learning is separable, but separability is a standard assumption in the peer effects literature (e.g., Todd and Wolpin, 2003; Burke and Sass, 2013).

Table A21: Learning from High-Type Teammates

	Tax Compliance					Tax Revenue				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cumulative High-Type Exposure	3.53			2.51		83.02			69.77	
	(1.66)			(1.31)		(36.75)			(29.67)	
	[0.03]			[0.05]		[0.02]			[0.02]	
High-Type Exposure Lag 1		2.34	3.41		2.52		50.56	71.70		41.19
		(1.62)	(2.00)		(1.47)		(37.39)	(48.15)		(32.15)
		[0.15]	[0.09]		[0.09]		[0.18]	[0.14]		[0.20]
High-Type Exposure Lag 2			0.40					22.26		
			(0.92)					(19.94)		
			[0.66]					[0.26]		
Cumulative High-Type Exposure × HH				5.90					167.89	
				(7.52)					(170.57)	
				[0.43]					[0.33]	
Cumulative High-Type Exposure × LH				-38.05					-36.53	
				(2.32)					(48.82)	
				[0.69]					[0.44]	
High-Type Exposure Lag 1 × HH					2.13					91.28
					(4.62)					(104.39)
					[0.64]					[0.38]
High-Type Exposure Lag 1 × LH					-2.58					-51.63
					(2.00)					(43.55)
					[0.20]					[0.24]
Mean	7.92	7.92	6.54	7.92	7.92	236.00	236.00	212.62	236.00	236.00
Observations (Holdout Sample)	11,732	11,732	11,732	11,732	11,732	11,732	11,732	11,732	11,732	11,732
Obervations (Analysis Sample)	7,665	7,665	5,166	7,665	7,665	7,665	7,665	5,166	7,665	7,665

Notes: This table shows the impact of past exposure to high-type teammates on collectors' current tax collection performance, measured by a property tax compliance indicator in Columns 1–5 and by property tax revenue per property owner (in Congolese Francs) in Columns 6–10. The tax compliance outcome in Columns 1–5 is multiplied by 100, and the coefficients can be interpreted as percentage point changes. Columns 1–3 and 6–7 report estimates from equation (18), using the cumulative high-type exposure measure (Columns 1 and 6), one high-type exposure lag (Columns 2 and 7), or two high-type exposure lags (Columns 3 and 8). Columns 4–5 and 9–10 estimate equation (25), using the cumulative high-type exposure measure interacted with indicators for the type of the tax collectors' pair (Columns 4 and 9) and the first lag exposure measure interacted with indicators for the type of the tax collectors' pair (Columns 5 and 10). Standard errors are clustered at the neighborhood level and presented in parenthesis while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average tax compliance and the sample sizes are reported at the bottom of the table. We discuss these results in Section 8.3.

A10.2.2 Learning-by-doing.

Endogenous learning could also affect the average tax compliance function if collectors learn tax collection skills over time (i.e., learning-by-doing) and low-type collectors learn more tax collection skills than high-type collectors when assigned to a high-type household. Then, positive assortative matching on the collector-household dimension would lead to low levels of learning-by-doing among low-type collectors, which would lead us to overestimate the impact of the optimal assignment.

To test for learning-by-doing, we analyze the relationship between tax compliance in month t and the number of households assigned to collector teams involving collector c in previous months, which we denote $X_{c,t-1}$. Formally, we estimate the regression:

$$y_{hnt} = \gamma \left(X_{c_1(n),t-1} + X_{c_2(n),t-1} \right) + \lambda_t + \varepsilon_{hnt} \tag{26}$$

where $c_1(n)$ and $c_2(n)$ are functions indicating the collectors assigned to neighborhood n and λ_t is a vector of campaign month fixed effects. If learning-by-doing is important $\gamma > 0$ since more opportunities to learn (i.e., more past assignments) should be associated with better tax collector performance. The coefficient γ is unbiased given that collectors were randomly assigned to neighborhoods of different size, as described in Section 3.

We find limited evidence of learning-by-doing in this context. If anything, increasing the number of past assignments by 1 SD *decreases* tax compliance by 1.58 percentage points (Table A22, Column 1), although the estimate is not significant at conventional levels ($p = 0.10$). This could suggest that a higher number of assignments causes exhaustion rather than learning. However, collectors assigned to a larger number of assignments in previous campaign months do not appear to reduce their tax collection effort level, as proxied by an indicator for being visited by tax collectors ($p = 0.91$, Column 4) or the number of visits by tax collectors ($p = 0.94$, Column 7).[36] We find similar results when analyzing the relationship between tax compliance or visits in month t and the number of households assigned to teams involving collector c in the previous month $t-1$ (Columns 2, 5, 8) or in the two previous months $t-1$ and $t-2$ (Columns 3, 6, 9). Taken together, these results suggest a limited role for learning-by-doing in our setting.

[36]The negative coefficient in Column 1 is thus more likely to reflect exogenous decreases in households' compliance behavior over time, rather than collectors exerting less effort. As discussed in Balan et al. (2022), tax compliance decreased over the course of the 2018 tax campaign due to increasing discontent with the incumbent president Joseph Kabila, who was ousted in a contentious election just after the tax campaign ended.

Table A22: Learning-by-doing

	Tax compliance			Visit Indicator			Number of Visits		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past Nbhd Assignments	-1.584			0.425			-0.005		
Cumulative	(0.969)			(3.886)			(0.059)		
	[0.102]			[0.913]			[0.938]		
Past Nbhd Assignments		-0.345	-1.379		2.681	2.171		0.021	0.016
Lag 1		(0.880)	(0.975)		(1.604)	(1.712)		(0.025)	(0.038)
		[0.695]	[0.157]		[0.095]	[0.205]		[0.405]	[0.575]
Past Nbhd Assignments			-0.046			-1.372			0.000
Lag 2			(0.475)			(3.204)			(0.038)
			[0.924]			[0.534]			[0.998]
Mean	6.369	6.369	5.644	37.175	37.175	36.518	0.492	0.492	0.488
Observations	15,733	15,733	11,782	10,359	10,359	7,840	10,357	10,357	7,839

Notes: This table explores the relationship between tax collectors' performance and their number of assignments in the previous campaign months. We consider three outcomes: an indicator for tax compliance by the owner (Columns 1–3), an indicator for receiving a post-registration visit (Columns 4–6), and the number of post registration visits (Columns 7–9). In Columns 1, 4 and 7, we report results from equation (26) by estimating the relationship between the outcome of interest and the number of assignments received by each collector in the pair during all the previous tax campaign months. In Columns 2, 5, and 8, we show the relationship between the outcome of interest and the number of assignments received by each collector in the pair in the previous tax campaign month ($t-1$). In Columns 3, 6, and 9, we report the relationship between the outcome of interest and the number of assignments received by each collector in the pair in the previous tax campaign month ($t-1$) and the month prior ($t-2$). All regressions include campaign months fixed effects. We standardize the explanatory variable. We multiply the tax compliance and visit indicators by 100 and estimates for these variables are thus expressed in percentage points. Standard errors are clustered at the neighborhood level and presented in parentheses while the corresponding p-values are presented in brackets ($* = p < 0.1, ** = p < 0.05, *** = p < 0.01$). The average for each outcome is reported at the bottom of the table, which also report the corresponding sample size. We discuss these results in Section 8.3.

A11 Effects on Secondary Outcomes

This section explores in more detail the effects of implementing the optimal assignment policy on bribes, payment of other formal and informal taxes, and views of the government.

A11.1 Bribe Payments

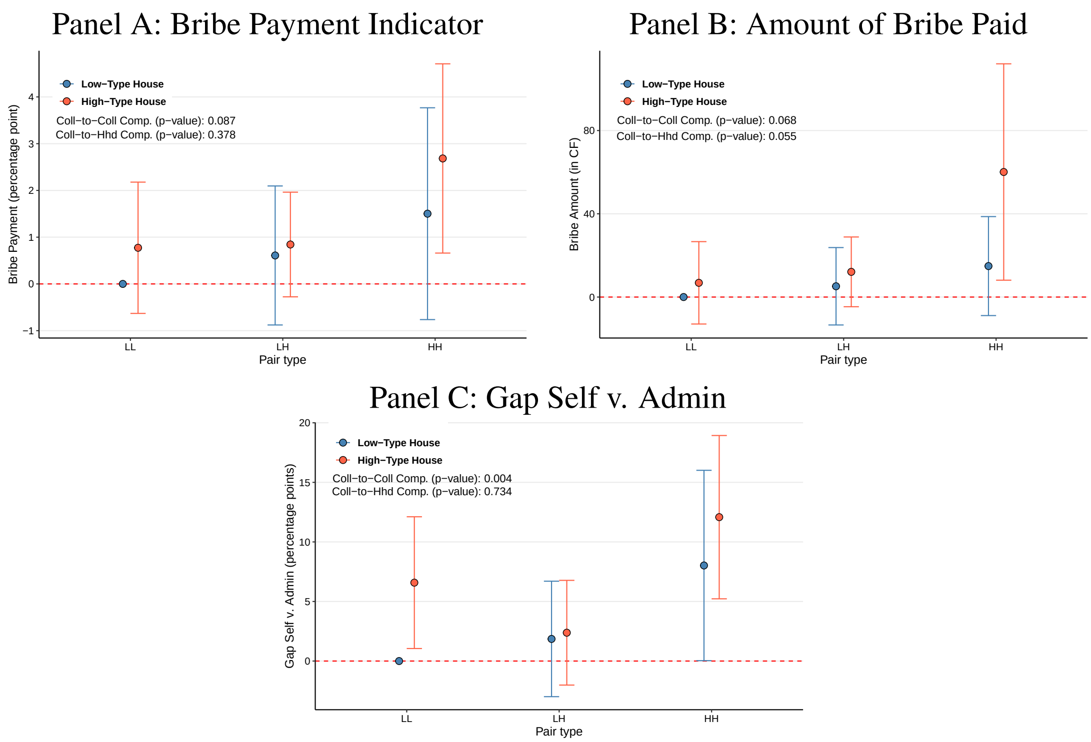
We test if the optimal assignment would impact bribery using three survey-based measures of bribes. First, households reported in the midline survey if they paid the "transport" of the collectors — a local code for bribes — and if so, how much they paid. Though self-reported, this bribe measure has been validated in past work in this same context.[37] Implementing the optimal assignment policy does not appear to significantly increase bribe payment on the extensive margin, though the coefficient is positive: 0.387 percentage points, $p = 0.268$ (Panel A of Table 2, Row 1). However, we find suggestive evidence of an increase of 13.896 CF ($p = 0.098$) — a 46% increase — in the amount of bribes paid per owner (Panel A of Table 2, Row 2). We find similar, albeit slightly larger, increases in amounts of bribes paid when the government aims at maximizing tax revenue per owner (Table A8, Column 2) and much smaller effects on bribe payments when the government's objective is to maximize tax revenues net of the amount of bribes paid per owner (Table A8, Column 4).

As a second measure, we consider the gap between administrative tax data and citizen self-reports of payment at midline. Although it likely picks up social desirability responses, this measure may capture instances in which a citizen unwittingly paid a bribe or the collector simply pocketed the tax money without printing a receipt. According to this measure, the optimal assignment policy would increase bribe payments on the extensive margin by 2.253 percentage points ($p = 0.059$), a 24% increase (Panel A of Table 2, Row 3).

On net, we find suggestive evidence that the optimal assignment would slightly increase bribe payments. This increase reflects complementarities in collector type rather than complementarities in collector-household type (Figure A25). Complementarity tests confirm that the average bribe payment function exhibits complementarities in collector type when measuring bribes using the bribe payment indicator ($p = 0.087$), the amount of bribes paid ($p = 0.068$), or the gap between administrative tax data and citizen self-reports of payment ($p = 0.004$). The results on the collector-household dimension are more mixed: we fail to reject that the average bribe payment function exhibits complementarity in collector-household type for (extensive margin) indicators of bribe payments ($p = 0.378$, $p = 0.734$) but not for (intensive margin) amount of bribes paid ($p = 0.055$).

[37] Reid and Weigel (2017) compare similar measures with less overt bribe measures in the context of motorcycle taxi drivers paying bribes at Kananga's roadway tools. They find that they line up closely and that it does not appear to be taboo to discuss small payments to officials in Kananga.

Figure A25: Bribe Payments by Collector and Household Types



Notes: This figure shows the estimates of bribe payments for different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis is either an indicator for bribe payment (Panel A), the amount of bribe paid (Panel B), or the gap between administrative tax data and citizen self-reports of payments (Panel C), all measured at midline. The coefficients for the high- and low-propensity households are shown in red and blue, respectively. The points estimates are estimated using equation 7 with bribe payments as the outcome and low-type households assigned to a LL pair of collectors as the excluded category. The vertical lines show the 95% confidence intervals for each of the estimates using standard errors clustered at the neighborhood level. We report the p-value associated with a test for each outcomes, denoted Y, exhibiting increasing differences in collector type (we test H_1: [$Y(H,H,v)$-$Y(L,H,v)$] - [$Y(H,L,v)$-$Y(L,L,v)$] > 0 against H_0: [$Y(H,H,v)$-$Y(L,H,v)$] - [$Y(H,L,v)$-$Y(L,L,v)$] ≤ 0) and in collector and household type (we test H_1: [$Y(H,H,1)$-$Y(L,L,1)$] - [$Y(H,H,0)$-$Y(L,L,0)$] > 0 against H_0: [$Y(H,H,1)$-$Y(L,L,1)$] - [$Y(H,H,0)$-$Y(L,L,0)$] ≤ 0). We discuss these results in Section 8.4.

A11.2 Compliance with Other Formal and Informal Taxes

By increasing compliance with the property tax, implementing the optimal assignment could reduce the payment of other taxes if payments of the property tax and payments of other formal or informal taxes are substitutes (Olken and Singhal, 2011).

In Kananga, the most common contribution is an informal labor levy called *salongo*. *Salongo* is organized weekly by neighborhood chiefs and involves citizens contributing labor to public good projects, such as repairing roads. According to our midline survey data, 37% of citizens participated in *salongo* over a two week period, with participants contributing 4.3 hours on average. The optimal assignment does not appear to have significant effects on *salongo* participation on the extensive (3.890 percentage points, $p = 0.123$) or intensive margin (0.187 hours, $p = 0.299$) (Table 2, Panel B).

Other formal taxes paid by citizens in Kananga include the vehicle tax (3% of endline respondents reported paying), the market vendor fee (17%), the business tax (5%), the income tax (11%). These measures are self-reported but our endline survey included an obsolete poll tax to gauge potential reporting bias. Overall, we find no evidence that the optimal assignment would crowd out payments of other formal taxes (Table 2, Panel C).

A11.3 Views of the Government and Taxation

Finally, if high-type collectors' effectiveness in generating compliance reflects their use of coercion and threats of enforcement, the optimal policy could erode citizen's views of the government and of taxation. We investigate the effects on such beliefs using midline and endline survey data. The optimal assignment does not appear to significantly affect views of government (Table 2, Panel D). It appears to have mixed effects on citizens' view of taxation (Table 2, Panel E), slightly increasing citizen trust in the tax ministry ($p = 0.100$), while marginally reducing the perceived likelihood of enforcement and the perceived share of tax revenue spent on public goods ($p = 0.214$ and $p = 0.106$, respectively). We find no significant impact of the optimal assignment on tax morale ($p = 0.491$). Overall, then, there is little evidence of eroding views of the government or of taxation that might give the government pause in choosing the optimal assignment policy.

A12 Distributional Impacts Estimation

To estimate $\mathbb{E}_f[X_h|Y_h = 1]$ in Equation (12), we express it as a sum of different $\mathbb{E}_f[X_h|Y_h = 1, Z_h]$, where Z_h is the match-type for household h. If household h is of type v and was assigned to collectors of type a_1 and a_2, then $Z_h = (a_1, a_2, v)$. Formally,

$$\begin{aligned}\mathbb{E}_f[X_h|Y_h = 1] &= \sum_z \mathbb{E}[X_h|Y_h = 1, Z_h = z] \cdot \text{Pr}_f(Z_h = z|Y_h = 1) \\ &= \sum_z \mathbb{E}_f[X_h|y_h = 1, Z_h = z] \cdot w_f(z)\end{aligned}$$

where $w_f(z) = \dfrac{f(z)\text{Pr}(Y_h = 1|z)}{\sum_{z'} f(z')\text{Pr}(Y_h = 1|z')}$ is derived from Bayes' Rule. We can then estimate $\mathbb{E}_f[X_h|Y_h = 1]$ as:

$$\sum_z \sum_h \left(\frac{X_h \cdot 1[Y_h = 1] \cdot 1[Z_h = z]}{1[Y_h = 1] \cdot 1[Z_h = z]} \right) \cdot \widehat{w}_f(z)$$

where $\widehat{w}_{f^*}(z) = \dfrac{f^*(z)\widehat{\beta(z)}}{\sum_{z'} f^*(z')\widehat{\beta(z')}}$ and $\widehat{w}_{f^{SQ}}(z) = \dfrac{f^{SQ}(z)\widehat{\beta(z)}}{\sum_{z'} f^{SQ}(z')\widehat{\beta(z')}}$.

A13 Selection Policies

Using the notation introduced in Section 5, we define two types of selection policies that involve reallocating a share $\rho \in [0,1]$ of households previously assigned to low-type collectors. ρ captures the intensity of the selection policy. *Reallocation policies* reassign these households to currently employed high-type collectors while *hiring policies* reassign them to newly hired collectors. Selection policies thus consist in changing the number of assignments by collector type, and involve relaxing the workload constraint in the optimal assignment problem (Equation (4)).

The difference between *reallocation* and *hiring policies* can be summarized by λ, the probability that a household previously assigned to a low-type collector is re-assigned to a high-ability collector. For *reallocation policies*, $\lambda = 1$, while for *hiring policies*, $\lambda = \frac{1}{2}$.[38]

Under a selection policy characterized by ρ and λ, the number of assignments to high-type collectors is given by:

$$N^{asgmt}(H;\rho,\lambda) = N^{asgmt}_{f^{SQ}}(H) + N^{asgmt}_{f^{SQ}}(L)\rho\lambda \tag{27}$$

$N^{asgmt}_{f^{SQ}}(H)$ is the number of households assigned to high-type collectors under the status quo assignment function. $N^{asgmt}_{f^{SQ}}(L)\lambda\rho$ is the number of households reallocated from low-type collectors to high-type collectors under the selection policy characterized by ρ and λ.

Selection policies represent a change in the composition of collector types, but they leave the dependence structure of the assignment unchanged. The joint distribution of collector and household types under the selection policy characterized by ρ and λ is:

$$f^S(a_1,a_2,v;\rho,\lambda) = f^S(a_1;\rho,\lambda)f^S(a_2;\rho,\lambda)f^{SQ}(v) \tag{28}$$

with $f^S(a;\rho,\lambda) \equiv \dfrac{N^{asgmt}(a;\rho,\lambda)}{N^{asgmt}}$.

We can then estimate the impact of the selection policy characterized by ρ and λ by computing its ARE, which is the difference in average tax compliance under the selection policy and the status quo assignment:

$$\tau(\rho,\lambda) \equiv \sum_{v\in V}\sum_{a_1,a_2\in A^2}\left[f^S(a_1,a_2,v;\rho,\lambda) - f^{SQ}(a_1,a_2,v)\right]Y(a_1,a_2,v) \tag{29}$$

To estimate the impact of selection policies, $\tau(\rho,\lambda)$, we substitute the estimated average tax

[38] For *reallocation policies*, $\lambda = 1$ because households previously assigned to low-type collectors are reallocated to high-type collectors. For *hiring policies*, $\lambda = \frac{1}{2}$ because we assume newly hired collectors will be low-type with probability $\frac{1}{2}$ and high-type with probability $\frac{1}{2}$. The effect of similar *hiring policies* have been studied in the teacher value-added literature (e.g., Chetty et al., 2014).

compliance function $\widehat{\beta}(a_1, a_2, v)$ in Equation (29), which gives:

$$\widehat{\tau}(\rho, \lambda) \equiv \sum_{v \in V} \sum_{a_1, a_2 \in A^2} \left[f^S(a_1, a_2, v; \rho, \lambda) - f^{SQ}(a_1, a_2, v) \right] \widehat{\beta}(a_1, a_2, v) \tag{30}$$

where the distributions $f^S(\rho, \lambda)$ and f^{SQ} in $\widehat{\tau}(\rho, \lambda)$ are the theoretical distributions.[39]

[39]This approach contrasts with the estimation of the optimal assignment ARE, which relies on an estimator of the assignment function.

A14 Detailed Survey-based Variable Descriptions

This section provides the exact text of the questions used to construct the survey-based variables considered in the paper.

A14.1 Property and Property Owner Surveys

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

2. *Roof Quality.* This is a Likert scale variable, increasing in the quality of the roof of the respondent's house. It was recorded in the midline and endline survey in response to the prompt: 'Observe the principal material of the roof.' [thatch/ straw, mat, palms/ bamboos, logs (pieces of wood), concrete slab, tiles/slate/eternit, sheet iron]

3. *Wall Quality.* This is a Likert scale variable, increasing in the quality of the walls of the respondent's house. It was recorded in the midline and endline survey in response to the prompt: 'Observe the principal material of the walls of the main house.' [sticks/palms, mud bricks, bricks, cement]

4. *Fence Quality.* This is a Likert scale variable, increasing in the quality of the fence of the respondent's house. It was recorded in the midline and endline survey in response to the prompt: 'Does this compound have a fence? If so, select the type of fence.' [no fence, bamboo fence, brick fence, cement fence]

5. *Erosion Threat.* This is a Likert scale variable, increasing in the threat to the respondent's house caused by erosion. It was recorded in the midline survey in response to the prompt: 'Is this compound threatened by a ravine?' [no, yes - somewhat threatened, yes - gravely threatened]

6. *Distance of the property to state buildings/ health institutions/education institutions.* These distances were based on a survey that recorded the GPS locations of all the important buildings in Kananga. The shortest distance between the respondent's property and each type of location was then computed using ArcGIS.

7. *Distance of the property to the nearest road / to the nearest ravine.* These distances were also measured using GIS. The locations of roads and ravines were digitized on GIS by the research office enabling computation of the distance between the respondent's property and the nearest road or ravine.

8. *Gender.* This is a variable reporting the respondent's gender. It was recorded in the midline survey in response to the prompt: 'Is the owner a man or a woman? '

9. *Age.* This is a variable reporting the respondent's age. It was recorded in the midline survey in response to the question: 'How old were you at your last birthday?'

10. *Employed Indicator.* This is a dummy variable that equals 1 if the respondent reports any job (i.e., is not unemployed). It was recorded in the midline survey in response to the question: 'What type of work do you do now?' [Unemployed-no work, Medical assistant, Lawyer, Cart pusher, Handyman, Driver (car and taxi moto), Tailor, Diamond digger, Farmer, Teacher, Gardener, Mason, Mechanic, Carpenter, Muyanda, Military officer/soldier or police officer, Fisherman, Government personnel, Pastor, Porter, Professor, Guard, Work for NGO, Seller (in market), Seller (in a store), Seller (at home), Student, SNCC, Other]

11. *Salaried Indicator.* This is a dummy variable that equals 1 if the respondent reports one of the following jobs: medical assistant, lawyer, teacher, military officer/soldier or police officer, government personnel, professor, guard, NGO employee, bank employee, brasserie employee, Airtel (telecommunication services) employee, SNCC (national railway company of the Congo) employee. It was recorded in the midline survey in response to the question 'what type of work do you do now?' [responses noted above]

12. *Work for the Government Indicator.* This is a dummy variable that equals 1 if the respondent reports having one of the following jobs: military officer/soldier or police officer, government personnel, or SNCC (national railway company of the Congo) employee. It was recorded in the midline survey in response to the question 'what type of work do you do now?' [responses noted above]

13. *Relative Work for the Government Indicator.* This is a dummy variable that equals 1 if the respondent reports that someone in her/his family works for the government. It was recorded in the midline survey in response to the question: 'Does a close member of the family of the property owner work for the provincial government, not including casual labor?' [no, yes]

14. *Main Tribe Indicator.* This is a dummy variable that equals 1 the respondent reports being Luluwa, the main tribe in Kananga. It was recorded in the midline survey in response to the question: 'What is your tribe?' [Bindi, Bunde, Dekese, Dinga, Kefe, Kele, Kete, Kongo, Kuba, Kuchu, Kusu, Lele, Lualua, Luba, Lubakat, Luluwa, Lunda/Rund, Luntu, Lusambo, Mbala, Mfuya, Mongo, Ndumbi, Ngwandji, Nyambi, Nyoka, Pende, Rega, Sakata, Sala, Shi, Songe, Tetela, Tshokwe, Tutsi, Utu, Uvira, Wongo, Yaka, Yeke, Other]

15. *Years of Education.* This is variable reports the respondent's years of education. It was calculated using responses to two baseline survey questions:

    - 'What is the highest level of school you have reached?' [never been to school, kindergarten, primary, secondary, university]
    - 'What is the last class reached in that level?' [1, 2, 3, 4, 5, 6, >6]

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

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

18. *Trust in Provincial Government / National Government / Tax Ministry / Chief.* This is a Likert scale variable, increasing in the level of trust the respondent reports having in different organizations. It was recorded in the baseline and endline survey in response to the question:

    - 'I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: no confidence at all, not much confidence, quite a lot of confidence, a great deal of confidence?'
    - Organizations:
      (a) 'Local leaders'
      (b) 'The national government (in Kinshasa)'
      (c) 'The provincial government'
      (d) 'The tax ministry'

19. *Paid Bribe.* This is a variable providing the respondent's self-reported bribe payments. The underlying exact midline and endline survey questions are as follows:

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

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

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

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

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

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

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

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

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

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

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

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

30. *Perception of Enforcement.* This is a variable reporting the respondent's perception of how likely it is that one gets sanctioned for not paying property tax. The underlying midline survey question is as follows: 'In your opinion, do you think a public authority will pursue and enforce sanctions among households that did not pay the property tax in 2018? With which point of you do you agree?' [they will definitely sanction them,

they will probably sanction them, they will probably not sanction them, they will definitely not sanction them] We use this variable to construct a dummy that equals 1 if the respondent answered either 'they will definitely sanction them' or 'they will probably sanction them' and 0 otherwise.

31. *Perception of Public Goods Provision.* This is a variable reporting the respondent's perception of how likely it is that property tax revenue is spent on providing public goods in Kananga. The underlying midline survey question is as follows: 'In your opinion, how much of the money collected in property taxes will be spent on public infrastructure, for example the roads in your neighborhood or elsewhere in Kananga?' [All of it, most of it, some of it, none of it] We use this variable to construct a dummy that equals 1 if the respondent answered either 'all of it' or 'most of it' and 0 otherwise.

32. *Collector Messages.* We construct dummy variables that equal 1 if a message was used by the tax collectors during property tax collection, according to household self reports. It was recorded in the midline survey in response to the question: 'Now let's talk about the messages used by the property tax collectors in 2018 to convince property owners to pay the property tax. For each of the following messages, please indicate if you heard the tax collectors say this, or if you heard that they said this to other people.'

    - 'If you refuse to pay the property tax, you may be asked to go to the chief for monitoring and control.' [no, yes]
    - 'If you refuse to pay the property tax, you may be asked to go to the provincial tax ministry for monitoring and control.' [no, yes]
    - 'The Provincial Government will only be able to improve public infrastructure in your community if its residents pay property taxes.' [no, yes]
    - 'The Provincial Government will only be able to improve public infrastructure in Kananga if residents pay property tax.' [no, yes]
    - 'Pay the property tax to show that you have confidence in the state and its officials.' [no, yes]
    - 'It is important.' [no, yes]
    - 'Payment is a legal obligation.' [no, yes]
    - 'Many households are paying; you should pay to avoid embarrassment in your community.' [no, yes]
    - 'If you don't pay, there could be violent consequences.' [no, yes]

33. *Tax Visits.* This is a variable reporting tax collectors' visits to households. The exact midline survey questions are as follows:

    - 'Has your household been visited by a tax collector or another authority in 2018 to raise awareness for collection of the property tax (even if no one was home)?'
    - 'How many times did they come in total since June, including the visit to assign a code?' (Intensive margin)

A14.2 Tax Collectors Surveys

1. *Female.* This is a dummy variable that equals 1 if the respondent is female. It was recorded in the baseline collector survey in response to the prompt: 'Select the sex of the interviewee.' [female, male]

2. *Age.* This is a variable reporting the respondent's age. It was recorded in the baseline collector survey in response to the question: 'How old were you at your last birthday?'

3. *Main Tribe Indicator.* This is a dummy variable that equals 1 the respondent reports being Luluwa, the main tribe in Kananga. It was recorded in the baseline collector survey in response to the question: 'What is your tribe?' [Bindi, Bunde, Dekese, Dinga, Kefe, Kele, Kete, Kongo, Kuba, Kuchu, Kusu, Lele, Lualua, Luba, Lubakat, Luluwa, Lunda/Rund, Luntu, Lusambo, Mbala, Mfuya, Mongo, Ndumbi, Ngwandji, Nyambi, Nyoka, Pende, Rega, Sakata, Sala, Shi, Songe, Tetela, Tshokwe, Tutsi, Utu, Uvira, Wongo, Yaka, Yeke, Other].

4. *Years of Education.* This variable reports the respondent's years of education. It was calculated using responses to two baseline collector survey questions:

   - 'What is the highest level of school you have reached?' [never been to school, kindergarten, primary, secondary, university]
   - 'What is the last class reached in that level?' [1, 2, 3, 4, 5, 6, >6]

5. *Math Score.* This variable is a standardized index increasing in the respondent's math ability. The exact baseline collector survey questions used to create the standardized index are: 'Now we would like to ask you some math problems. Don't worry if you are not sure of the answer, just do your best to answer them.'

   - 'Can you tell me what 2 plus 3 equals?'
   - 'Can you tell me what 6 plus 12 equals?'
   - 'Can you tell me what 32 minus 13 equals?'
   - 'Can you tell me what 10 percent of 100 is?'

6. *Literacy .* This variable is a standardized index increasing in the respondent's ability to read Tshiluba. The exact baseline collector survey questions used to create the standardized index are: 'Now we would like to ask you if you could read two separate paragraphs about tax collection by the provincial government. The first paragraph is in Tshiluba and the second paragraph is in French. Don't worry if you're not sure of certain words, just do your best to read the paragraphs.'

   - 'How well did they read the Tshiluba paragraph?' [could not read, read with lots of difficulty, read with a little difficulty, read perfectly]
   - 'How confidently did they read the Tshiluba paragraph?' [not at all confident, not very confident, a bit confident, very confident]

- ‘How well did they read the French paragraph?’ [could not read, read with lots of difficulty, read with a little difficulty, read perfectly]
- ‘How confidently did they read the French paragraph?’ [not at all confident, not very confident, a bit confident, very confident]

7. *Monthly Income.* This variable is the self-reported income of the respondent. It was recorded in response to the baseline collector survey question: ’What was the household’s total earnings this past month?’ [amount in USD]

8. *Number of Possessions.* This variable report the number of possessions owned by the collector’s household. The exact baseline collector survey question is as follows: ‘In your household, which (if any) of the following do you own?

   - A motorbike [no, yes]
   - A car or a truck [no, yes]
   - A radio [no, yes]
   - A television [no, yes]
   - An electric generator [no, yes]
   - A sewing machine [no, yes]
   - None.’ [no, yes]

9. *Born in Kananga.* This is a dummy variable that equals 1 if the respondent was born in Kananga. The exact baseline collector survey question is as follows: ‘Were you born in Kananga?’ [no, yes]

10. *Trust in Provincial Government / National Government / Tax Ministry.* This is a Likert scale variable increasing in the level of trust the respondent reports having in each organization. The exact baseline collector survey question is as follows:

    - ‘I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: no confidence at all, not much confidence, quite a lot of confidence, a great deal of confidence?’
    - Organizations:
      (a) ‘The national government (in Kinshasa)’
      (b) ‘The provincial government’
      (c) ‘The tax ministry’

      The values were reversed to code this variable.

11. *Provincial Government Capacity.* This is a dummy variable equal to 1 if the collector believes that the government has the capacity to respond to an urgent situation. The exact baseline collector survey question is as follows: ‘Imagine that many of the roads in central Kananga have been badly damaged due to bad weather. Do you think the local government would fix this problem within three months?’ [no, yes]

12. *Provincial Government Responsiveness.* This is a Likert scale variable increasing in the respondent's perception of how responsive the provincial government is. The exact baseline collector survey question is as follows: 'To what degree does the provincial government respond to the needs of your avenue's inhabitants?' [Not very hard working, Hard working, Somewhat hard working, Not hard working]

13. *Provincial Government Performance.* This is a variable increasing in the respondent's perception of the overall performance of the provincial government. The exact baseline collector survey question is as follows: 'How would you rate the performance of the provincial government in Kananga?' [terrible, very poor, poor, fair, very good, excellent]

14. *Provincial Government Corruption.* This is a variable that reports what fraction of the tax revenues from the 2018 property tax campaign the respondent thinks the Provincial Government will put to good use. The exact baseline collector survey question is as follows: 'Now I would like to ask you what you think the provincial government will do with the money it receives from the property tax campaign this year. Imagine that the Provincial Government of Kasaï-Central receives $1000 thanks to this campaign. How much of this money will be put to good use, for example providing public goods?' [0-1000]

15. *Employed Through Connections.* This is a dummy variable equals to 1 if the respondent got his job as a tax collector for the Provincial Tax Ministry through a personal connection. The exact baseline collector survey question is as follows: 'How did you know that a position was available at the Provincial Tax Ministry?' [through a connection at the Provincial Tax Ministry, through a connection in the Provincial Government, I responded to job announcement from the Provincial Tax Ministry, I applied without knowing that the Provincial Tax Ministry was hiring]

16. *Relatives are Provincial Tax Ministry Employees.* This is a dummy variable that equals 1 if the respondent has a family member working at the Provincial Tax Ministry. The exact baseline collector survey question is as follows: 'Do you have a family member who is a Provincial Tax Ministry employee?' [no, yes]

17. *Relatives are Provincial Government Employee.* This is a dummy variable that equals 1 if the respondent has a family member working for the provincial government. The exact baseline collector survey question is as follows: 'Do you have a family member who is a Provincial Government employee?' [no, yes]

18. *Taxes are Important.* This is a Likert scale variable increasing in how important the respondent considers taxes to be. The exact baseline collector survey question is as follows: 'To what degree do you think that paying the property and rent taxes are important for the development of the province?' [not important, important, somewhat important, important, very important]

19. *Provincial Tax Ministry is Important.* This is a Likert scale variable increasing in how important the respondent considers the work of the Provincial Tax Ministry to be. The exact baseline collector survey question is as follows: 'To what degree do you think the work of the Provincial Tax Ministry is important for the development of the province?' [not important, important, somewhat important, important, very important]

20. *Paid Property Tax in the Past.* This is a dummy variable that equals 1 if if the respondent declared having paid the property tax in the past. The exact baseline collector survey question is as follows: 'Have you (or your family) paid your own property tax this year?' [no, yes]

21. *Importance of Progressive Taxes.* This is a dummy variable that equals 1 if the respondent reports that taxes in general should be progressive. The exact baseline collector survey question is as follows: 'Do you think all individuals should be taxed the same amount or should taxes be proportional to someone's income/wealth?' [everyone should pay the same amount, taxes should be proportional to someone's income/wealth]

22. *Importance of Progressive Property Taxes.* This is a dummy variable that equals 1 if the respondent reports that property tax rates should be progressive. The exact baseline collector survey question is as follows: 'According to you who should pay more property tax?' [only the poorest, mostly the poorest but also a little bit the rest of society, everyone should contribute the same amount, mostly the wealthiest but also a little bit the rest of society, only the wealthiest]

23. *Important to Tax Employed Individuals.* This is a Likert scale variable reporting respondent's view of the importance of taxing individuals with salaried jobs in Kananga. The exact baseline collector survey question is 'How important do you think it is to pay the property tax for property owners who are employed?' [not important, somewhat important, important, very important]

24. *Important to Tax Property Owners.* This is a Likert scale variable increasing in respondent's view of the importance of taxing property in Kananga. The exact baseline collector survey question is 'How important do you think it is to pay the property tax for property owners who have lived in a compound for many years?' [not important, somewhat important, important, very important]

25. *Important to Tax Property Owners with a Title.* This is a Likert scale variable reporting respondent's view of the importance of taxing property owners in Kananga. The exact baseline collector survey question is 'How important do you think it is to pay the property tax for property owners who have a formal land title?' [not important, somewhat important, important, very important]

26. *Extrinsic Motivation.* This variable is a standardized index increasing in tax collectors' extrinsic motivation to work as a tax collector. The exact endline collector survey questions used to create the standardized index are: 'Now, I want you to reflect on why you

worked as a tax collector for the IF campaign of 2018. I am going to give you a series of possible reasons for why you did this work. For each reason, indicate if you strongly disagree, disagree, neutral, agree, strongly agree that this is a reason why you worked on the property tax campaign of 2018. Responses:

- 'I did this work because of the income it provided me.'
- 'I did this work because it allowed me to earn money.'
- 'I did this work because it provided me financial security.'
- 'I accept any paid job opportunity that is offered to me.'

27. *Intrinsic Motivation.* This variable is a standardized index increasing in tax collectors' intrinsic motivation to work as a tax collector. The exact endline collector survey questions used to create the standardized index are: 'Now, I want you to reflect on why you worked as a tax collector for the IF campaign of 2018. I am going to give you a series of possible reasons for why you did this work. For each reason, indicate if you strongly disagree, disagree, neutral, agree, strongly agree that this is a reason why you worked on the property tax campaign of 2018.' Responses:

    - 'I did this work because I derived much pleasure from learning new things.'
    - 'I did this work for the satisfaction I experienced from taking on interesting challenges.'
    - 'I did this work for the satisfaction I experienced when I was successful at doing difficult tasks.'

28. *Introjection.* This variable is a standardized index increasing in tax collectors being motivated to work due to introjected regulation. The exact endline collector survey questions used to create the standardized index are: 'Now, I want you to reflect on why you worked as a tax collector for the IF campaign of 2018. I am going to give you a series of possible reasons for why you did this work. For each reason, indicate if you strongly disagree, disagree, neutral, agree, strongly agree that this is a reason why you worked on the property tax campaign of 2018. Responses:

    - 'I wanted to succeed at this job, otherwise I would have been very ashamed of myself.'
    - 'I wanted to be very good at this work, otherwise I would have been very disappointed.'
    - 'I did this work because I wanted to be a "winner" in life.'
    - 'I took this job because I thought it was prestigious.'

29. *Goal Orientation.* This variable is a standardized index increasing in tax collectors being motivated to work due to goal orientation. The exact endline collector survey questions used to create the standardized index are: 'Now, I want you to reflect on why

you worked as a tax collector for the IF campaign of 2018. I am going to give you a series of possible reasons for why you did this work. For each reason, indicate if you strongly disagree, disagree, neutral, agree, strongly agree that this is a reason why you worked on the property tax campaign of 2018. Responses:

- 'I did this work because I wanted to contribute to the economic development of Kananga.'
- 'I did this work because I wanted to help the government do more for the citizens of Kananga.'
- 'I did this work because I wanted to contribute to the increase in the collection of taxes.'

30. *Amotivation.* This variable is a standardized index increasing in tax collector being unmotivated to work as a tax collector. The exact endline collector survey questions used to create the standardized index are: 'In any job, it can also be hard sometimes to feel motivated to work. When reflecting back on the IF campaign of 2018, indicate if any of the following reasons offers explanatory power for feeling unmotivated. For each reason, indicate if you strongly disagree, disagree, neutral, agree, strongly agree that this is a reason why you may not have felt motivated to work on the IF campaign of 2018.' Responses:

    - 'I didn't seem able to manage the tasks the job required of me.'
    - 'We worked under unrealistic working conditions.'
    - 'Our bosses expected too much of us.'

References

Abadie, Alberto and Matias D. Cattaneo, "Econometric Methods for Program Evaluation," *Annual Review of Economics*, 2018, *10*, 465–503.

Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey, "Inference on Winners," *Working Paper*, 2021.

Ashraf, Nava and Oriana Bandiera, "Social Incentives in Organizations," *Annual Review of Economics*, 2018, *10*, 439–463.

Balan, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan Weigel, "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the DR Congo," *American Economic Review*, 2022, *112* (3), 762–797.

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

Bergeron, Augustin, Gabriel Tourek, and Jonathan Weigel, "The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC," *Working Paper*, 2021.

Bhattacharya, Debopam, "Inferring Optimal Peer Assignment from Experimental Data," *Journal of the American Statistical Association*, 2009, *104* (486), 486–500.

Bishop, C. M., *Pattern Recognition and Machine Learning*, Springer, 2006.

Blei, D. M., A. Kucukelbir, and J. D. McAuliffe, "Variational Inference: A Review for Statisticians,," *Journal of the American statistical Association*, 2017, *112* (518), 859–877.

Bonhomme, Stéphane, "Teams: Heterogeneity, sorting, and complementarity," *Working Paper*, 2021.

_, Thibaut Lamadon, and Elena Manresa, "A Distributional Framework for Matched Employer-Employee Data," *Econometrica*, 2019, *87* (3), 699–739.

Burke, Mary A and Tim R Sass, "Classroom Peer Effects and Student Achievement," *Journal of Labor Economics*, 2013, *31* (1), 51–82.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman, "Understanding Mechanisms Underlying Peer Effects: Evidence from a Field Experiment on Financial Decisions," *Econometrica*, 2014, *82* (4), 1273–1301.

Charrad, Malika, Nadia Ghazzali, Veronique Boiteau, and Azam Niknafs, "NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set," *Journal of Statistical Software*, 2014.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 2014, *104* (9), 2633–79.

Cullen, Alison and Christopher Frey, *Probabilistic Techniques in Exposure Assessment*, Springer, 1999.

Dantzig, George B., "Linear Programming," in "Problems for the Numerical Analysis of the Future," Institute for Numerical Analysis, U.S. Government Printing Office, 1951.

Graham, Bryan S., Guido W. Imbens, and Geert Ridder, "Identification and Efficiency Bounds for the Average Match Function under Conditionally Exogenous Matching," *Journal of Business and Economic Statistics*, 2020, *38* (2), 303–316.

Lang, Kevin, "A Language Theory of Discrimination," *The Quarterly Journal of Economics*, 1986, *101* (2), 363–382.

Lloyd, Stuart P., "Least Squares Quantization in PCM," *IEEE Transactions on Information Theory*, 1982.

Luenberger, David G., *Linear and nonlinear programming*, Springer, 1984.

Olken, Benjamin A and Monica Singhal, "Informal taxation," *American Economic Journal: Applied Economics*, 2011, *3* (4), 1–28.

Reid, Otis and Jonathan Weigel, "Citizen Participation in Corruption: Evidence from Roadway Tolls in the Democratic Republic of the Congo," *Working Paper*, 2017.

Rousseeuw, Peter, "Silhouettes: A Graphical Aid to Interpretation and Validation of Cluster Analysis," *Computational and Applied Mathematics*, 1987.

Thorndike, Robert, "Who Belongs in the Family?," *Psychometrika*, 1953, *18* (4), 267–276.

Todd, Petra E and Kenneth I Wolpin, "On the Specification and Estimation of the Production Function for Cognitive Achievement," *The Economic Journal*, 2003, *113* (485), F3–F33.

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