DISCUSSION PAPER SERIES

DP16771

OPTIMAL ASSIGNMENT OF BUREAUCRATS: EVIDENCE FROM RANDOMLY ASSIGNED TAX COLLECTORS IN THE DRC

Augustin Bergeron, Pedro Bessone, John Kabeya Kabeya, Gabriel Tourek and Jonathan Weigel

DEVELOPMENT ECONOMICS

PUBLIC ECONOMICS

CEPR

ISSN 0265-8003

OPTIMAL ASSIGNMENT OF BUREAUCRATS: EVIDENCE FROM RANDOMLY ASSIGNED TAX COLLECTORS IN THE DRC

Augustin Bergeron, Pedro Bessone, John Kabeya Kabeya, Gabriel Tourek and Jonathan Weigel

Discussion Paper DP16771
Published 04 December 2021
Submitted 25 November 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Development Economics
- Public Economics

OPTIMAL ASSIGNMENT OF BUREAUCRATS: EVIDENCE FROM RANDOMLY ASSIGNED TAX COLLECTORS IN THE DRC

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 through the optimal assignment of its tax collectors. We study the two-stage random assignment of property tax collectors (i) into teams and (ii) 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. Implementing the optimal assignment would increase tax compliance by an estimated 37% relative to the status quo (random) assignment. By contrast, to achieve a similar increase under the status quo assignment, the government would have to replace 63% of low-ability collectors with high-ability ones or to increase collectors' performance wages by 69%.

JEL Classification: D73, H11, M50, H20

Keywords: bureaucracy, taxation, assortative matching

Augustin Bergeron - abergeron@stanford.edu
Stanford

Pedro Bessone - pedrobtepedino@gmail.com
Uber

John Kabeya Kabeya - johnkabeya82@gmail.com
Direction Générale des Recettes du Kasaï-Central

Gabriel Tourek - gzt1@pitt.edu
University of Pittsburgh

Jonathan Weigel - j.weigel@lse.ac.uk
Haas School of Business, UC Berkeley and CEPR

Acknowledgements

We thank Timothy Armstrong, Oriana Bandiera, Pablo Balan, Vittorio Bassi, Anne Brockmeyer, Katherine Casey, Arun Chandrasekhar, Thomas Chaney, Raj Chetty, Lucas de Lima, Esther Duflo, Pascaline Dupas, Marcel Fafchamps, Edward Glaeser, Nathan Hendren, Matthew Kahn, Asim Khwaja, Toru Kitagawa, Gabriel Kreindler, Pablo Kurlat, Isabela Manelici, Robert Metcalfe, Joana Naritomi, Nathan Nunn, Paulina Oliva, Benjamin Olken, Imran Rasul, Frank Schillbach, Jeffrey Weaver, and Noam Yuchtman for invaluable suggestions. For outstanding research assistance, we thank Mert Akan, Manon Delvaux, Samih Ferrah, Arnaud Fournier, 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.

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

November 17, 2021

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 through the optimal assignment of its tax collectors. We study the two-stage random assignment of property tax collectors (*i*) into teams and (*ii*) 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. Implementing the optimal assignment would increase tax compliance by an estimated 37% relative to the status quo (random) assignment. By contrast, to achieve a similar increase under the status quo assignment, the government would have to replace 63% of low-ability collectors with high-ability ones or to increase collectors' performance wages by 69%.

Affiliations: Bergeron – Stanford University (abergeron@stanford.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 CEPR (jweigel@berkeley.edu).
We thank Timothy Armstrong, Oriana Bandiera, Vittorio Bassi, Anne Brockmeyer, Katherine Casey, Arun Chandrasekhar, Thomas Chaney, Raj Chetty, Lucas de Lima, Esther Duflo, Pascaline Dupas, Marcel Fafchamps, Edward Glaeser, Nathan Hendren, Matthew Kahn, Asim Khwaja, Toru Kitagawa, Gabriel Kreindler, Pablo Kurlat, Isabela Manelici, Robert Metcalfe, Joana Naritomi, Nathan Nunn, Paulina Oliva, Benjamin Olken, Imran Rasul, Frank Schillbach, Jeffrey Weaver, and Noam Yuchtman for invaluable suggestions. For outstanding research assistance, we thank Mert Akan, Manon Delvaux, Samih Ferrah, Arnaud Fournier, 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 wage or promotion incentives, are typically unavailable to governments because of seniority-based civil service regulations.[2] Moreover, there is growing recognition that public sector workers explain much of the variation in state performance across sectors and regions (Finan et al., 2015; Best et al., 2019; Fenizia, 2019). Yet, we have little evidence on whether the assignment of public sector employees to postings or teams can enhance state effectiveness.[3]

This paper examines front-line bureaucrat assignment 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 Kasai Central randomized tax collectors to teammates and neighborhoods to minimize collusion between collectors and households. 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. The median collector worked with 6 different teammates and in 12 different neighborhoods (covering 1,200 properties) during the campaign.

We use this two-stage randomization to estimate the optimal assignment — of collectors to teammates, and of teams to households — and its impact on tax compliance, i.e.,

[1] See, e.g., Shapley and Shubik (1971); Becker (1973); Crawford and Knoer (1981) on the role of assignment theoretically and, e.g., Graham (2011); Graham et al. (2014); Bhattacharya and Dupas (2012); Bonhomme (2021) on estimation for different classes 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) provide direct evidence that rigid bureaucratic promotion rules constrain the performance of public sector workers.

[3] Khan et al. (2019) provide evidence on a similar but distinct question: can the reward of future performance-based postings create incentives for bureaucrats to improve outcomes? By contrast, we focus on the direct effects of assignments on bureaucrat performance.

the probability that households pay taxes.[4] 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 prior to tax collection in 78 randomly selected neighborhoods (our *analysis sample*).[5] Chiefs' estimates are highly correlated with subsequent tax compliance during the campaign, providing a convenient pre-treatment measure of each household's type. Similarly, we partition tax collectors into two types.[6] 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 households' payment propensity (our *holdout sample*). Specifically, we define collector types (high and low) as whether they were above or below the median in terms of average tax compliance achieved across all neighborhoods they were randomly assigned to in this holdout sample. We estimate the average compliance associated with each collector using a fixed effects model and Empirical Bayes adjustment (Morris, 1983) to increase precision.

Having defined tax 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 optimal assignment function: the assignment of collectors to teammates and households that maximizes tax compliance subject to status quo constraints on team workload and size.[7] Finally, we estimate the effect of implementing the optimal assignment — relative to the status quo random assignment — on tax compliance and revenue.

It is not obvious, ex ante, what assignment function would maximize tax compliance in this setting.[8] If collection from households characterized by a high tax payment propensity is a simple task, then it could be optimal to assign them to low-ability collectors. If instead collection from high tax payment propensity households requires effort and per-

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

[5]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.

[6]We use two types to maximize power, but the results are robust to allowing for more types (Table A7).

[7]The optimal assignment similarly holds constant the random reshuffling of collectors into new teams each month to prevent the emergence of collusion/covering, as in the status quo assignment.

[8]Past empirical work on optimal matching (e.g., Carrell et al., 2009, 2013; Aucejo et al., 2019; Bhattacharya, 2009; Fenizia, 2019; Graham et al., 2020b; Marx et al., 2021) also reaches mixed conclusions, as we discuss in Section 7.1.

suasion skills, then assigning them to high-ability collectors could be optimal. Similarly, when forming collector teams, if only one high-ability collector is required to ensure that all essential tasks are completed, then one might expect that pairing a high-ability with a low-ability collector (i.e., mixed teams) would prove optimal. However, there could also be scope for complementarities between collectors' effort or skills that would justify grouping high-ability collectors together and low-ability collectors together (i.e., homogeneous teams). What assignment function maximizes tax compliance is thus an empirical question.[9]

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 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 and collector-to-household match 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 on days and times when they have the cash on hand to pay.

Implementing the optimal assignment would increase tax compliance by an estimated 2.941 percentage points relative to the status quo random assignment. This amounts to a 37% increase in compliance relative to the status quo average of 8%. Tax revenue would increase by 26% under the optimal assignment. Each dimension of the optimal assignment — collector-to-collector and collector-to-household — appears to contribute roughly equally to the total effect of the optimal assignment. Specifically, optimizing only on the assignment of collectors to teammates would increase compliance by 16%, while optimizing only on the assignment of collectors to households would increase compliance by 10%. Concerning incidence, the increase in tax compliance under the optimal policy would be progressivity-enhancing, largely falling on wealthier households with more valuable prop-

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

erties.

We consider a range of robustness checks, including using alternative definitions of household and collector type, optimizing with three collector types (rather than two), redoing the analysis with neighborhood-level (rather than household-level) assignments, assuming alternative government maximands, and providing estimates robust to overfitting and the "winner's curse." None of these exercises qualitatively change the main results. We also investigate several spillover/SUTVA concerns, including the possibility that changing collectors' assignments could directly impact effort levels or opportunities for learning over the course of the campaign by match type. According to the available evidence, these concerns are unlikely to be a source of bias in our estimates.

To benchmark the magnitude of these effects, we compare the optimal assignment policy to selection policies, which consist of reallocating households assigned to low-type collectors to high-type collectors (*reallocation policies*) or to newly hired collectors (*hiring policies*).[10] 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 not achieve compliance gains comparable to those from the optimal assignment, even if all low-type collectors' households were reallocated.[11]

As a further benchmark, we compare our results to the effect of performance-based financial incentives to 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.[12] 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.

Finally, we 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. States often rotate tax collectors to prevent collusion with tax-

[10] When studying replacing a low-type collector with a newly hired tax collector, we assume that the new hire is low-type with probability 0.5 and high-type with probability 0.5. Similar policies have been used as a benchmark in the literature on teacher value-added (e.g., Chetty et al., 2014).

[11] These are conservative estimates because they factor in neither possible negative externalities on high-type collectors due to the increase in workload, nor the search and training costs of hiring new collectors.

[12] We describe the randomization of piece-rate wages in Section 2 and explore the effects of piece-rate wages on compliance and revenue in further detail in Bergeron et al. (2020b).

payers (Brewer, 1990), and bribery was an explicit concern of the tax authority during the 2018 property tax campaign. Using survey data on bribe payment, we find suggestive evidence that the optimal policy would increase bribe payments to tax collectors. However, it would not affect citizens' compliance with other taxes, their view of the government, or their tax morale. Faced with these mixed results, the government would need to weight the social cost of $1 paid in bribes about four times higher than the value of $1 in tax revenue to favor the status quo over the optimal assignment.

We contribute to three strands of literature. First, we provide some of the first estimates of the importance of bureaucrat assignment in shaping state effectiveness in revenue mobilization. While past work examines the importance of selection (Dal Bó et al., 2013; Callen et al., 2015; Hanna and Wang, 2017; Xu, 2018; Ashraf et al., 2020; Dahis et al., 2020), 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; Rasul et al., 2021; Bandiera et al., 2021) of public-sector workers, less attention has been paid to the assignment of bureaucrats as a source of state effectiveness. Two closely related papers are Best et al. (2019) and Fenizia (2019), which exploit the rotation of bureaucrats across sites to study the role of bureaucrat quality in explaining public sector performance.[13,14] We build on these studies by exploring the optimal assignment of bureaucrats to teams and postings,[15] leveraging the random assignment of tax collectors and studying more objective performance measures (tax compliance and revenue) than are typically available for bureaucrats. Finally, we advance this literature by exploiting rich survey data to explore the mechanisms explaining the optimal assignment of collectors and to consider other policy-relevant response margins, such as tax incidence, corruption, fiscal externalities, and citizens' views of the tax authority.

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

[14] We also quantify the importance of tax collectors in explaining tax compliance in Kananga. Our results suggest that collectors explain 36% of the variance in compliance across neighborhoods. In comparison, Fenizia (2019) finds that public sector managers explain 9% of the total variance in productivity, while Best et al. (2019) show that bureaucrats who manage procurement processes explain over 24% of the variation in quality-adjusted public procurement prices.

[15] Fenizia (2019) 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.

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; Kleven et al., 2016), 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., 2021), and the use of large taxpayer offices to increase the staff-to-taxpayer ratio (Basri et al., 2019).[16] 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 studied 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).[17] While these papers consider uni-dimensional assignment problems, we study the bi-dimensional problem of assigning collectors to teammates and households. In our context, considering only 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.[18] Finally, we make a small methodological contribution by applying the median-unbiased estimators developed by Andrews et al. (2019) to address possible "winner's curse" upward bias that can arise in optimization problems like those considered in this literature.

This paper is organized as follows. Sections 2, 3, and 4 respectively review the setting, design, and data. Section 5 introduces the conceptual framework, before presenting how it is empirically estimated in Section 6. Section 7 describes the optimal assignment policy and discusses potential mechanisms explaining the matching of collectors to teammates and households under the optimal assignment. Section 8 explores the impacts of the optimal

[16] 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., 2019; Bergeron et al., 2020b; Brockmeyer et al., 2020).

[17] Another related paper is Marx et al. (2021), which studies how ethnic heterogeneity in teams impacts the performance of a canvassing nonprofit in Kenya.

[18] Carrell et al. (2009) study peer effects using the random assignment of students to peer groups, and Graham et al. (2020a) study the optimal assignment of teachers to classrooms by leveraging random assignment.

assignment policy on tax compliance and revenue. Section 9 explores the effects of the optimal assignment policy on bribery, payments of other taxes, and citizens' views of the government and of taxation, before concluding in Section 10.

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.[19] 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 revenue mobilization by local governments (Franzsen and McCluskey, 2017), the provincial government has turned to the property tax to increase tax revenue, conducting a series of citywide door-to-door collection campaigns since 2016 (Weigel, 2020; Balan et al., 2021).

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 main arteries are paved, and even these 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.[20] They were drawn from a pool of aspiring bureaucrats who frequently perform contract work for different arms of the provincial government.[21] They did not receive a regular salary outside of the piece-rate

[19] The tax-GDP ratio was 7.7% in 2018, compared to an African average of 16.5% (OECD, 2020). Globally the tax-GDP ratio ranks 188 out of 200 countries, including oil-rich countries.

[20] In some neighborhoods, which are excluded from this analysis, tax collection was conducted by the neighborhood chiefs, as described in Balan et al. (2021).

[21] Such contract work typically consists of public administration tasks like tax collection, land titling, and

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 (Table A3). 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 this tax and all types of tax collection for two reasons. First, the government believes that receiving a visit from two collectors is likely to project greater authority.[22] Second, it assumes that working in teams reduces the opportunities for collusion between collectors and households because hiding illegal financial transactions is potentially harder when another tax collector is present.[23] In this way, collection by teams could 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 (e.g., Burgess et al., 2010; Khan et al., 2016; Ashraf and Bandiera, 2018; Banerjee et al., 2021; Marx et al., 2021), and field-based visits from tax collectors/inspectors are a cornerstone in tax authorities' enforcement arsenal (e.g., Khan et al., 2016; Cogneau et al., 2020; Krause, 2020; Okunogbe, 2021).

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. During registration visits, collectors assigned a unique tax ID to each property and issued official tax notices showing the tax liability and other information about the tax.[24] Collectors assessed each property's tax liability based

vaccination campaigns, among others.

[22]Anecdotally, there was also a strong norm among collectors to work in teams, again because they felt "stronger" in demanding payment of the tax — i.e., they believed it enabled them to present a more credible threat of enforcement.

[23]This logic is consistent with the discussion of collusion in hierarchies in Tirole (1986), as well as the notion that tax evasion should be less common in large firms with multiple potential whistleblowers (Kleven et al., 2016).

[24]Additionally, owners were informed that they could always pay at the provincial tax ministry, if they preferred. In total, 38 property owners — about 1% of taxpayers — paid at the ministry, even though paying in this manner increased the transaction costs of tax compliance.

on the principal house's construction, as described below, or whether it was exempt.[25] 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.[26]

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 or door of the house (adjacent to the property code). The in-person nature of tax collection thus left much to the discretion of collectors: which properties to revisit, how many times to revisit them, what persuasion tactics and messages to use to try 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.

When a property owner paid the tax, collectors used handheld receipt printers to issue receipts. The transaction-level receipt data was automatically uploaded to the government's tax database when the collector returned the device to the tax ministry every few days. Any persistent discrepancies between deposited tax revenues and transactions in the receipt data would be deducted from collectors' compensation or cause for suspension (and was rare in practice).

Collector Compensation. Collectors earned piece-rate wages with two components. First, they received 30 Congolese Francs (CF) per property registered. Second, they earned compensation proportional to the amount of tax they individually submitted to the state account.[27] Individual compensation diminished incentives for free-riding.[28] Collectors

[25]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.

[26]Only 3.5% of taxpayers paid during property registration. The remaining 96.5% of taxpayers paid during follow-up tax collector visits.

[27]As discussed by Khan et al. (2016), performance pay is often used among tax collectors in settings like Pakistan, Brazil, and elsewhere. Specifically, the compensation scheme in Kananga varied randomly on the property level between (*i*) 30% of the amount of tax collected, and (*ii*) a constant 750 CF per property (independent of the rate). We explore this variation in Section 8.

[28]In practice, collectors rarely worked alone (unless their partner was sick or absent for some reason). When working together, they were instructed to alternate which collector takes the payment of different households. We observe in the receipt data that they followed this alternation norm closely.

were also reimbursed for one round trip per day from the tax ministry to their assigned neighborhoods. On top of the monetary compensation, collectors also had career incentives to perform well: after the previous property tax campaign, the tax ministry hired the best tax collectors for more secure, 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 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 consisted of 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 flat, 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,[29] 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.[30,31] 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 annual deadline in theory owe 250% of the original liability plus the possibility of a 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 collectors may prove more effective at collecting taxes than others, which we consider in Section 7.2.1.

[29] We estimate property value using machine learning as described in Bergeron et al. (2020a).

[30] There were 45,162 registered properties in Kananga according to the 2018 property register. 40,183 were classified in the low-value band, and 4,979 were classified in the high-value band.

[31] An additional 285 higher-value properties, classified as villas, were taxed according to a different schedule and by different collectors and thus are excluded from our analysis.

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. 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 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.[32,33] These 180 neighborhoods span two randomly selected sub-samples — in which the same state tax collectors worked — that we leverage in our analysis. First, in 78 neighborhoods (6,904 properties), which we refer to as 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.[34] We will use the chiefs' predictions to estimate household type (cf. Section 6.1).[35] Second, in the 102 remaining neighborhoods (11,732 properties), which we refer to as the *holdout sample*, we estimate collector types using a fixed effects approach (cf. Section 6.2).[36] After defining types, we will then estimate the

[32] The tax campaign was active in 364 neighborhoods across Kananga, but we exclude 184 neighborhoods from the analysis: (*i*) 8 neighborhoods where a logistics pilot took place, (*ii*) 111 neighborhoods where city chiefs collected taxes ("Local" neighborhoods in Balan et al. (2021)), (*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. (2021)), (*iv*) 5 neighborhoods with no door-to-door collection (the pure control in Balan et al. (2021)), and (*v*) 10 neighborhoods where one of the collectors in the team never worked in any other neighborhood. We exclude these neighborhoods from our analysis because tax collectors were not randomly assigned to neighborhoods or to teammates (*i - iii*), no citizens paid taxes (*iv*), or because collectors only worked with a single teammate (*v*), preventing us from obtaining fixed effect estimates of collector type (as discussed in Section 5).

[33] In total, 47 state collectors were involved in the 2018 property tax campaign. However, we exclude from our analysis 13 state collectors who were randomly assigned to work with neighborhood chiefs (the "Central X Local" treatment arm in Balan et al. (2021)) or who only worked with one teammate during the tax campaign.

[34] Balan et al. (2021) describes in further detail the random assignment of 78 neighborhoods to this treatment arm to compare city chiefs as tax collectors to state collectors provided with local information.

[35] These neighborhoods are called "Central + Local Information" in Balan et al. (2021).

[36] The number of properties in the analysis and holdout samples (18,636) does not correspond exactly to the number of registered properties (19,600) described in Section 4.1. This is due to missing values in chiefs' predictions for 964 (12%) of the 7,868 registered properties in the analysis sample.

average tax compliance function and the optimal assignment in the analysis sample (cf. Section 6.3).

The provincial tax ministry has used this randomized assignment approach 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.[37] 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.[38] By randomly reassigning collectors to teammates monthly and teams to neighborhoods, the government sought to minimize such collusion.[39]

Many tax authorities deliberately reshuffle collectors in a similar fashion to prevent collusion. 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 (Xu, 2018), 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.[40] 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.[41] The status quo (random) assignment is thus an informative benchmark to compare to the optimal assignment policy because it is clearly defined and because it resembles the practices of tax authorities in many developing countries.

3.2 Balance

Table 1 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. (2020a). Panel B considers property owner characteristics

[37] In particular, in 2018, the tax authority compared state agents to city chiefs as property tax collectors, and the randomization of state agents enabled a cleaner comparison. Balan et al. (2021) provides further detail.

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

[39] Additionally, randomly reshuffling teams each month may prevent collectors from covering for one another if such collusion is easier to sustain with repeated interactions.

[40] 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).

[41] By contrast, if the status quo assignment were opaque, it would be difficult to assess the external validity of our analysis of the optimal policy.

collected at midline that are unlikely to be 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 1 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.[42] This is in line with what one would expect under random assignment. Table 1 also reports tests of the omnibus null hypothesis that the treatment effects are all zero using parametric F-tests for bilateral comparisons. In all cases, we fail to reject the omnibus null hypothesis for the property and property owner 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 tax ID numbers, geographic coordinates, property owner names, property classifications (cf. Section 2.1), exemption status, and tax rates.[43] The handheld receipt printers used by tax collectors during both stages of the campaign stored details of each transaction in their memory.[44] 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.

[42] 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.

[43] The universe of registered properties in Kananga is 45,162. But, as noted in Section 3, we exclude neighborhoods without random assignment of collectors. We also exclude exempt properties. These two restrictions reduce the number of registered properties to 19,600.

[44] If citizens chose to visit the tax ministry themselves to pay, which was possible everywhere, an official there similarly issued a receipt, such that these transactions appear in the administrative data.

4.2 Household Surveys

Enumerators working for the research team administered baseline surveys to 1,404 households from July to December in 2017.[45] 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.[46,47]

4.3 Collector Surveys

Before the tax campaign, enumerators administered a baseline survey with collectors covering demographics, trust in the government, perceived performance of the government, views of taxation, and preferences for redistribution.[48] Enumerators surveyed the 34 collectors who comprise our analysis sample.

5 Conceptual Framework

5.1 Household and Collector Types

We consider an economy with N_h households and N_c tax collectors. Households are characterized by observable type $v_h \in V$ and collectors by observable type $a_c \in A$, where A

[45] The baseline survey was conducted with a total of 4,343 respondents. But, after restricting to neighborhoods with random assignment of collectors and excluding exempt households, we have 1,404 baseline respondents.

[46] The midline survey was conducted with 36,314 total respondents. In the restricted sample studied in this paper, we have 16,346 midline surveys in total, 11,448 of which were conducted with the owner.

[47] Attrition between registration and the midline survey (15%) is balanced across treatments (Table 1).

[48] We also rely on data from an endline survey conducted with collectors after tax collection when analyzing collectors' motivation in Section A8.1.

and V are finite ordered sets. In the context of tax collection, we define each household's type as its likelihood of paying the property tax and each collector's type as their ability to collect taxes.[49] This section refers to finite sets A and V of arbitrary size, but to maximize power, our main estimating equation will assume that households are either low-type ($v = l$) or high-type ($v = h$), i.e., $v = \{l, h\}$. Similarly, we assume that tax collectors are either low-type ($a = L$) or high-type ($a = H$), i.e., $A = \{L, H\}$.[50]

Tax collectors work in pairs. Each neighborhood — and thus each household — is assigned to a collector pair. 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), indicating the type of the collectors and the household.[51] The order of the collectors is arbitrary given that they perform an identical task.

5.2 Average Tax Compliance Function

We assume the government seeks to maximize tax compliance, i.e., the probability that households pay taxes conditional on collector and household types:[52]

$$Y(a_1, a_2, v_h) = \mathbb{E}[y_h(c_1, c_2) | a_{c_1} = a_1, a_{c_2} = a_2, v_h],$$

The government's problem is to pick an assignment function f, a probability mass function that gives the distribution of each match type (a_1, a_2, v_h) that determines both the collector-to-collector and the collector-to-household dimensions of the assignment.[53]

5.3 Status Quo Assignment

Throughout the paper, we compare the optimal assignment function to the status quo assignment function. In our setting, the status quo assignment consists of randomly assigning collectors to teammates and collector pairs to neighborhoods.[54] We can therefore write the status quo assignment function as $f^{SQ}(a_1, a_2, v) = f_a^{SQ}(a_1) f_a^{SQ}(a_2) f_v^{SQ}(v)$.

[49]We describe how household and collector types are estimated in Sections 6.1 and 6.2, respectively.

[50]We show robustness to optimizing with three collector types (rather than two) in Figure A15 and Table A7.

[51]When a pair of collectors of types a_1 and a_2 work together, we denote their team as a a_1-a_2 team or pair. For example, if a team of collectors c_1 and c_2 are of type H, we refer to them as an *H-H* team or pair.

[52]We also consider the case where the government maximizes tax revenues or tax revenues net of bribe payments in Section 8.2.

[53]Since the order of the collectors is arbitrary, we assume that $f(a_1, a_2, v_h) = f(a_2, a_1, v_h)$.

[54]As noted in Section 3, frequently reshuffling teams and postings is a common strategy among tax authorities to reduce collusion between tax collectors and households. Note also that the optimal policy we study holds constant the random reshuffling of collectors to new partners each month while varying whether they work with low- or high-type partners.

We focus primarily on the status quo assignment function with two collector types and two household types, i.e., $a \in \{L, H\}$ and $v \in \{l, h\}$. For our definition of collector type introduced in Section 6.2, collector types L and H are equally distributed, i.e., $f_a^{SQ}(H) = f_a^{SQ}(L) = \frac{1}{2}$. As a result $f^{SQ}(H, H, v) = f^{SQ}(L, L, v) = f^{SQ}(L, H, v) = f^{SQ}(H, L, v) = \frac{1}{4} f_v^{SQ}(v)$, with $f_v^{SQ}(v)$ the share of v-type households in the population. We characterize $f^{SQ}(v)$ empirically in Section 6.1, where we describe the definition of household type v and show that for our definition $f^{SQ}(l) \approx 1/3$ and $f^{SQ}(h) \approx 2/3$.

5.4 Optimal Assignment

We study the optimal assignment, which is the assignment that maximizes expected tax compliance while keeping the marginal distributions in collector and household type the same as under the status quo assignment. To formally define the optimal assignment, we need to introduce additional notation. First, consider $N_f^{asgmt}(a, v)$, the number of v-type households assigned to a-type collectors under assignment function f:

$$N_f^{asgmt}(a, v) = N_h \left[2f(a, a, v) + \sum_{a' \neq a} \big(f(a, a', v) + f(a', a, v) \big) \right]$$

For (a, a, v) matches, a-type collectors are assigned twice to a v-type household, and the number of such assignments is $2N_h f(a, a, v)$. For (a, a', v_h) or (a', a, v_h) matches, a-type collectors are assigned to one v-type household, and the number of such assignments is $N_h \cdot \sum_{a' \neq a} (f(a, a', v) + f(a', a, v))$.[55] Second, we denote $N_f^{asgmt}(a)$ the total number of households assigned to a-type collectors, i.e., their total workload. Third, consider $N^{asgmt} = 2N_h$, the total number of collector assignments.[56] Fourth, we define the marginal distribution of a-type collectors as $f_a(a) = N_f^{asgmt}(a)/N^{asgmt}$, the share of assignments allocated to a-type collectors. Lastly, we define the marginal distribution of v-type households as $f_v(v) = N_h(v)/N_h$, the share of v-type households in the population.

Using this notation, we can define the optimal assignment problem as:

[55] As an example, consider the case where there are 100 households ($N_h = 100$) and all of them are of the same type. Assume two L-type and two H-type collectors. Lastly, assume that the assignment f is uniform: i.e., $f(a_1, a_2, v) = 1/4 \; \forall (a_1, a_2)$. In this example, 25 households are assigned to an *H*-*H* pair, 50 to an *L*-*H* pair (ignoring the order of types), and 25 households to an *L*-*L* pair. As a consequence, there are 50 times in which an H-type collector is assigned to a household while working as part of an HH-pair (i.e., the two type H collectors are assigned to 25 households), 50 times in *L*-*H* pairs, and $N_f^{asgmt}(H, v) = 100$.

[56] N^{asgmt} is equal to two times the total number of households since each household is assigned to two tax collectors.

Problem 1. *Optimal Assignment*

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

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

$$\sum_{v \in V} N_f^{asgmt}(a, v) = N_{f^{SQ}}^{asgmt}(a) \qquad \forall a \in A \tag{3}$$

The Optimal Assignment Problem consists in finding the assignment function f^* that maximizes expected tax compliance in Equation (1) under the constraints described in Equations (2) and (3).[57] Equation (2) is a non-overlapping assignment constraint. It requires that the number of assignments of tax collector pairs to v-type households under f is equal to the total number of v-type households. In other words, the government can only assign one team of collectors to each household. Equation (3) is a workload constraint. It requires that the total number of households assigned to a-type collectors is equal under f and under the status quo assignment. In other words, the government must keep a constant workload by collector type.[58] We discuss the uniqueness and asymptotic properties of the optimal assignment function in Appendix Sections A2.1 and A2.2-A2.3, respectively.

Having defined the optimal assignment, we can estimate the impact of the optimal assignment by computing the *Average Reallocation Effect* (ARE, Graham et al. (2014)), which is the difference in average tax compliance under the optimal and the status quo assignment:

$$ARE = \sum_{v \in V} \sum_{a_1, a_2 \in A^2} \Big[f^*(a_1, a_2, v) - f^{SQ}(a_1, a_2, v) \Big] Y(a_1, a_2, v) \tag{4}$$

6 Estimation

To characterize the optimal assignment function and estimate the return to the optimal assignment empirically, we first need to estimate household and collector types.

[57] There is implicitly one additional constraint, which is that the order of the tax collector is irrelevant, i.e., $f(a_1, a_2, v) = f(a_2, a_1, v)\ \forall a1, a2 \in A^2, v \in V$.

[58] This constraint ensures that the optimal assignment is resource-neutral by ruling out policies that change the distribution of collector types or the number of assignments by collector type relative to the status quo.

6.1 Household Type

When estimating household type, the goal is to capture each household's pre-treatment propensity to pay the property tax.[59] We estimate household type by leveraging a unique feature of the field experiment we study. As described in Section 3, in the 78 neighborhoods of the analysis sample, local chiefs reported each property owner's ability to pay the property tax before tax collection started in the neighborhood. During consultations with state collectors, chiefs went line by line through the neighborhood property roll, guided by the property 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.[60] As shown in Balan et al. (2021), chiefs' estimates were highly predictive of property tax payment (Figure A1), even controlling for household characteristics.[61]

We classify households as low-type ($v = l$) if deemed "unlikely" to be able to pay the property tax according to their neighborhood chief, or high-type ($v = h$) if deemed "likely" or "very likely" to be able to pay.[62] According to this definition, 67% of households are high-type. The optimal assignment estimation therefore relies on the 78 neighborhoods in the analysis sample for which we have chiefs' estimates of household type. Other than these estimates, these 78 neighborhoods are identical to the 102 neighborhoods in the holdout sample where state collectors also worked, given that they were randomly selected.[63]

Although we prefer using the chief estimates because they were elicited before tax collection and are the best available predictor of tax compliance,[64] predicting household types using observable house and property owner characteristics might be easier for some governments.[65] Section 8.2 explores robustness to estimating household types based on

[59] We unfortunately cannot use prior tax compliance because properties' unique tax ID numbers were reassigned in 2018 during the first step of the campaign (cf. Section 3).

[60] Chiefs also reported the willingness to pay of each household, separate from their ability to pay. However, this measure was introduced in the second month of consultations and is thus only available for a smaller sample. We therefore use only the ability to pay measure in our estimation of household type.

[61] On average a one-unit increase in the neighborhood chief's ability-to-pay ranking is associated with an 4.32 percentage-point increase in the probability of subsequent tax payment.

[62] This is the most natural partition with two types of collectors since the gap in compliance is much larger between owners who are "unlikely" and "likely" to pay than between owners who are "likely" and "very likely" to pay (Figure A1).

[63] Balan et al. (2021) show that the assignment of neighborhoods to what we call the analysis and holdout samples is orthogonal to observable characteristics of the property and of the property owner.

[64] Indeed, the correlation between tax compliance and household type is higher when household type is based on chiefs' estimates (0.1017) than when it is based on house characteristics from surveys (0.0481).

[65] For instance, city chiefs might not exist at a local level where they would have rich information about

the relationship between house characteristics and tax compliance in the holdout sample of 102 neighborhoods. The estimated impacts of the optimal assignment policy are similar in magnitude but estimated with less precision.[66]

6.2 Collector Type

We have no informative pre-treatment measure of collector type. To solve this problem, we use a sample splitting approach and estimate collector type in the holdout sample of 102 neighborhoods for which we don't have chiefs' estimates of household type. This partitioning of neighborhoods allows us to avoid estimating collector types within the analysis sample, which could lead to overfitting (i.e., attributing collector type partly based on noise) and might mechanically generate complementarity in collector types (Mullainathan and Spiess, 2017).

In this holdout sample of 102 neighborhoods (11,732 properties), we estimate collector type, q_c, as the average tax compliance collector c achieved across all randomly assigned neighborhoods:

$$q_c = \mathbb{E}\left[Y_h(c_1, c_2, v_h) | c_1 = c\right] \tag{5}$$

which we can estimate using the following fixed-effect regression:

$$y_{hnt} = \sum_{c'} \alpha_{c'} 1_{[c' \in c(n)]} + \lambda_t + \varepsilon_{hnt} \tag{6}$$

where y_{hnt} is an indicator for household h in neighborhood n paying the property tax during the 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. As discussed in Section 3, collectors worked simultaneously in two neighborhoods during successive month-long periods of the property tax campaign.[67] We therefore introduce tax campaign month fixed effects λ_t to net out any time-varying components of tax compliance that might affect the analysis.[68] We cluster standard errors at the neighborhood level, the level at which collector pairs were randomly assigned.

potential taxpayers, or they might have a more competitive relationship with the formal state such that they would be unwilling to provide information about household compliance propensities.

[66] These results are presented in Table A8.

[67] Specifically, collectors worked in the 102 holdout sample neighborhoods during campaign months 1, 3, 5, and 7 and in the 78 analysis sample neighborhoods during months 2, 4, and 6. Balan et al. (2021) provides further detail on the staggered rollout of both treatment arms.

[68] In estimating Equation 5, we subtract the average tax compliance across collectors, $\mathbb{E}\left[Y_h(c_1, c_2, v_h)\right]$, as otherwise the level of q_c would not be identified after including month fixed effects.

The coefficient of interest is $\boldsymbol{\alpha}$, the vector of collector fixed effects.[69] 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.[70] We increase the precision of our collector fixed-effect estimator using an Empirical Bayes approach (e.g., Morris, 1983; Kane and Staiger, 2008). More specifically, we consider the OLS estimator $\hat{\alpha}_c^{OLS}$ as an unbiased but noisy measure of q_c — i.e., $\hat{\alpha}_c^{OLS} = q_c + \nu_c$, where ν_c represents noise — and estimate α_c as the posterior mean of q_c:[71]

$$\begin{aligned}\hat{\alpha}_c^{EB} &= \mathbb{E}\left[q_c | \hat{\alpha}_c^{OLS}\right] \\ &= \left(\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}_0^2 + \widehat{\sigma}_c^2}\right)\hat{\alpha}_c^{OLS}\end{aligned}$$

where $\widehat{\sigma}_c^2$ is the variance of $\hat{\alpha}_c^{OLS}$ and σ_0^2 is the variance of q_c. We estimate σ_0^2, using the approach described in Morris (1983).[72]

To motivate our investigation into collector assignments, we illustrate the importance of collectors in shaping tax compliance behavior in this setting. Using the estimated $\hat{\alpha}_c^{EB}$, we find that tax collectors explain 36% of the variance in tax compliance across neighborhoods.[73] By contrast, Fenizia (2019) finds that public sector managers in Italy explain 9%

[69] Without time fixed effects, random assignment of collectors to teammates and to neighborhoods implies that $\alpha_c = q_c$ in large samples. Because we include month fixed effects, α_c may slightly differ from q_c. In particular, if collectors' tax enforcement ability changes over time, then 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, noting that in case this assumption is violated, we are still able to identify a meaningful measure of collectors' enforcement ability.

[70] Even though neighborhoods are randomly assigned to collector pairs, implying that neighborhood characteristics are identically distributed across collectors, the differences in neighborhood characteristics across collectors could be large due to the small number of neighborhoods assigned to each tax collector. Similarly, even though tax collectors are randomly assigned to teammates, and teammates' characteristics are identically distributed across collectors, the difference in teammates' characteristics could be large across collectors due to the small number of teammates assigned to each collector.

[71] We assume that ν_c follows a normal distribution with mean zero and variance σ_c^2. We also assume that q_c itself follows a normal distribution with mean zero and variance σ_0^2.

[72] The Empirical Bayes estimator shrinks the OLS estimator towards the common mean, which is normalized to zero, by a factor that depends on the sample noise σ_c^2 of $\widehat{\alpha}_c$ and the variability of α_c across collectors, captured by σ_0^2. When the ratio σ_c^2/σ_0^2 is large, the OLS estimator is relatively imprecise in comparison to the heterogeneity in ability σ_0^2. In that case, we shrink the estimator closer to the common mean. Conversely, if this ratio is small, the OLS estimator is relatively precise and closer to the Empirical Bayes estimator. The Empirical Bayes estimator has a smaller mean squared error than the OLS estimator, so it will yield, on average, better predictions than the OLS estimator (Morris, 1983).

[73] Specifically, we compute $Var(\hat{\beta}_c^{EB})/Var(\overline{Y}_n)$, where $Var(\hat{\beta}_c^{EB})$ is the sample variance of the Empirical Bayes estimates across collectors and $Var(\overline{Y}_n)$ is the sample variance of the average tax compliance across neighborhoods.

of the total variation in the efficiency of filing insurance claims, and Best et al. (2019) find that bureaucrats who manage procurement processes in Russia explain 24% of the variation in quality-adjusted public procurement prices. A likely explanation for why our estimate is larger is that field-based tax collectors in Kananga have a high degree of discretion over key dimensions of tax collection: the intensity of enforcement effort, the tactics and arguments used to persuade households to pay, the possibility of paying a bribe, etc. This contrasts with office-based positions in government bureaucracies, which are more easily monitored by supervisors and governed by rules intended to standardize processes. Yet field-based tax collectors/inspectors are central to the operations of most tax authorities in developing countries (e.g., Khan et al., 2016; Cogneau et al., 2020; Krause, 2020; Okunogbe, 2021) and thus worthy of closer scrutiny.

To define collector types, we rank and partition collectors into discrete groups using $\hat{\alpha}_c^{EB}$. This dimensionality reduction allows us to estimate the average compliance function non-parametrically in Section 6.3.[74] Our main specification defines two types of collectors: low-ability, L, or high-ability, H, depending on their $\hat{\alpha}_c^{EB}$ rank, $r_c = \text{rank}(\hat{\alpha}_c^{EB})/N_c$. Collectors with $r_c < 0.5$ are categorized as low-type, while collectors with $r_c > 0.5$ are categorized as high-type.

This non-parametric approach to ranking collectors — based on the compliance they achieved across randomly selected neighborhoods — remains agnostic about the underlying average tax compliance function.[75] It is possible that assuming that tax collector fixed effects are additive constitutes a misspecification.[76] However, this would not compromise our objective, which is to define a sensible metric for collector type that enables us to analyze the returns to the optimal assignment of collectors while making as few assumptions as possible and without imposing a specific functional form on the average tax compliance

[74]This approach as also used by Bhattacharya (2009) and Graham et al. (2020a) in the context of optimally assigning teachers to students, and by Carrell et al. (2013) in the context of assigning students to platoons at the Naval Academy.

[75]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. As randomization ensures that horizontal differences — in this example, ethnicity — are uncorrelated with collector assignments, the observed compliance rank will capture a meaningful signal of collector effectiveness to support estimation of an optimal assignment based on this measure of collector ability.

[76]Any paper using a mover design (e.g., Abowd et al., 1999) also implicitly assumes that types are additive when estimating worker and firm fixed effects.

function.[77,78]

High-type collectors differ from low-type collectors in many ways beyond their ability to collect taxes (Table A3). 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.

In Section 8.2, we discuss robustness to alternative definitions of collector types. The results are qualitatively similar when tax collectors are partitioned into three categories based on their rank r_c.[79] Results are also similar when we estimate collector type in the holdout sample using baseline collector characteristics, an approach that might be more easily employed by governments than estimating a fixed effects model. [80]

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 (6,904 properties). We follow Bhattacharya (2009) and Graham et al. (2020a) and estimate it non-parametrically using the following regression:

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

[77]By contrast, if our objective were to precisely estimate the value added (i.e., fixed effect) associated with each tax collector, potential misspecification would be greater cause for concern. Misspecification would also complicate our estimate of the share of the variance in tax compliance that is explained by collectors (36%), though this concern is not unique to our setting and applies in general to work relying on mover designs. Following the literature, we view this estimate as a first-order approximation (and not the primary focus of our empirical analysis).

[78]Given that Section 7 shows complementarity in collector type, a natural question is whether our approach to ranking collectors using separable fixed effects could constitute a source of bias in our ultimate estimates. As noted, we prefer our non-parametric approach because it remains agnostic about functional form and thus remains valid under different possible compliance (production) functions. In the particular case of complementarity in collector type, our estimated collector fixed effects would be inflated among the high-type collectors, who would look like better individual collectors than they actually are because part of their observed "effectiveness" comes from the complementarity. However, this potential source of bias would not jeopardize the application of our collector-type estimation approach because we do not seek to recover the exact causal effect associated with each collector; rather, we seek a sensible ranking of them. In this case, upward bias on high-type collectors would not impact our ranking because we have random assignment of collectors into teams and measure the average compliance levels across multiple neighborhoods in which they work.

[79]While increasing the number of collector types mechanically improves the efficiency of collector assignment, it also leads to noisier estimates of collector types and of the optimal assignment (Table A7). For this reason, the main results presented in Table 2 use two collector types.

[80]These results are presented in Table A7.

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 five dummies: $(H,H,h),(L,H,h),(L,L,h),(H,H,l),(L,H,l)$, reflecting matches of collectors and households of two types ($A=\{L,H\}$ and $V=\{l,h\}$).[81] We also include campaign month fixed effects λ_t, as discussed above. Standard errors are clustered at the neighborhood level.

6.4 Impact of the Optimal Assignment

We now turn to the estimation of the optimal assignment function f^*. Again following Bhattacharya (2009) and Graham et al. (2020a), we use our estimates of the average tax compliance function $\hat{\beta}(a_1,a_2,v)$ and plug them into the empirical analog of the Optimal Assignment Problem (Problem 1):[82,83]

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 estimated optimal assignment function and average tax compliance function to obtain the ARE estimator:

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

Our main specification reports conventional standard errors clustered at the neighborhood level, as discussed above. However, collector type might be estimated with noise in

[81] The intercept is not identified when campaign month fixed effects are included, so we need to exclude one of the type dummies. Here we exclude the dummy for matches of type (L,L,l).

[82] As noted, in Section 8.2 we examine a government maximizing revenue, or revenue net of bribes, in lieu of tax compliance.

[83] Although $\hat{\beta}$ identifies Y up to a constant, the solution to Problem 1 is the same if we substitute Y for $Y+c$ for any constant c. To see that, note that $(\mathbf{Y}+c)'\mathbf{f} = \mathbf{Y}'\mathbf{f} + c\mathbf{Y} = \mathbf{Y}'\mathbf{f} + c$, where the last equality derives from the fact that f is a probability mass function.

the first step of our analysis (Section 6.2) due to sampling error, which would mean that clustered standard errors are too small. To take into account the sampling error associated with the estimation of tax collector type, we also report standard errors from Bayesian bootstrap re-sampling (Rubin, 1981) at the neighborhood level when estimating average tax compliance and revenue by collector and household type (Figure A4) and the effect of the optimal assignment on tax compliance and revenue (Table A6). Because we identify collector type by exploiting their assignment to a relatively small set of neighborhoods, Bayesian bootstrap — in which we resample weights for neighborhoods in each iteration and use a weighted least squares estimator — is better suited to the context than the standard bootstrap.[84]

7 Optimal Assignment

7.1 Characterizing the Optimal Assignment

We begin by characterizing the composition of tax collector teams and of team-to-household matches under the optimal assignment.

Ex ante, it is not obvious what assignment function would maximize tax compliance and revenue.[85] If collection from households characterized by a high tax payment propensity simply involved showing up and soliciting payment, then it could be optimal to assign them to low-ability collectors. Alternatively, if collection from high tax payment propensity households requires persuasion skills or conscientiousness in making follow-up visits at times when owners have liquidity, then the government may do better by assigning them to high-ability collectors.

[84]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 (Efron, 1982). 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. By randomly sampling neighborhood weights in each iteration, which does not require dropping neighborhoods altogether, the Bayesian bootstrap is preferable in our setting.

[85]Past empirical work also reaches mixed conclusions. Carrell et al. (2009) predicted negative assortative matching of students would optimize test scores, but Carrell et al. (2013) found contrasting evidence when implemented in real life. Bhattacharya (2009) finds that positive assortative matching of students in dorms has little average impact on test scores. Graham et al. (2020a) and Aucejo et al. (2019) both find evidence of modest complementarities in teacher and student characteristics. Marx et al. (2021) find that the effect of ethnic homogeneity on productivity is positive on a peer-to-peer level but negative on a worker-to-manager level.

Similarly, in forming teams, if only one high-ability collector is required to ensure that all essential tasks are completed, then one might expect that pairing a high-ability with a low-ability collector (mixed teams) would maximize compliance. However, there could also be scope for complementarity between collectors' effort or skills that would justify grouping high-ability collectors together and low-ability collectors together (homogeneous teams).

7.1.1 Collector-to-Collector Assignment

According to our estimation approach, the optimal assignment of collectors to teams involves positive assortative matching. Specifically, the provincial tax ministry would only form pairs of high-type collectors (*H-H* teams) — 50% of total teams — or pairs of low-type collectors (*L-L* teams) — remaining 50% of teams. It would never form pairs of mixed type, *L-H* teams (Figure 1). This contrasts with the status quo which is characterized by 50% of *L-H* pairs, 25% of *L-L* pairs, and 25% of *H-H* pairs, due to random assignment.[86]

Such positive assortative matching derives from complementarities in collector type in the average tax compliance function (Figure 2). Assigning a low-type collector to a high-type teammate increases tax compliance by 1.5 percentage points relative to assignment to another low-type teammate. By contrast, assigning a high-type collector to a high-type teammate increases compliance by 9.5 percentage points relative to assignment to a low-type teammate. A formal test of complementarity confirms that the average tax compliance function is convex in collector type ($p = 0.037$).[87] Given that tax revenue is equal to compliance multiplied by a constant (the tax rate) in this context, this same pattern of complementarity in collector type mechanically appears when studying average tax revenue per owner ($p = 0.090$, Figure A2).[88] We find even stronger evidence of complementarity

[86] Under the optimal assignment, the random reshuffling of collectors into new teams each month — to prevent the emergence of collusion/covering — could occur in a similar fashion as under the status-quo assignment. For example, every campaign month, the 17 high-type (low-type) collectors could be randomly matched with a high-type (low-type) teammate. During the course of the six-months long tax campaign, each high-type (low-type) collector would thus be randomly assigned to 6 high-type (low-type) teammates.

[87] We test that $Y(a_1, a_2, v)$ has increasing differences in collector type, i.e., that $Y(H, a, v)$-$Y(L, a, v)$ increases with collector's type a. Formally we test the hypothesis H_1: $[Y(H, H, v)$-$Y(L, H, v)]$ - $[Y(H, L, v)$-$Y(L, L, v)] > 0$ against the null hypothesis H_0: $[Y(H, H, v)$-$Y(L, H, v)]$ - $[Y(H, L, v)$-$Y(L, L, v)] \leq 0$. For simplicity we only report the p-value of this test for high-type households ($v = h$). A more general test for non-linearity consists in testing $[Y(H, H, v)$-$Y(L, H, v)]$-$[Y(H, L, v)$-$Y(L, L, v)] \neq 0$ for $v = h$. Such a test has the advantage of allowing to detect both increasing and decreasing differences in collector type. Results for this test confirm that the tax compliance function is non-linear in collector type ($p = 0.074$). In the remainder of the paper we primarily report tests for complementarity (i.e., increasing differences) to facilitate direct comparisons of patterns in mechanism-related outcomes with the observed complementarity in compliance.

[88] Tax revenue is obtained by multiplying tax compliance by the tax liability and thus mechanically results in

when controlling for property characteristics ($p = 0.019$ for compliance and $p = 0.053$ for revenue, Figure A3).[89]

7.1.2 Collector-to-Household Assignment

The optimal assignment also involves positive assortative matching on the collector-to-household dimension. Under the optimal assignment, the government would only assign *H-H* teams to high-type households and *L-L* teams to low-type households. Specifically, it would assign 75% of high-type households to *H-H* teams, 25% of high-type households to *L-L* teams, and all low-type households to *L-L* teams (Figure 1). Some high-type households are assigned to *L-L* teams because 67% of households are high-types, while only 50% of collector pairs are high types, and the workload constraint means the *H-H* teams cannot take on more total households than under the status quo assignment.

Positive assortative matching here reflects complementarities in collector-to-household match type. Assigning an *L-L* team to a high-type household would increase compliance by 3.5 percentage points relative to assigning the team to a low-type household. By contrast, assigning an *H-H* team to a high-type household would increase compliance by 13.4 percentage points relative to assigning the team to a low-type household. A formal test of complementarity confirms the convexity in the compliance function with respect to collector-to-household match type ($p < 0.001$).[90] As before, the same pattern of complementarity in collector-to-household match type applies to the average tax revenue per owner ($p = 0.004$, Figure A2). Again, we find stronger evidence of complementarity when controlling for property characteristics ($p < 0.001$ for compliance and $p = 0.002$ for revenue, Figure A3).[91]

7.2 Mechanisms

Before turning to the impact of the optimal assignment policy (Section 8), we first explore mechanisms behind the complementarities in collector-to-collector and in collector-

less precise estimates and slightly weaker evidence of convexity in collector type.

[89] Additionally, complementarity tests using standard errors from Bayesian bootstrap re-sampling to account for sampling errors associated with the estimation of tax collector type return similar though slightly weaker evidence of convexity ($p = 0.109$ for compliance and $p = 0.174$ for revenue, Figure A4).

[90] We test that $Y(H,H,v)$-$Y(L,L,v)$ increases with household type v. Specifically, we test the hypothesis H_1: $[Y(H,H,1)$-$Y(L,L,1)]$ - $[Y(H,H,0)$-$Y(L,L,0)] > 0$ against the null hypothesis H_0: $[Y(H,H,1)$-$Y(L,L,1)]$ - $[Y(H,H,0)$-$Y(L,L,0)] \leq 0$ and report the associated p-value. A general test for non-linearity — i.e., that $[Y(H,H,v)$-$Y(L,H,v)]$-$[Y(H,L,v)$-$Y(L,L,v)] \neq 0$ — also confirms non-linearity ($p = 0.001$ for compliance and $p = 0.008$ for revenues).

[91] Complementarity tests using standard errors from Bayesian bootstrap re-sampling also confirm convexity in compliance ($p = 0.004$) and revenue ($p = 0.013$) (Figure A4).

to-household match type. We focus here on two key potential mechanisms: collector skill and effort.[92]

7.2.1 Collector Skill

A first possible mechanism is that *H-H* teams were more skillful in convincing households to pay. The in-person mode of tax collection in Kananga left much at the discretion of collectors, including what types of messages and other persuasion techniques to use. It could be that high-type collectors are 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.

First, we study post-taxation beliefs about enforcement and tax morale. If *H-H* teams 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* teams would perceive a higher probability of enforcement among delinquent properties after the campaign.[93] Using midline survey data (collected after tax collection was completed in each neighborhood), we find that high-type collectors cause households to perceive a higher likelihood of sanctions for tax delinquency on average. However, *H-H* teams do not differentially increase property owners' beliefs about sanctions relative to *L-H* teams (Figure A5, Panel A).[94] Similarly, *H-H* teams do not appear to differentially increase citizens' perceptions that tax revenues are spent on public goods relative to *L-H* or even *L-L* teams (Figure A5, Panel B).[95]

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, etc. We therefore examine if *H-H* teams differentially relied on certain messages compared to *L-L* and *L-H* teams but find

[92] Appendix Section A5 explores other possible mechanisms.

[93] Alternatively, if they were more effective in appealing to households' tax morale, we would expect that those assigned to *H-H* teams to be more confident that tax revenues would be spent on public goods.

[94] A complementarity test fails to reject that citizens' beliefs about sanctions are non-convex in collector-to-collector match type ($p = 0.964$) or collector-to-household match type ($p = 0.268$).

[95] A complementarity test fails to reject that citizens' beliefs that tax revenue is spent on public good are non-convex in collector-to-collector match type ($p = 0.993$). Though we find suggestive evidence of complementarity in the collector-to-household match type ($p = 0.091$), it is driven by *L-L lowering* citizens' perceptions about public spending when assigned to collect from high-type households relative to low-type households rather than *H-H* teams increasing such perceptions.

little evidence of complementarities in collector type in this dimension (Figure A6).[96,97] 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).

7.2.2 Collector Effort

A second explanation is that high-type collectors exerted greater effort when matched with another high-type collector (e.g., Mas and Moretti, 2009; Brune et al., Forthcoming).[98] To explore this possibility, we investigate the number of distinct days and the number of 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,[99] and (*ii*) the date and time of visits that led to a tax payment, which is systematically recorded by the tax receipt data.[100] Although collectors were supposed to work for an entire month in each assigned neighborhood, whether they actually did so and for how long were left at their discretion. According to both measures, we find that *H-H* teams exerted disproportionately more effort than *L-L* or *L-H* teams (Figure A7).[101]

While the chalk dates and tax receipt data offer objective measures of collector effort, they may also be recorded with error that could lead us to overstate the extent to which *H-H*

[96]Messages used by the tax collectors to convince property owners to pay included emphasizing: sanctions (Panels A–B), public goods provision (Panels C–D), showing trust in the government (Panel E), the importance of paying the tax (Panel F), the legal obligation to pay (Panel G), the potential social embarrassment of evading taxes (Panel H), and other threats for tax delinquents (Panel I).

[97]Complementarity tests systematically fail to reject that the messages used by the tax collectors are non-convex in collector-to-collector type (p-value between 0.219 and 0.993) or collector-to-household type (p-value between 0.149 and 0.794)

[98]A simple model that generates complementarity in effort provision is as follows. Assume the tax compliance probability $y = e_1 + e_2$ is a function of the effort exerted by each collector, e_1 and e_2. Additionally, assume that $e_i = a_i + \beta a_i e_j$ for $(i, j) = (1, 2), (2, 1)$, where a_i is collector i's type. This effort function could easily result from a utility function where the effort of a collector depends on the effort of the teammate and where the marginal effect of the teammate's effort is increasing in collector i's type a_i.

[99]Enumerators recorded these dates in the midline survey.

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

[101]A test of complementarity in collector and household type shows convexity in days and hours worked with respect to collector-to-collector match type ($p = 0.032$ for days worked and $p = 0.051$ for hours worked) and collector-to-household match type ($p = 0.078$ for days worked and $p = 0.097$ for hours worked).

teams' performance is explained by effort.[102] For this reason, we also examine midline survey data asking the property owners about the number of visits made by the 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* teams indeed conducted more visits than *L-L* and *L-H* teams. They did so both on the extensive margin (Figure A8, Panel A) — the share of households that received any post-registration visits — 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-to-collector and collector-to-household type.[103,104]

Why would collecting taxes on more distinct days and for more hours increase tax compliance? One explanation is that it might have increased the chances that property owners had the cash on hand to pay the tax when the collectors solicited payment.[105] It is well-documented that liquidity constraints impact property tax compliance, even in middle- and high-income countries like Mexico and the United States (Brockmeyer et al., 2020; Wong, 2020). If property owners in Kananga, a low-income setting, faced time-varying cash-on-hand constraints, then collector visits on different days, and on different times over the course of the day, might have increased the probability that property owners had cash on hand when collectors visited.

We provide two pieces of evidence consistent with this cash on hand interpretation. First, we examine heterogeneity in collector effort by the neighborhood employment rate. Property owners with some source of employment are more likely to have cash on hand than

[102] For the chalk marks, tax collectors might have forgotten or chosen not to record their visits, and such omissions could vary by collector type. In addition, the receipt data only capture visits that resulted in tax payments. If H-H teams collected payments on more days and hours for other reasons than effort, then relying only on the receipt-based measure could overestimate the role of effort. Note that the chalk date was meant to be recorded for *all* visits, which already alleviates this type of endogenous measurement error concern to some extent.

[103] A complementarity test fails to reject that the visit indicator and the number of visits is non-convex in collector-to-collector match type ($p = 0.520$ and $p = 0.131$, respectively) or collector-to-household match type ($p = 0.712$ and $p = 0.336$, respectively).

[104] Another potential explanation, which would be consistent with collector effort as the mechanism, is that *L-L* and *L-H* teams exempt more properties from the tax, which then translates into lower levels of tax compliance. To investigate this issue, we include exempted properties in the holdout and analysis sample and estimate tax exemption status by collector and household type (Figure A9). Tax exemption by collectors does not appear to exhibit convexity in collector-to-collector or in collector-to-household match type and is thus unlikely to explain the complementarities documented in section 7.1.

[105] 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 are not convex in collector type (Section 7.2.1).

the unemployed. If the additional days and hours of tax collection by *H*-*H* teams boosted tax compliance by relaxing time-varying cash-on-hand constraints, then the increase in collector effort should have been concentrated in neighborhoods with higher employment rates where such constraints are less likely to always bind. The data bear out this prediction (Figure A10). Second, in an economy with many day laborers, property owners might be more likely to have cash on hand later in the day. Collecting taxes later in the day would thus boost tax payment if cash-on-hand constraints are a key impediment to compliance. To test this prediction, we use the receipt data to estimate the average time of collection across collector types. We find suggestive evidence that *H*-*H* teams did more of their tax collection later in the day compared to *L*-*H* or *L*-*L* teams (Figure A11). *H*-*H* teams 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.

A natural question is why all collector teams did not simply work for longer hours if this could relax household liquidity constraints and boost tax revenue (and thus collector compensation)? Anecdotal evidence suggests that collectors and their supervisors were aware that working on more days and visiting later in the day could increase the chances that cash on hand constraints were non-binding and lead to more tax payments. This point was stressed during collector training sessions by tax ministry supervisors when advising collectors on field strategies. Thus, rather than a knowledge gap between *H*-*H* and other teams, the mechanism more likely concerns coordination between collectors. As noted above, collectors viewed tax collection as a joint task and exhibited a strong preference to work in teams rather than alone.[106] Thus, if their partner were unreliable and did not show up for work on time (or at all), even a high-type collector might likely choose not to work that day. This production process may thus exhibit O-ring properties (Kremer, 1993), in which either collector failing to show up for work leads tax revenue for that team to go to zero. Such coordination problems are a common feature of joint production tasks (Alchian and Demsetz, 1972; Olson, 1989), which, as noted above, often characterize the work of frontline agents in the public and private sector in developing countries (e.g., Burgess et al., 2010; Khan et al., 2016; Ashraf and Bandiera, 2018; Banerjee et al., 2021; Marx et al., 2021).

In sum, *H*-*H* collector teams appear to achieve disproportionately higher compliance

[106] Indeed, we observe very few collectors working alone in the data. Collectors explained this preference by arguing that, relative to a solo collector, a pair of collectors would lead households to perceive a payment as more likely to reach the government account and non-payment to face a higher risk of enforcement.

than *L-H* and *L-L* teams 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* teams likely reflects their ability to solve the coordination problem inherent in team-based tax collection, rather than by overcoming knowledge constraints or other frictions.[107]

8 Impact of the Optimal Assignment

We now estimate the increase in tax compliance and revenue under the optimal assignment policy, examine a series of robustness checks, explore distributional implications, and compare the effect of the optimal assignment with the impact of alternative policies such as collector selection and wage increases.

8.1 Main Results

According to our estimation approach outlined in Section 6, the optimal assignment policy would increase tax compliance by 2.941 percentage points ($p = 0.024$) (Table 2, Row 1, Column 1). This represents a 37% increase in compliance relative to the status quo assignment. The policy would also lead to a 54.471 Congolese Franc (CF) increase in tax revenue per owner ($p = 0.074$), a 26% increase (Column 2). The effect of the optimal assignment remains significant when controlling for property characteristics (Table A4, Columns 3 and 4), when including exempted properties (Table A5, Columns 3 and 4), and when accounting for sampling errors associated with the estimation of tax collector type (Table A6, Columns 3 and 4).[108] As discussed in the previous section, these increases in compliance and revenue reflect the complementarities in collector-to-collector and collector-to-household match type (Figure 2), which are fully exploited under the optimal assignment policy.

To assess how each margin of the optimal assignment — collector-to-collector and collector-to-household — contributes to the total effect of the policy, we estimate the return to alternative policies optimizing on each of these margins separately (Table 2, Rows

[107] In Appendix Section A5, we consider other possible mechanisms, including homophily and social incentives. We find little evidence that these channels account for *H-H* teams' greater effectiveness.

[108] The p-values associated with the effects of the optimal assignment policy are slightly higher when estimating standard errors of the tax compliance function from Bayesian bootstrap re-sampling at the neighborhood level: 0.080 for tax compliance (Table A6, Columns 3) and 0.150 for tax revenue (Table A6, Columns 4). The larger standard errors result from taking into account sampling error when estimating collector types.

2–3).[109] For instance, if the government optimizes the assignment of collectors to teammates but assigns teams to households at random, it 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. By comparison, if the government optimizes the assignment of collectors to households but forms collector teams at random, it 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. Both dimensions of assignment appear important in raising tax compliance, and the government does substantially better by jointly optimizing.

8.2 Robustness Checks

We examine a number of alternative estimation approaches and robustness checks, which reinforce our main results.

Alternative Definition of Collector Type. The optimal assignment thus far relies on the government's ability to estimate collector type using their performance (in the holdout sample) during the tax campaign. However, in practice the government might seek to predict types by correlating observable collector characteristics with performance in a past tax campaign. While this approach to estimating collector type might be less precise, it has the practical advantage of allowing the government to predict type for new collectors.

We implement a version of this approach by predicting collector type using an OLS regression of tax compliance on collector characteristics in the holdout sample. We then define the predicted collector's type (high and low) based on whether they are above or below the median in terms of their predicted tax compliance in the holdout sample.[110] With this alternative estimation of collector type, we still observe complementarity in collector type and in collector and household type for tax compliance (Figure A13) and tax revenue (Figure A14).[111] Similarly, the optimal assignment would still increase tax compliance by 2.688 percentage points ($p = 0.030$) and tax revenue per owner by 56.926 CF ($p = 0.048$), a 34% and 28% increase, respectively (Table 2, Columns 3 and 4).

Three Collector Types. We also show results when partitioning tax collectors into three

[109] Figure A12 characterizes the resulting assignments of these uni-dimensional optimized policies.

[110] We focus on a subset of the collector characteristics described in Panel A of Table A3: gender, age, ethnicity, level of education (never been to school, kindergarten, primary, secondary, university), math score, literacy (in Tshiluba and French), income, and possessions.

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

types — low (L), middle (M), and high (H) — instead of two. The optimal assignment still involves positive assortative matching due to complementarities in collector-to-collector and collector-to-household match type for tax compliance (Figure A15) and revenue (Figure A16).[112,113] Moreover, the optimal policy would have larger effects, increasing compliance by 4.411 percentage points ($p = 0.032$) or 55% relative to the status quo assignment (Table A7, Column 1) and tax revenue per owner by 62.212 CF ($p = 0.202$) or 30% (Column 2). With a finer partition of types, the estimated impacts of the optimal assignment are larger but also noisier due to fewer observations for each type of collector.

Alternative Definition of Household Type. The optimal assignment policy thus far assumes that the government has access to the neighborhood chief's prediction of each household's ability to pay. In settings without such chiefs, or in which chiefs have a more competitive relationship with the formal state (Henn, 2020), the government might prefer to 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 the predicted household's type (high and low) based on whether they are above or above the median in terms of their predicted tax compliance in the analysis sample.[114] Under this alternative definition, we still find evidence of complementarity in collector type and in collector and household type for tax compliance (Figure A18) and tax revenue (Figure A19), although the standard errors are larger.[115] Tax compliance would increase by 2.759 percentage points ($p = 0.067$) and tax revenue by 50.417 CF per owner ($p = 0.148$) under the optimal assignment, a 34% and 24% increase, respectively (Table A8, Columns 3–4).

Revenue-Maximization Objective. Thus far, we have assumed that the government's objective function is to maximize tax compliance.[116] However, a government might in-

[112] We define these groups so that the bottom tercile of collectors in terms of ability q_c are of type L, the top tercile are of type H, and the intermediate tercile are of type M.

[113] As is the case with two collector types, teams with a L-type collector perform considerably worse. The optimal assignment thus consists of constituting *L-L* teams and *M-H* teams. The pairing of M-type collectors with H-type collectors is driven by the fact that *M-H* teams significantly outperform *M-M* teams and somewhat outperform *H-H* teams.

[114] We focus on the characteristics described in Panel A of Table 1: distance to state buildings, distance to health institutions, distance to education institutions, distance to roads, distance 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.)

[115] As a consequence, the optimal assignment again involves positive assortative matching (Figure A17).

[116] We focus on tax compliance rather than revenue as the government's objective as revenue is equal to tax

stead prefer to maximize tax revenue, or tax revenue net of bribes. The results are similar when adopting these alternative objectives (Table A9). The revenue-maximizing assignment policy, for instance, would increase tax revenue per owner by 61.014 CF ($p = 0.020$) or 30% relative to the status quo assignment (Column 1). This is in fact slightly larger than the comparable estimate (54.471 CF) from the compliance-maximizing policy, though the two are not statistically different. We also find similar effects, albeit smaller in magnitude, when the government aims at maximizing tax revenue net of bribe payments per owner (Column 3).

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). Assigning tax collectors on the neighborhood level might therefore be more policy relevant, even if it likely reduces the effectiveness of the collector-to-household matching. In Table A10, we therefore consider two neighborhood-level optimal assignment policies: categorizing neighborhoods as high or low type based on (*i*) their share of high and low type households (Columns 1–2), (*ii*) or their total number of high and low type households (Columns 3–4).[117] The optimal assignment policy would increase tax compliance by 1.764 percentage points ($p = 0.085$) under (*i*) and by 2.906 percentage points ($p = 0.048$) under (*ii*). This latter estimate is just shy of that from our main specification involving household-level assignments (2.941 percentage points). One reason is that policy (*ii*) in fact partly relaxes the constraint on the marginal distribution by collector type in Equation (3): high-type collectors in effect receive more households under this assignment than under the status quo.[118] Taking neighborhoods' size into account thus allows the government to increase the number of high-type households assigned to *H*-*H* teams — and thereby 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 create an

compliance multiplied by the tax rate and thus potentially a noisier empirical object for the optimization problem.

[117] Appendix Section A6 provides more details about the estimation of these neighborhood-level assignments.

[118] One concern is that a larger assignment load could cause collector exhaustion and lower productivity, meaning we would be overestimating the impact of this counterfactual policy. However, as discussed in Section A8.1, we find no evidence that collectors face binding time constraints or that they visit a smaller share of households in larger neighborhoods. These observations suggest that collectors would be able to work in larger neighborhoods (with more households on average) without lowering their productivity.

overfitting problem, i.e., we may be selecting the optimal assignment based on noise.[119] In particular, because we select the best of many possible assignments using tax compliance by match type, which is estimated with noise, the effect of the optimal assignment might be biased upward. This is an example of the "winner's curse" in optimization problems (Andrews et al., 2019).

We implement the methodology introduced by Andrews et al. (2019), which relies on optimal confidence intervals and median-unbiased estimators that are valid conditional on the policy selected and so overcome this winner's curse. A challenge in applying their estimator in our context is that it only applies to discrete policy spaces.[120] However, the policy space in the optimal assignment problem is non-finite, consisting of all 6-dimensional distribution probabilities (i.e., a 5-simplex) satisfying the constraints in Problem 1. Fortunately, we can reduce the policy space to a set of three policies in two steps.[121] 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, 1948) — 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.

Our results are robust to possible winner's curse bias.[122] Table A11 reports the conditional and hybrid median-unbiased estimators and optimal confidence intervals proposed in Andrews et al. (2019) using the set of solutions discussed above.[123] Reassuringly, the

[119] Ex ante, we would not anticipate this problem being very severe in our context because we have so few variables in our model: five dummies for the different combinations of collector and household types, plus dummies for the three months of campaign activity in the analysis sample. This essentially restricts the degrees of freedom we have to fit noise.

[120] Andrews et al. (2019) focuses on common empirical problems that involve inference on parameters selected through optimization over a finite set of candidates. For example, in a randomized trial with multiple treatments, one might want to learn about the true average effect of the treatment that performed best.

[121] We are deeply grateful to Toru Kitagawa for helpful discussions on how to apply Andrews et al. (2019) in our context.

[122] To our knowledge, the optimal matching literature has not considered the winner's curse as a potential source of bias and our solution to this issue might be of independent interest.

[123] Another solution to this problem would be to split the sample in three instead of two, enabling out-of-sample estimation of the impact of the optimal policy. However, this approach would be costly in terms of power, since we would have to split in two the sample of 78 neighborhoods for which we observe household type.

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. The results are also similar and statistically significant when the objective is to maximize tax revenue instead.

Spillovers and the SUTVA Assumption. The analysis implicitly assumes that potential outcomes by match type would be unaffected by changes in the assignment function. This assumption, sometimes known as the stable unit treatment value assumption (SUTVA), is essential in identifying average compliance under different assignment functions. Section A8 explores potential sources of SUTVA violations in this context.

First, changing collectors' assignments could impact effort levels by match type and thereby affect households' tax compliance. The most worrying scenario for our analysis would be if (*i*) collectors target high-type households for tax visits, and (*ii*) collectors are time constrained, i.e., unable to do all the tax visits that would have a positive return during the month-long campaign period. If both conditions were met, then implementing the optimal assignment could decrease the probability that high-type households are visited and thus reduce compliance.[124] However, while we find some evidence that collectors target visits to high-type households (especially *L*-*L* teams), there is no evidence that tax collectors are time constrained across multiple measures (Section A8.1). Endogenous effort of this form therefore does not appear to be a major concern in our setting.

Alternatively, low-type collectors could become demoralized under the optimal assignment if they realize they will only work with low type teammates and only be assigned low type households in the future. We provide evidence by exploring whether low type collectors assigned to a higher share of low-type teammates and households during the 2018 campaign appear more demoralized in an endline survey using standard motivation questions from the psychology literature. Although low-type collectors have weaker motivation overall, those assigned to work with a higher fraction of low-type teammates or assigned to a higher fraction of low-type properties do not appear to be differentially demoralized. Similarly, we don't find evidence that assignment to low-type teammates or households is associated with a higher probability that the tax collectors dropped out of the tax campaign (Section A8.1).[125] The empirical evidence thus suggests that the assignment of low-type

[124] Under the status quo assignment, high-type households are equally allocated across high- and low-type collectors. By contrast, under the optimal assignment, the majority of high-type households would be assigned to high-type collectors. If collectors were time constrained, high-type households would thus be less likely to be visited under the optimal assignment than under the status quo assignment.

[125] Only three tax collectors in our sample (8.82%) did not complete the full 2018 tax campaign.

collectors to low-type teammates and households under the optimal assignment would not undermine their motivation.[126]

Second, if collectors learn throughout the tax campaign, then changes in the assignment function could alter collectors' learning path and thereby affect potential outcomes by match type. One potentially concerning form of learning for our analysis is learning-by-doing. Learning-by-doing could, for example, justify first assigning collectors to households from whom they will learn the most about tax collection and then deploying collectors to other households.[127] 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 (Section A8.2.1).

Another potentially concerning form of learning is learning from high-type teammates (who may share techniques that are effective at convincing households to pay, for example). For instance, if low-type collectors learned more than high-type collectors from working with a high-type teammate, we would likely overestimate the impact of the optimal policy (because mixed teams would collect more tax than we expect them to). If, by contrast, high-type collectors learned more from working with a high-type teammate, we would likely underestimate the impact of the optimal policy (because imposing positive assortative matching would fuel greater learning). We do find evidence of learning from high-type teammates: past assignment to a high-type teammate has a positive effects on tax collectors' subsequent collection (Section A8.2.2). However, if anything, learning from high-type teammates is more pronounced among high-type collectors, consistent with our main results underestimating the true impact of the optimal assignment. That said, the coefficients are not significant at conventional levels. The most we can confidently infer from this analysis is thus that potential learning from teammates appears unlikely to cause our main estimation to *over*estimate the impact of the optimal policy.

[126]Section A8.1.2 also considers a more extreme case of demoralization: low-type collectors dropping out of the tax campaign under the optimal assignment. We show that the effect of the optimal assignment would remains positive relative to the status quo assignment even for non-trivial dropout rates among low-type collectors (Figure A23).

[127]For instance, if collectors are more likely to learn when assigned to a high-type household, and low-type collectors more so, then the results presented in Section 8.1 would likely overestimate the effect of the optimal assignment as imposing positive assortative matching on the collector-to-household dimension would diminish learning-by-doing among low-type collectors.

8.3 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. Appendix Section A7 describes the estimation of $\mathbb{E}_f[X_h|Y_h = 1]$.

Importantly, 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 A2, 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).

8.4 Comparison with Selection Policies and Wage Increases

8.4.1 Effects of Selection Policies

To benchmark the effect of the optimal assignment, we turn to estimating the increase in tax compliance and revenue associated with two types of selection policies: (*i*) *reallocation policies*, which consist in reallocating a fraction ρ of households previously assigned to low-type collectors to currently employed high-type collectors, and (*ii*) *hiring policies*, which consist in reassigning them to newly hired collectors of average ability instead.[128,129]

Figure 3 shows the effect of both selection policies on tax compliance relative to the status quo assignment (see Section A4 for a description of the ARE estimation for selection policies) when a fraction ρ of the low-type collector households are reallocated to high-type collectors (reallocation policies) or to newly hired collectors (hiring policies). According to Panel A, reallocation policies would surpass the optimal assignment only for large values of ρ. In particular, the provincial tax ministry would have to reassign at least 63% of low-type collectors' households to high-type collectors to achieve the same increase in compliance as under the optimal assignment policy.[130] As shown in Panel B, hiring policies, by contrast, would never rival the optimal assignment. At most, the government could increase tax compliance by 2.237 percentage points if it were to reallocate all low-type collectors' households to newly hired collectors. This is 0.704 percentage points less than the effect of the optimal assignment (2.941 percentage points).[131] 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.[132]

8.4.2 Effects of Collector Financial Incentives

As a second benchmark, we compare the effect of the optimal assignment policy with another intervention frequently used to motivate frontline state agents like tax collectors in developing countries: performance-based financial incentives.[133] We leverage the randomization of 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.[134] This wage structure introduced exogenous vari-

[128]Specifically, we assume that newly hired collectors are low-type with probability 0.5 and high-type with probability 0.5. This type of policy has been explored in the literature on teacher quality (Chetty et al., 2014) and public sector manager quality (Fenizia, 2019).

[129]The formal definition of reallocation policies and hiring policies, using the notation introduced in Section 5, is given in Section A4.

[130]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.

[131]Figure A20 shows similar results when relying on the predicted collector types based on their survey characteristics introduced in Section 8.2.

[132]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 A22), but they might be important when ρ is large.

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

[134]The piece-rate wage associated with each property was written on the property register used by the tax collectors, along with the property tax rate and information about the owner. This randomization is explored

ation in collector compensation within each tax rate, which we use to estimate the effect of stronger collector financial incentives on tax compliance (Figure 4, Panel A).[135] We find that the government would have to increase collectors' piece-rate wages by 69% to achieve the same compliance increase as the optimal assignment.

While the size of this necessary wage increase might be enough to give the government pause in contemplating this policy tool, a further consideration is the policy's cost-effectiveness. Specifically, paying collectors a larger share of the tax revenue they collect will only generate more revenue if the compliance response to stronger performance incentives is sufficiently positive. To explore the cost-effectiveness of increasing performance-based wages, we estimate the effect of changes in collector incentives on tax 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 elasticity of tax compliance with respect to collector wages is not sufficiently large to offset the mechanical decrease in revenues from paying higher piece-rate wages. The likely decrease in tax revenue associated with higher performance-based financial incentives underscores a key advantage of the optimal assignment: its cost neutrality. Given the tightness of budget constraints facing governments in low-income countries, increasing collector performance by optimizing their assignment, which leverages existing human and financial resources, seems a promising approach for raising fiscal capacity.[136]

9 Effects on Secondary Outcomes

The optimal policy maximizes tax compliance, but teams of high-type collectors might be more likely to accept bribes as well as taxes,[137] or they might undermine tax morale if they achieve compliance through threats and coercion. This section explores these potential costs of implementing the optimal assignment policy.

in further detail in Bergeron et al. (2020b).

[135] Specifically, predicted compliance reflects the coefficients from an OLS regression of tax compliance on collector wage (Table A12, Column 1).

[136] We can also compare the effect of the optimal assignment policy with another standard intervention frequently used to stimulate tax compliance in rich and poor countries alike: 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. (2020b). Enforcement messages increase tax compliance by 1.4 percentage points relative to placebo messages about how paying the property tax is important (Table A13), 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 we estimate for the optimal assignment policy.

[137] Recent work on the building of the modern Chinese tax system indeed suggests that leakage often increases in tandem with revenue (Cui, 2021).

9.1 Bribe Payments

We first examine if the optimal assignment policy would impact bribe payment by households. In Kananga's door-to-door tax collection system, collectors have discretion over key margins of tax administration and enforcement — assessment, exemptions, and enforcement intensity — that open scope for collusive bribery: i.e., households making a smaller payment to collectors directly in lieu of paying the full tax liability to the state.[138] 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, as might happen with repeated interactions. Under the optimal assignment policy, the increase in the homogeneity of teams could therefore potentially fuel collusion and bribe payment.

We test this possibility using three survey-based measures of bribes. First, for our preferred measure, 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.[139] 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 4, 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 4, 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 A9, 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 A9, 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 4, Row 3).

[138]The scope for collusion in property taxation exists in many settings (e.g., Khan et al., 2016).

[139]Reid and Weigel (2017) compare this measure with less overt bribe measures and find they line up closely. It does not appear to be taboo to discuss making small payments to officials in Congo. Indeed, nearly half of motorcycle taxi drivers openly admitted to paying bribes at Kananga's roadway tolls using similar local codes for bribes (Reid and Weigel, 2017).

On net, we find suggestive evidence that the optimal assignment would slightly increase bribe payment on the extensive and intensive margin. These increases reflect complementarities in collector-to-collector type rather than complementarities in collector-to-household type (Figure A21).[140] In light of this increase in bribe payments, should the government still implement the optimal policy? On the one hand, it would raise tax revenue and most of the incidence would fall on rich households. On the other hand, it could lead to higher bribe payments. For simplicity let's assume that the government's welfare function is given by $U(T, B) = T - \lambda B$, where T is tax revenue, B is the amount of bribes extracted by the tax collectors, and $\lambda > 0$ is a parameter capturing the marginal rate of substitution between a dollar of taxes and bribes from the government's perspective. Since the optimal assignment is associated with 54.471 CF increase in tax revenue per owner and 13.896 CF higher bribe payment per owner, implementing the optimal assignment would only decrease government's welfare (relative to the status quo assignment) if $\lambda > 3.920$. Thus, the government's marginal disutility from bribe payments would need to be close to four times larger than its marginal utility from tax revenue for the status quo to be preferable to the optimal policy.

9.2 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 4, 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

140 Complementarity tests confirm that the average bribe payment function is convex 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-to-household dimension are more mixed: we fail to reject that the average bribe payment function exhibits complementarity in collector-to-household for (extensive margin) indicators of bribe payments ($p = 0.378$, $p = 0.734$) but not for (intensive margin) amount of bribes paid ($p = 0.055$).

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 4, Panel C).

9.3 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 4, Panel D). It appears to have mixed effects on citizens' view of taxation (Table 4, 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.

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. According to our 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 neighbourhoods appears a promising way for governments in low-income settings 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., 2019) 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, 2009; Kleven et al., 2011; 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.[141] 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 would implement the optimal assignment or whether political economy factors would prevent them from doing so. For instance, if low-type collectors have powerful patrons, they might lobby in favor of mixed teams, which allow them to take home higher revenues by free-riding on their more productive peers. We view understanding how tax authorities respond to information about the potential returns to positive assortative matching under the optimal assignment, as well as the role of political economy constraints in sustaining more idiosyncratic assignments, as fertile ground for future research.

[141] As noted, important recent exceptions include Keen and Slemrod (2017); Khan et al. (2016, 2019); Basri et al. (2019).

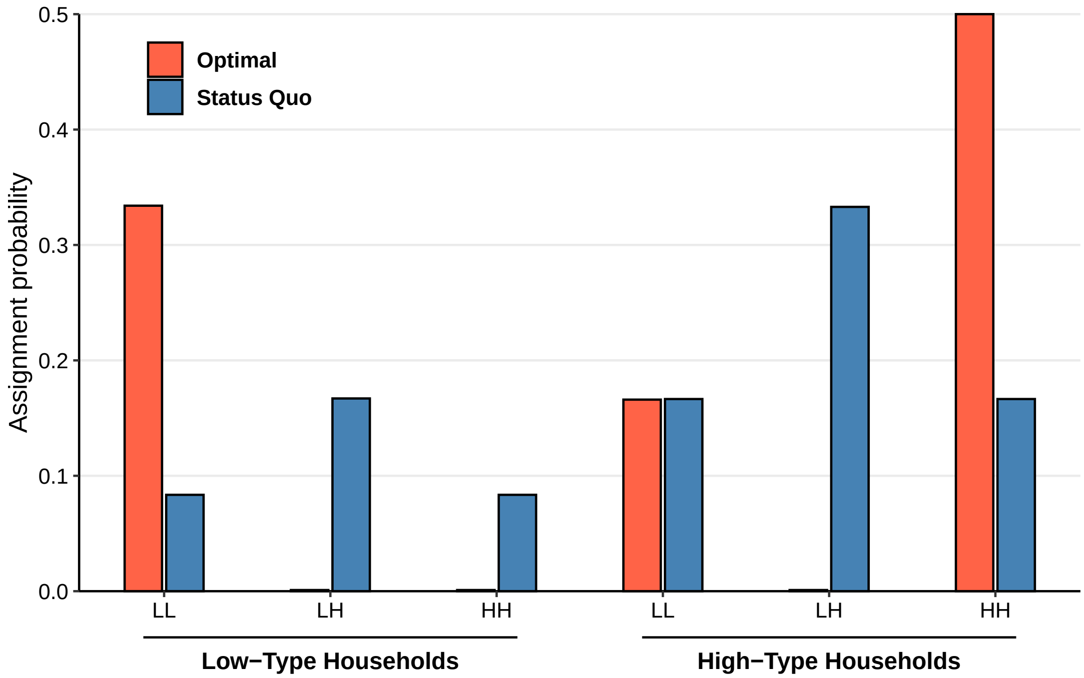
11 Tables and Figures

TABLE 1: 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)

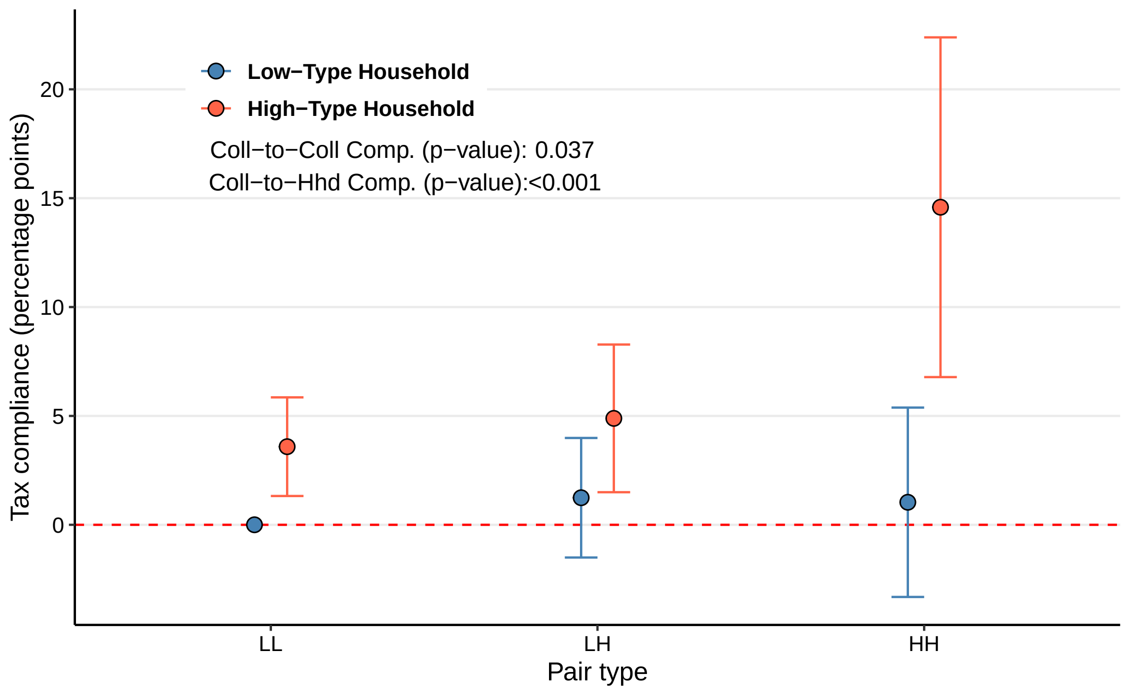
Notes: This table reports coefficients from balance tests conducted by regressing baseline and midline characteristics of properties (Panel A), property owners (Panels B and C), and neighborhoods (Panel D) on an indicator for the type of the collector pair (low-high or LH, high-high or HH, with low-low 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. 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. (2021) and exempted properties. The results are summarized in Section 3.2. The variables are described in detail in Section A9.

FIGURE 1: OPTIMAL VS. STATUS QUO ASSIGNMENTS



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 7.1.1 and 7.1.2.

FIGURE 2: 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 households' 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. We report the p-value associated with a test for compliance exhibiting an increasing difference in collector type and, separately, in collector and household type. We discuss these results in Section 7.1.1 and 7.1.2.

Table 2: Effects of the Optimal Assignment: 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

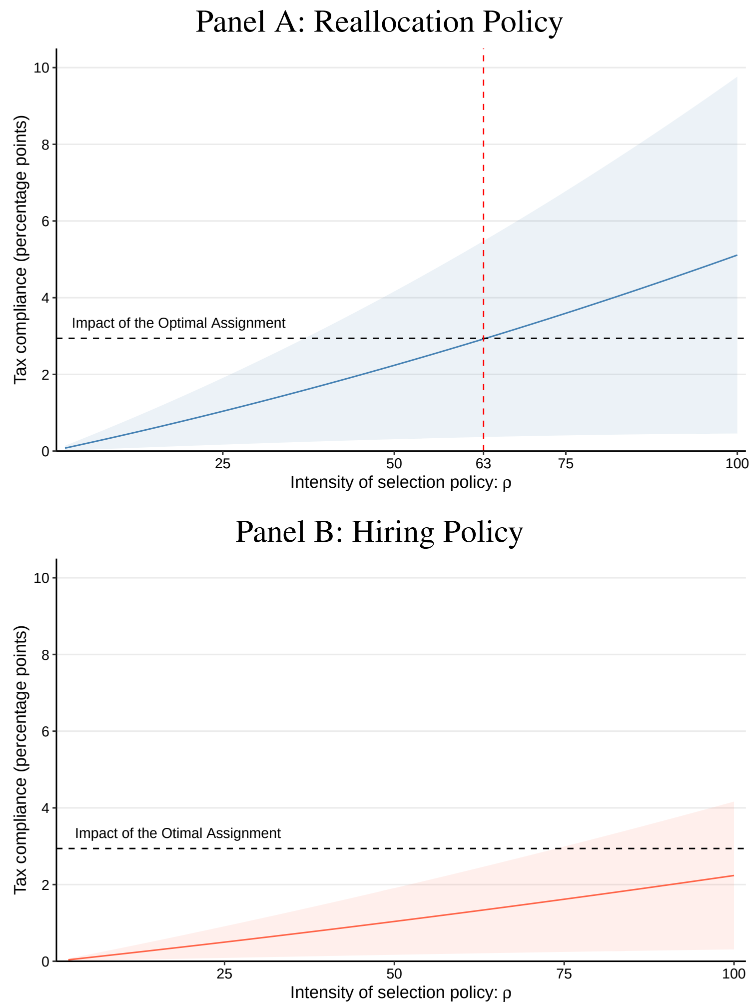
Notes: This table reports estimates of the impact of the counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1 and 3 show estimates for the probability of compliance, i.e., a dummy indicating whether households paid the property taxes (multiplied by 100). 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 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-household and the collector-to-collector dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis; p-values are presented in brackets. 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 Sections 8.1 and 8.2.

TABLE 3: OPTIMAL ASSIGNMENT: INCIDENCE

	Average Taxpayers Optimal Assignment	Average Taxpayers Random Assignment	Average All	Difference (1) vs. (2)	p-value	Observations Taxpayers	Observations All	Sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(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 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. Column 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., 2020a). Panel C analyzes characteristics of the property owner measured at midline and endline. The variables are described in detail in Section A9. We discuss these results in Section 8.3.

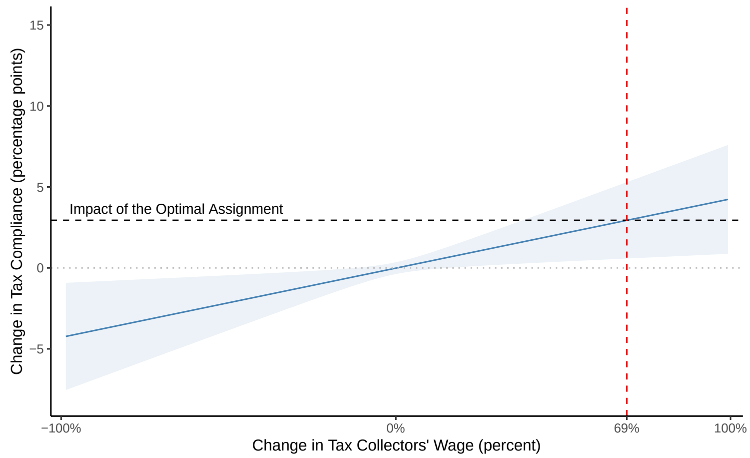
FIGURE 3: EFFECTS OF SELECTION POLICIES



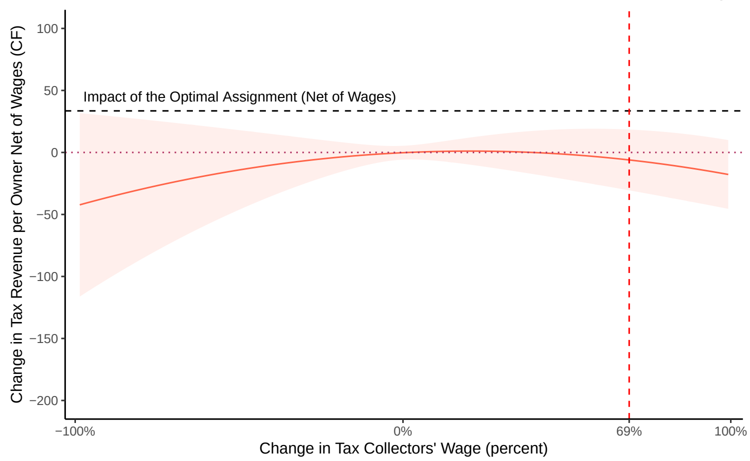
Notes: This figure shows the impact of the selection policies on the probability of tax compliance (y-axis). Selection policies involve reassigning $\rho\%$ (x-axis) of the assignments that a low-ability collector would receive under the status quo assignment to other collectors. Panel A shows the estimated 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 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 2. We discuss these results in Section 8.4.1.

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 A12. 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 impact of the optimal assignment policy on tax compliance as reported in Column 1 of Table 2. In Panel B, the dashed horizontal black line indicates the 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 optimal assignment policy from the effect on tax revenue reported in Column 2 of Table 2. The shaded areas represent 95% confidence intervals using standard errors bootstrapped (with 1,000 iterations). We discuss these results in Section 8.4.2.

TABLE 4: EFFECTS OF 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
Bribe Amount	13.896*	8.408	0.098	30.431	11,732	4,691	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 optimal assignment policy on secondary outcomes. In Panel A, the outcome in row 1 and 2 are self-reported bribe payment and bribe amounts as measured during the midline survey. The outcome in row 3 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. p-values are presented in Column 3. 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 9.

References

Abadie, Alberto and Matias D Cattaneo, "Econometric methods for program evaluation," *Annual Review of Economics*, 2018, *10*, 465–503.

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

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

Andrews, Isaiah, Toru Kitagawa, and Adam McCloskey, "Inference on winners," Technical Report, National Bureau of Economic Research 2019.

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, Zachary Mozenter, and Bates White, "Match Effects in the Teacher Labor Market: Teacher Effectiveness and Classroom Composition," 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," *CEPR Discussion Paper No. DP15138*, 2021.

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," 2021.

Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh, "Improving police performance in Rajasthan, India: Experimental evidence on incentives, managerial autonomy, and training," *American Economic Journal: Economic Policy*, 2021, *13* (1), 36–66.

Basri, M Chatib, Mayara Felix, Rema Hanna, and Benjamin A Olken, "Tax Administration vs. Tax Rates: Evidence from Corporate Taxation in Indonesia," Technical Report, National Bureau of Economic Research 2019.

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," Technical Report 2020.

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

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," Technical Report 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," 2019.

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

— and Pascaline Dupas, "Inferring welfare maximizing treatment assignment under budget constraints," *Journal of Econometrics*, 2012, *167* (1), 168–196.

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," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2021, (2021-15).

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*, 2020.

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

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.

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.

Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaee, "Personalities and public sector performance: Evidence from a health experiment in Pakistan," Technical Report, National Bureau of Economic Research 2015.

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: Evidence from Dakar, Senegal," *Working Paper*, 2020.

Crawford, Vincent P and Elsie Marie Knoer, "Job matching with heterogeneous firms and workers," *Econometrica: Journal of the Econometric Society*, 1981, pp. 437–450.

Cui, Wei, "The Administrative Foundations of the Chinese Fiscal State," 2021.

Dahis, Ricardo, Laura Schiavon, and Thiago Scot, "Selecting Top Bureaucrats: Admission Exams and Performance in Brazil," 2020.

Dantzig, George B., *Linear Programming*, Problems for the Numerical Analysis of the Future, Proceedings of Symposium on Mondern Calculating Machinery and Numerical Methods, UCLA, July 29-31, 1948., 1948.

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," 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," Technical Report, Technical report, Mimeo 2019.

Finan, Frederico, Benjamin A Olken, and Rohini Pande, "The personnel economics of the state," Technical Report, National Bureau of Economic Research 2015.

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," Technical Report, National Bureau of Economic Research 2020.

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

_ , _ , and _ , "Identification and efficiency bounds for the average match function under conditionally exogenous matching," *Journal of Business & Economic Statistics*, 2020, *38* (2), 303–316.

Granovetter, Mark S, "The strength of weak ties," *American journal of sociology*, 1973, *78* (6), 1360–1380.

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*, 2020.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

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," Technical Report, National Bureau of Economic Research 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.

_ , Asim Ijaz Khwaja, and Benjamin A Olken, "Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings," *American Economic Review*, 2019, *109* (1), 237–70.

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez, "Why can modern governments tax so much? An agency model of firms as fiscal intermediaries," *Economica*, 2016, *83* (330), 219–246.

_ , Martin B Knudsen, Claus Thustrup Kreiner, Søren Pedersen, and Emmanuel Saez, "Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark," *Econometrica*, 2011, *79* (3), 651–692.

Kranton, Rachel, Matthew Pease, Seth Sanders, and Scott Huettel, "Identity, Groups, and Social Preferences," *Mimeo, Duke University*, 2013.

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.

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

Luenberger, David G, Yinyu Ye et al., *Linear and nonlinear programming*, Vol. 2, Springer, 1984.

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

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and team performance in a Kenyan organizationâ," *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.

Naritomi, Joana, "Consumers as Tax Auditors," *American Economic Review*, 2019.

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

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

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

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," *The 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.

_, _, and Martin J Williams, "Management, organizational performance, and task clarity: Evidence from ghanaâs civil service," *Journal of Public Administration Research and Theory*, 2021, *31* (2), 259–277.

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

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.

Shapley, Lloyd S and Martin Shubik, "The assignment game I: The core," *International Journal of game theory*, 1971, *1* (1), 111–130.

Tirole, Jean, "Hierarchies and bureaucracies: On the role of collusion in organizations," *JL Econ. & Org.*, 1986, *2*, 181.

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/Revue canadienne des sciences du comportement*, 2009, *41* (4), 213.

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

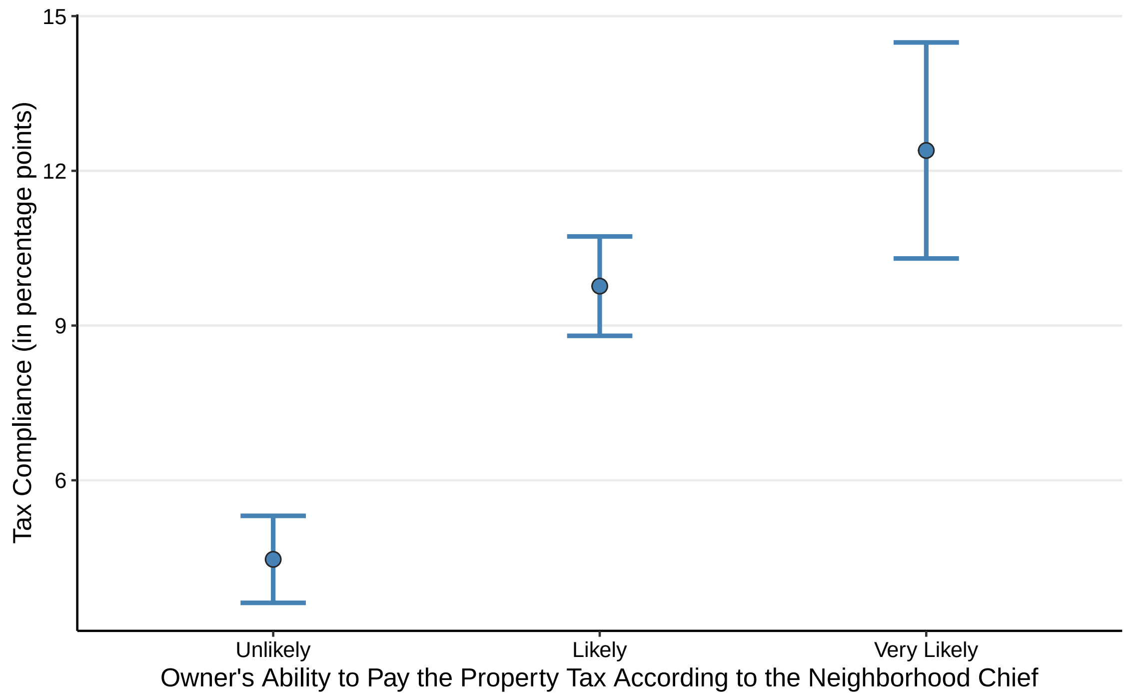
A1 Additional Tables and Figures

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.

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 A2: 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

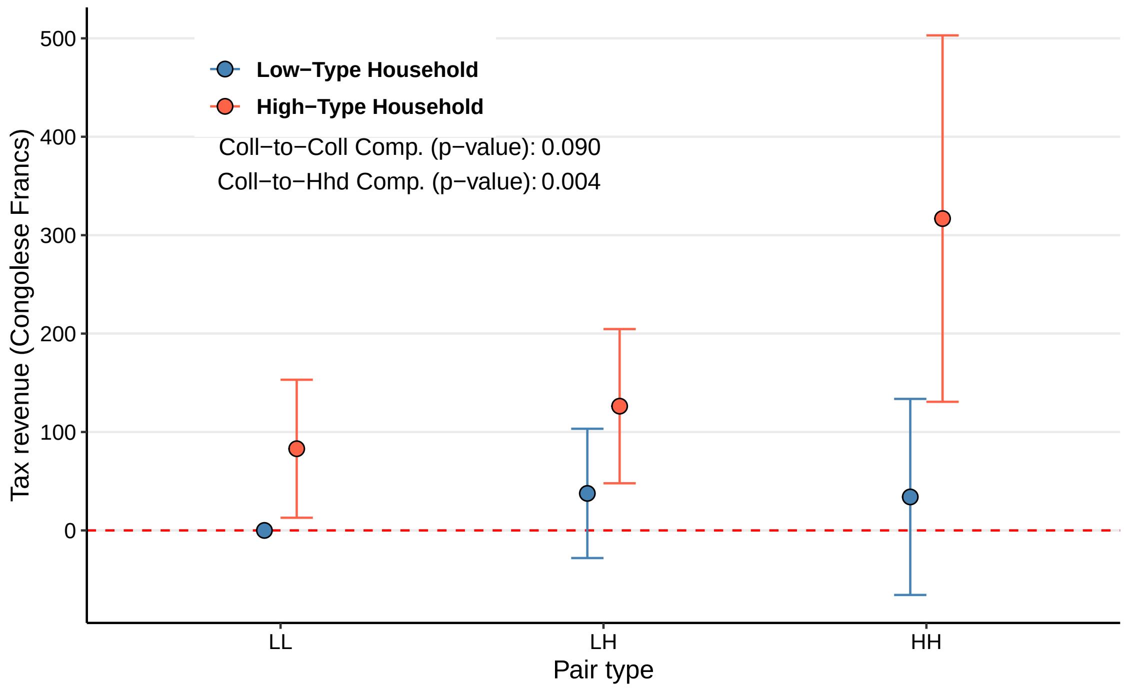
Notes: This table reports the relationship between household type (low or high) and property or property owner's characteristics. More specifically, we regress each property or property owner's characteristic on an indicator for the household being high type. Columns 1–6 report the correlation coefficient, standard errors (robust to heteroskedasticity), p-value, 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 A9. We discuss these results in Section 6.1.

TABLE A3: 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

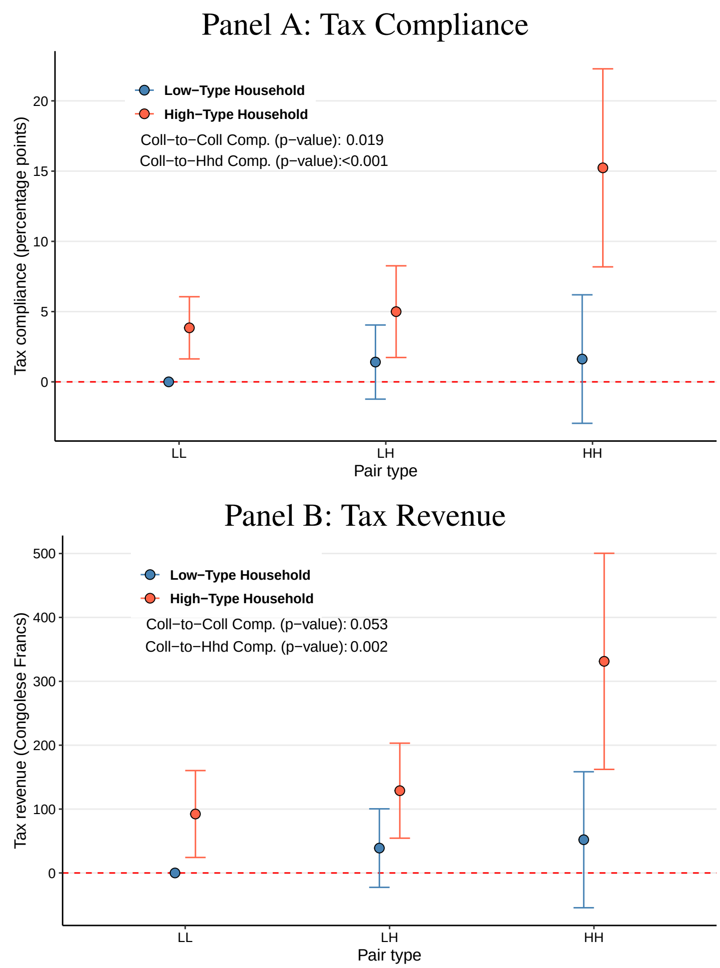
Notes: This table reports the relationship between characteristics and the type (low or high) of the tax collector. More specifically, we regress each collector's characteristic on an indicator for the collector being high type. Columns 1–6 report the correlation coefficient, standard error (robust to heteroskedasticity), p-value, 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 A2: 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. We report the p-value associated with a test for tax revenue exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.1.1 and 7.1.2.

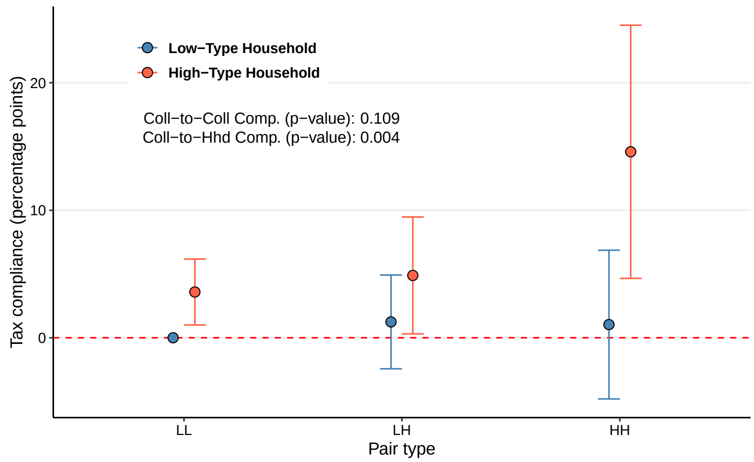
FIGURE A3: TAX COMPLIANCE AND TAX REVENUE BY COLLECTOR AND HOUSEHOLD TYPES — HOUSEHOLD CHARACTERISTICS CONTROLS



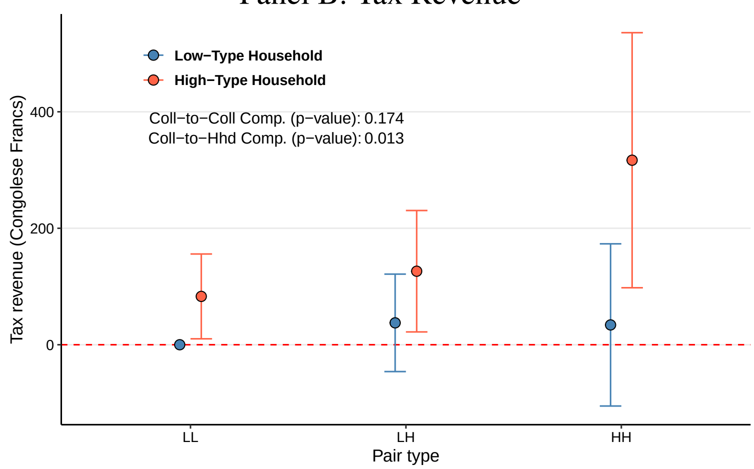
Notes: This figure shows the estimates of the average tax compliance (Panel A) and tax revenue per owner (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures tax compliance (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 clustered at the neighborhood level. All estimates include household characteristics controls from Panel A of Table 1 (distance to state buildings, distance to health institutions, distance to education institutions, distance to roads, distance to eroded areas and property value) when estimating tax collector type and tax compliance or revenue by collector and household type. We report the p-value associated with a test for the outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.1.1 and 7.1.2.

FIGURE A4: TAX COMPLIANCE AND TAX REVENUE BY COLLECTOR AND HOUSEHOLD TYPES — BOOTSTRAPPED STANDARD ERRORS

Panel A: Tax Compliance



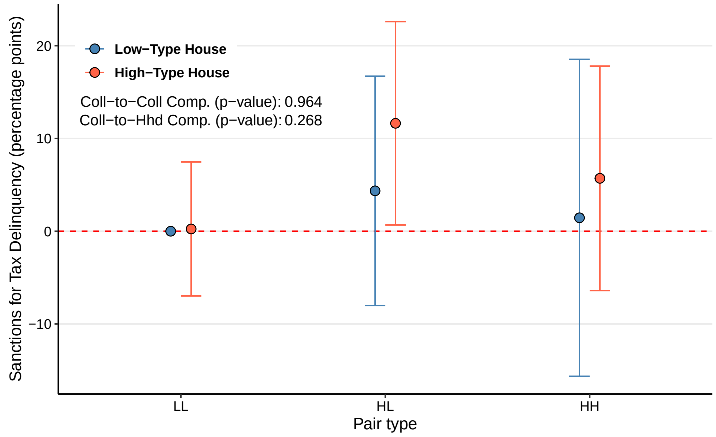
Panel B: Tax Revenue



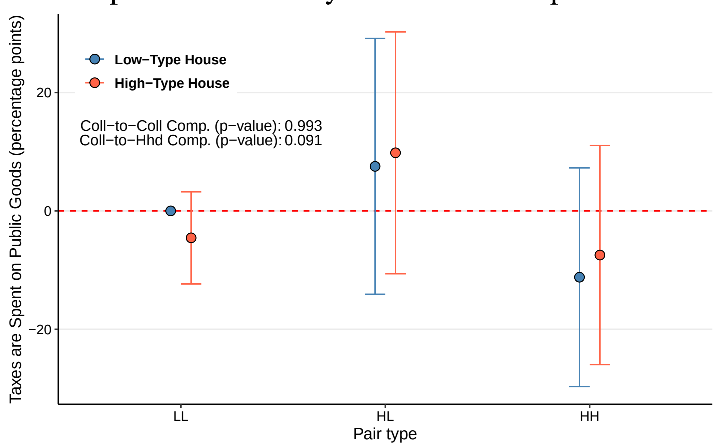
Notes: This figure shows the estimates of the average tax compliance (Panel A) and tax revenue per owner (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different types of collector 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 for the outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.1.1 and 7.1.2.

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

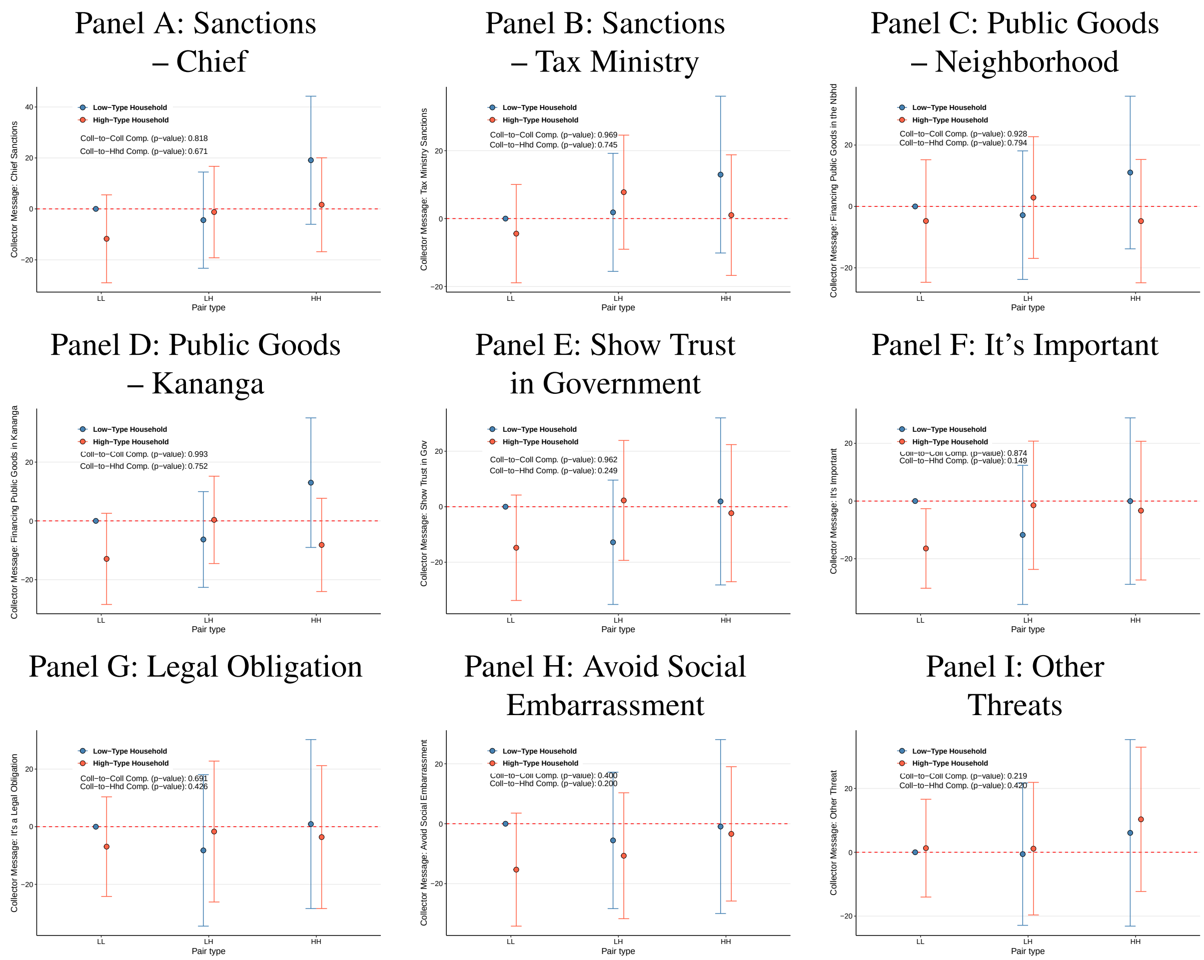


Panel B: Self-Reported Probability that Taxes are Spent on Public Goods



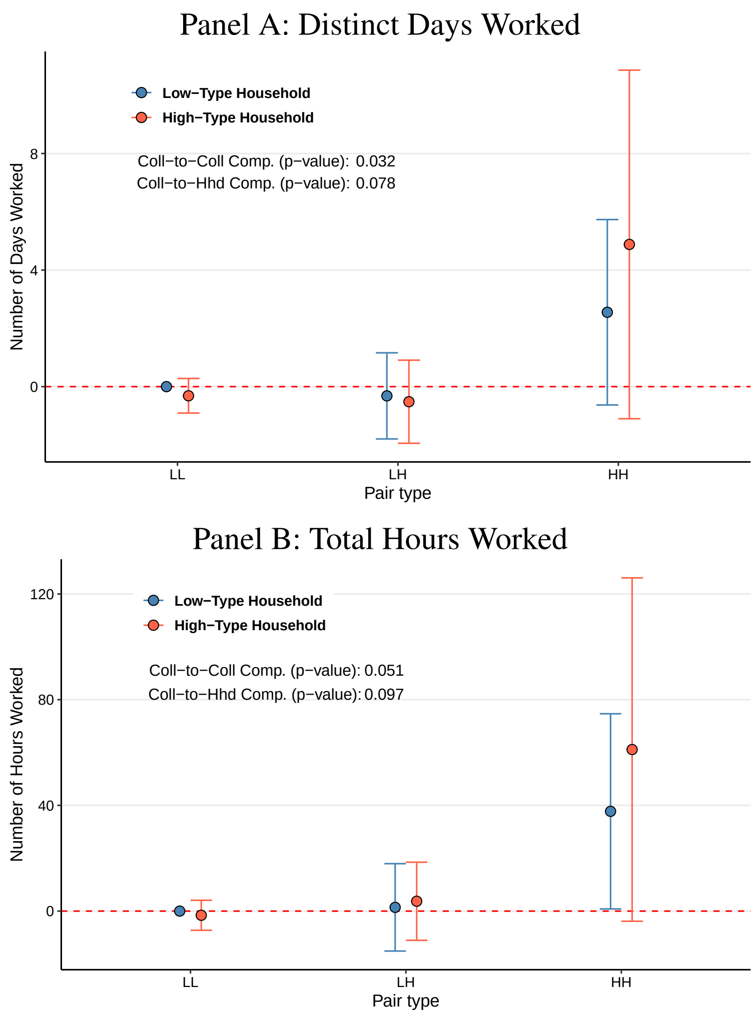
Notes: This figure shows the estimates of the average perception of enforcement and spending of tax revenues on public goods measured when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows the three different 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 for each outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.2.1.

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 for each outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.2.1.

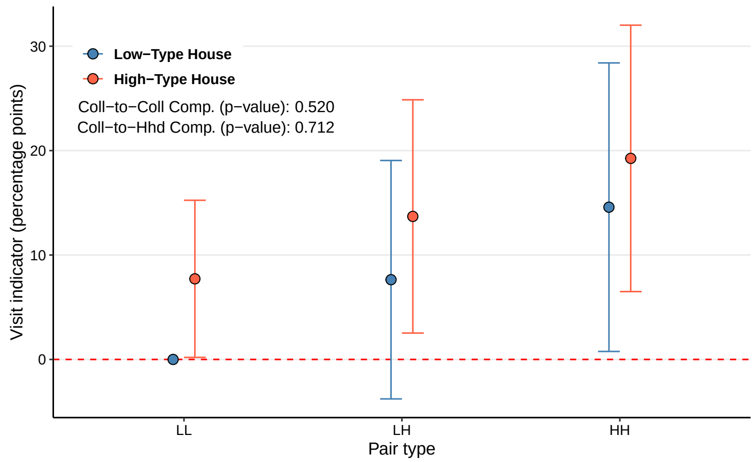
FIGURE A7: DAYS AND HOURS COLLECTORS WORKED BY COLLECTOR AND HOUSEHOLD TYPES



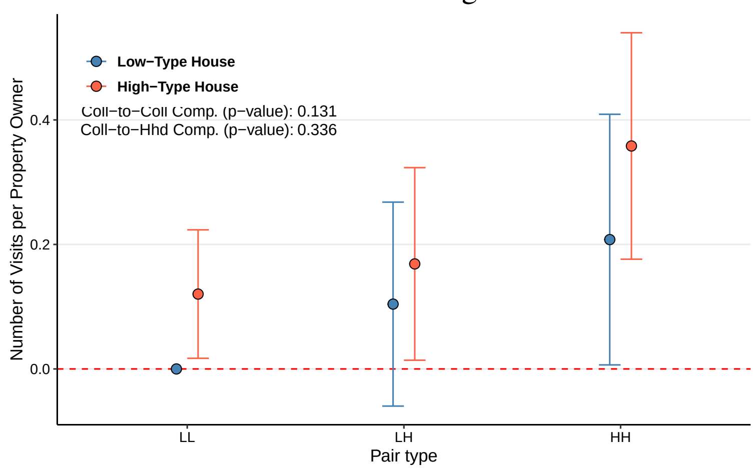
Notes: This figure shows the estimates of distinct days worked by the tax collectors (Panel A) and the total number of hours worked by the tax collectors (Panel B) for different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The x-axis shows 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 for each outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.2.2.

FIGURE A8: VISITS BY COLLECTOR AND HOUSEHOLD TYPES

Panel A: Post-Registration Visit Indicator

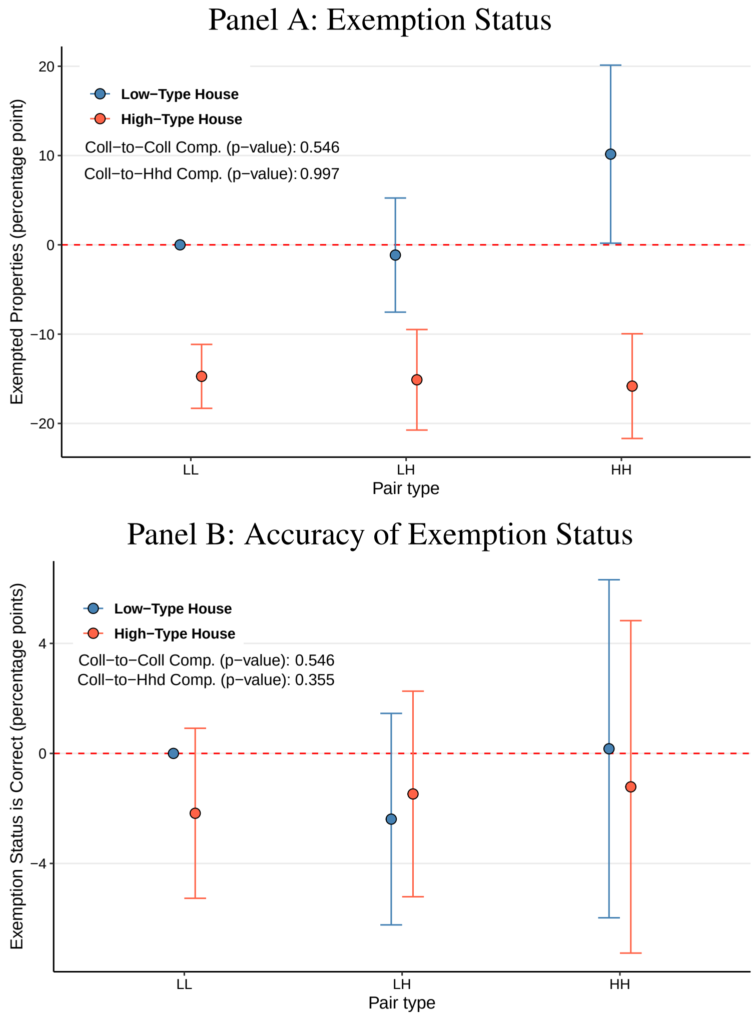


Panel B: Number of Post-Registration Visits



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 for each outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.2.2.

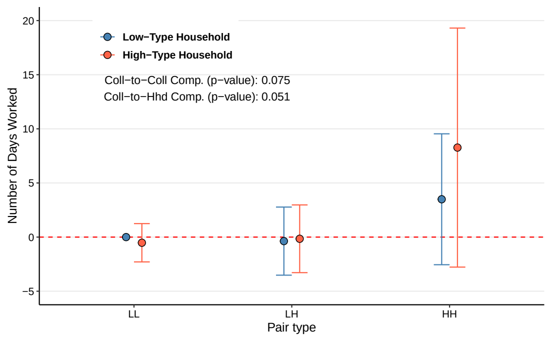
FIGURE A9: EXEMPTION BY COLLECTOR AND HOUSEHOLD TYPES



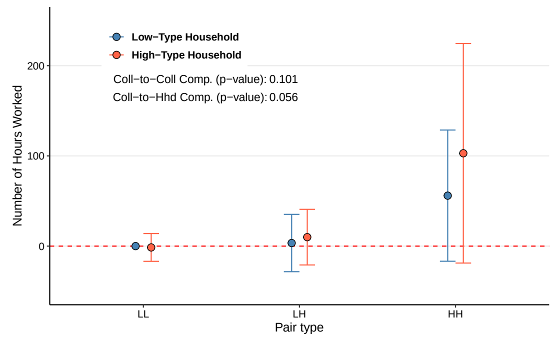
Notes: This figure shows the estimates of the property's tax exemption status at registration (Panel A) and whether this exemption status was deemed accurate by the enumerator during the registration survey (Panel B) when assigned to different types of collector pairs (low-low or LL, low-high or LH, 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 for each outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.2.2.

FIGURE A10: DAYS AND HOURS COLLECTORS WORKED BY COLLECTOR TYPES, HOUSEHOLD TYPES, AND EMPLOYMENT RATES

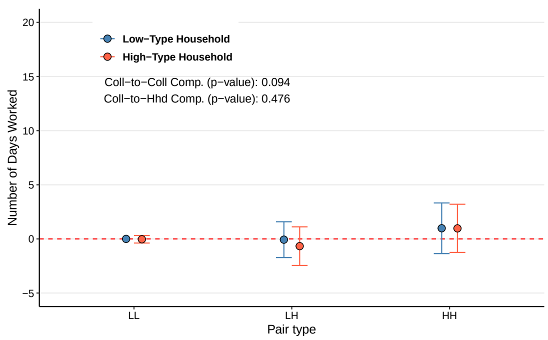
Panel A: Distinct Days Collectors Worked Above Median Employment Rate Nbhd



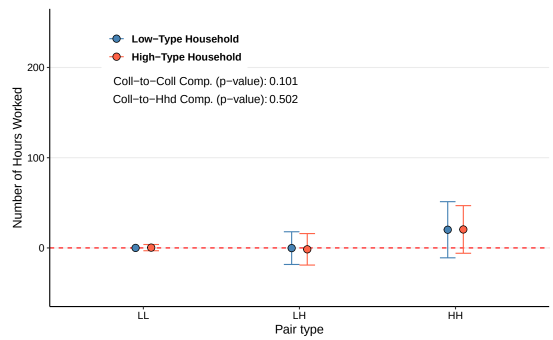
Panel B: Total Hours Collectors Worked Above Median Employment Rate Nbhd



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

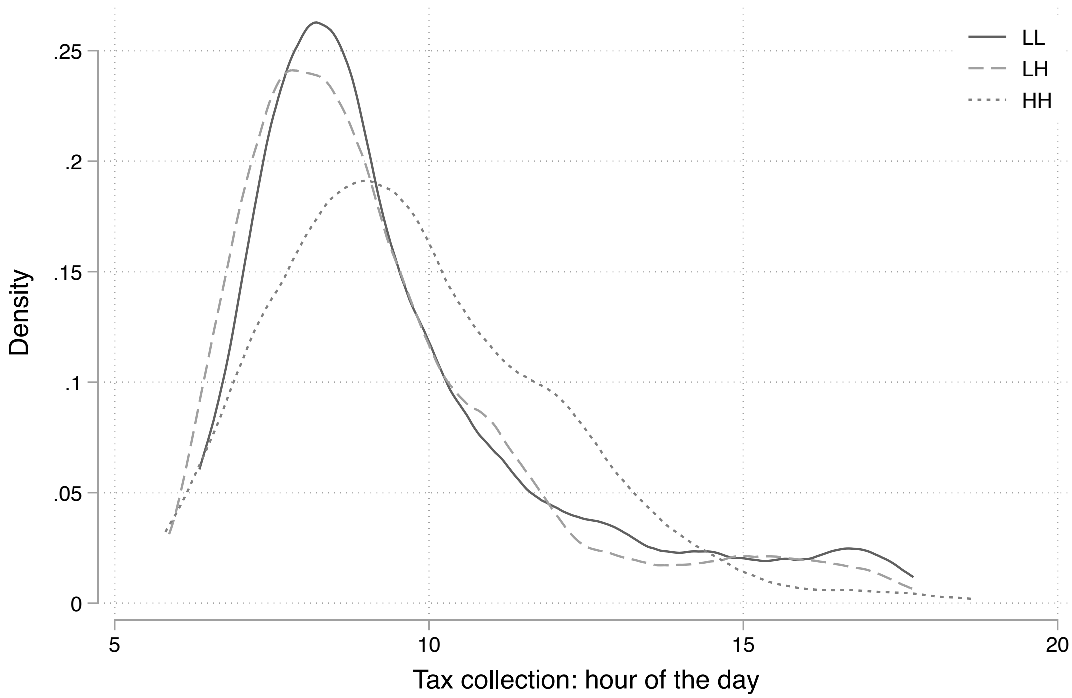


Panel D: Total Hours Collectors Worked Below Median Employment Rate Nbhd



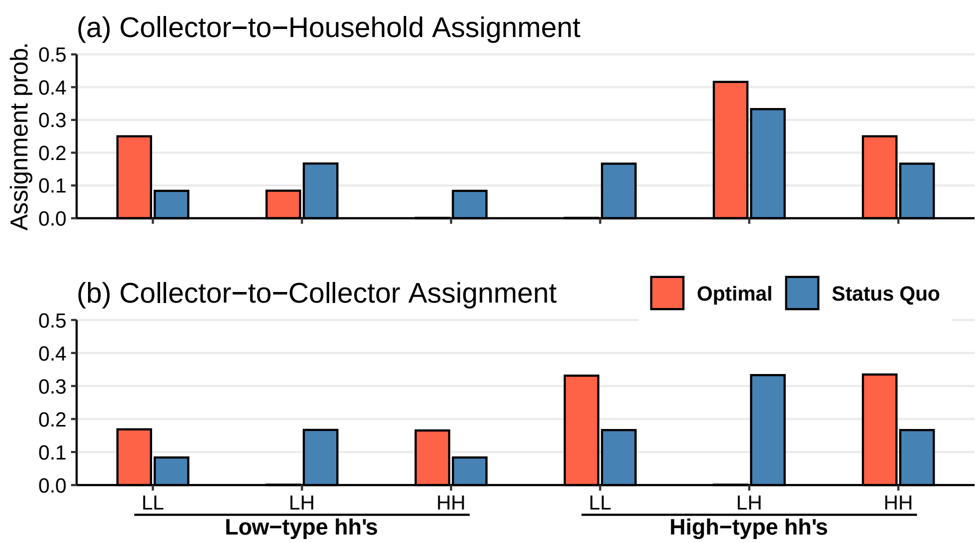
Notes: This figure shows the estimates of distinct days worked by the tax collectors (Panel A and C) and the total number of hours worked by the tax collectors (Panel B and D) for different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). The 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 for each outcome exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 7.2.2.

FIGURE A11: 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 7.2.2.

FIGURE A12: COLLECTOR-TO-HOUSEHOLD AND COLLECTOR-TO-COLLECTOR OPTIMAL ASSIGNMENTS



Notes: This figure shows the assignment function from two alternative optimal assignment mechanisms in comparison to the status quo assignment. Panel A shows the collector-to-household-only optimal assignment. Panel B shows the collector-to-collector-only 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 A4: EFFECTS OF THE OPTIMAL ASSIGNMENT: COMPLIANCE AND REVENUES – PROPERTY CHARACTERISTICS CONTROLS

	Household Types: Household Propensity to Pay			
	Collector Types: Fixed Effects Model		Collector Types: Fixed Effects 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	3.043	56.700
	(1.239)	(30.52)	(1.150)	(27.058)
	[0.024]	[0.074]	[0.008]	[0.036]
Collector-to-Collector Only	1.294	21.444	1.430	25.794
	(0.947)	(21.675)	(0.891)	(20.449)
	[0.172]	[0.322]	[0.109]	[0.207]
Collector-to-Household Only	0.837	17.156	0.875	16.017
	(0.312)	(8.520)	(0.369)	(7.112)
	[0.007]	[0.044]	[0.018]	[0.024]
Mean	8.000	206.213	8.000	206.213
Observations (Holdout Sample)	11,732	11,732	11,486	11,486
Observations (Analysis Sample)	6,904	6,904	6,903	6,903
Property Characteristic Controls	No	No	Yes	Yes

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. All columns present results when collector types are estimated using a fixed effects model as described in Section 6.2. Column 3 and 4 include household characteristics controls from Panel A of Table 1 (distance to state buildings, distance to health institutions, distance to education institutions, distance to roads, distance to eroded areas and property value) when estimating tax collector type and the effect of the optimal assignment on tax compliance and revenue. 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-household and the collector-to-collector dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis; p-values are presented in brackets. 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 Sections 8.1 and 8.2.

TABLE A5: EFFECTS OF THE OPTIMAL ASSIGNMENT: COMPLIANCE AND REVENUES – INCLUDING EXEMPTED PROPERTIES

	Household Types: Household Propensity to Pay			
	Collector Types: Fixed Effects Model		Collector Types: Fixed Effects 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	2.811	52.382
	(1.239)	(30.52)	(1.344)	(31.693)
	[0.024]	[0.074]	[0.036]	[0.098]
Collector-to-Collector Only	1.294	21.444	0.732	12.589
	(0.947)	(21.675)	(0.884)	(20.423)
	[0.172]	[0.322]	[0.407	[0.538]
Collector-to-Household Only	0.837	17.156	1.123	20.936
	(0.312)	(8.520)	(0.316)	(8.284)
	[0.007]	[0.044]	[0.000]	[0.011]
Mean	8.000	206.213	7.025	181.064
Observations (Holdout Sample)	11,732	11,732	13,535	13,535
Observations (Analysis Sample)	6,904	6,904	7,868	7,868

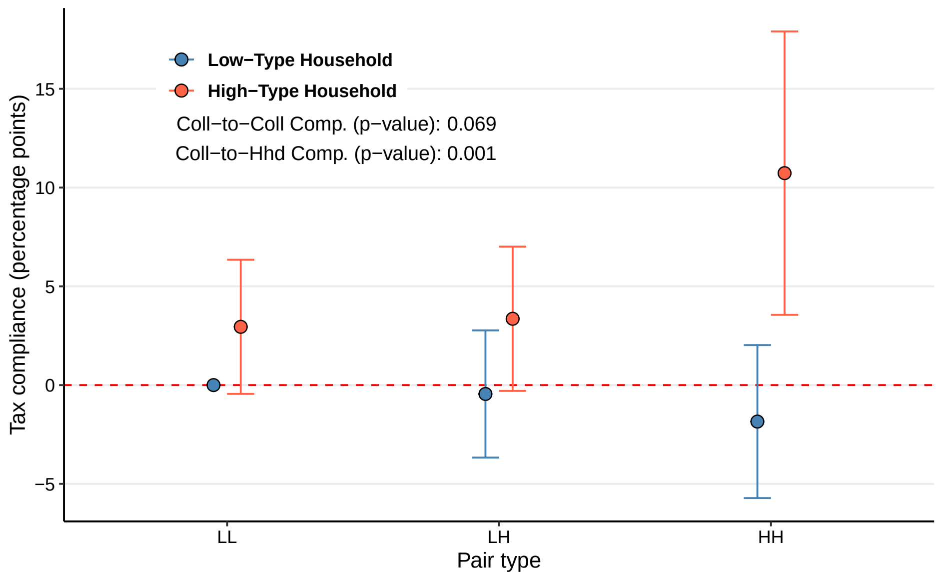
Notes: This table reports estimates of the impact of the counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1 and 3 show estimates for the probability of compliance, i.e., a dummy indicating whether households paid the property taxes (multiplied by 100). 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. All columns present results when collector types are estimated using a fixed effects model as described in Section 6.2. Column 3 and 4 include exempted properties to the holdout and analysis sample. 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-household and the collector-to-collector dimension of the assignment, respectively. Standard errors are clustered at the neighborhood level and presented in parenthesis; p-values are presented in brackets. 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 Sections 8.1 and 8.2.

TABLE A6: 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

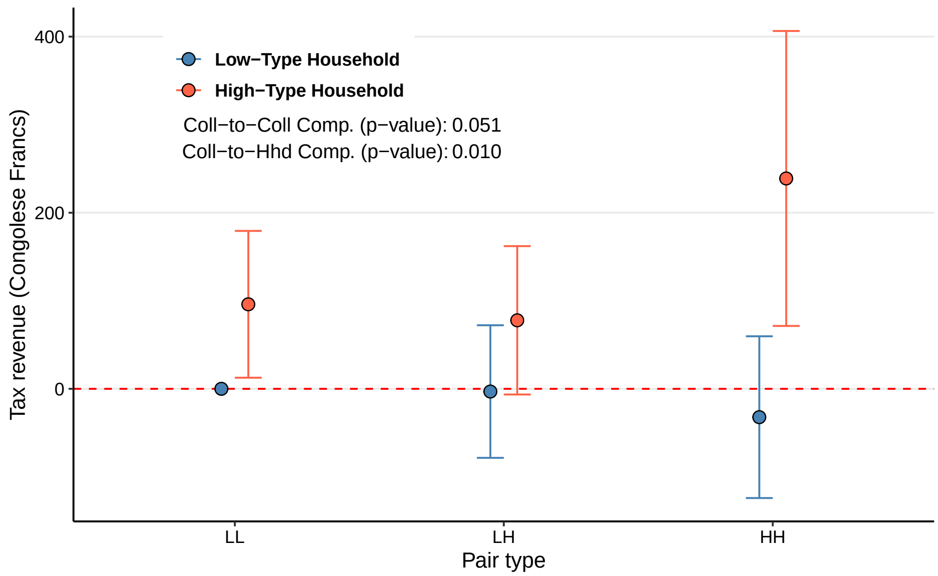
Notes: This table shows the impact of the counterfactual optimal assignment policy in comparison to the status quo (random) assignment. Columns 1 and 3 show results for probability of compliance, i.e., a dummy indicating whether households paid the property tax (multiplied by 100). The point estimates 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 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-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). p-values are presented in brackets. 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 A13: TAX COMPLIANCE BY COLLECTOR AND HOUSEHOLD TYPES – COLLECTORS' TYPE: COLLECTOR CHARACTERISTICS MODEL



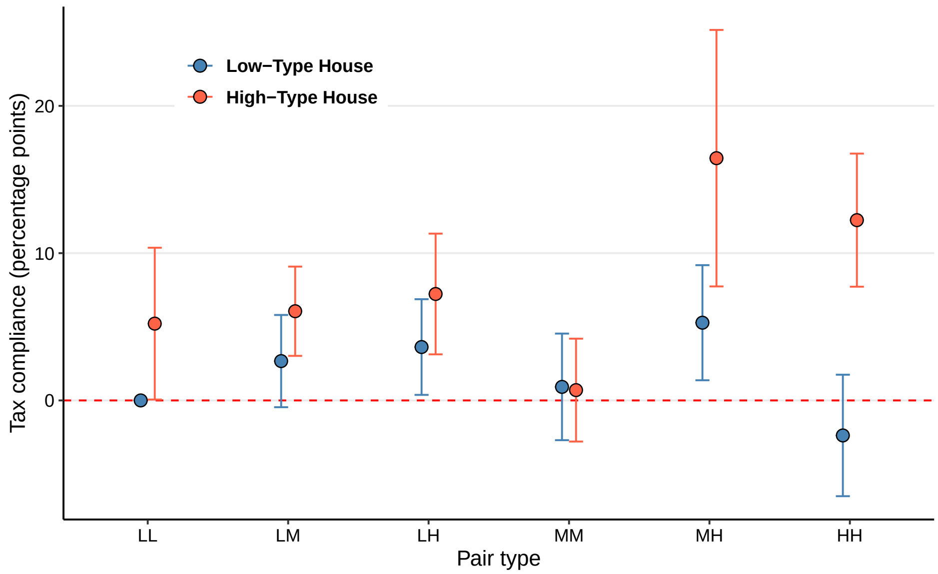
Notes: This figure shows the estimates of average tax compliance when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). Collectors' types are estimated from tax collectors' characteristics as described in Section 8.2. The x-axis shows the three different types of collector pairs: LL, LH, HH. The y-axis captures the tax compliance probability 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 for compliance exhibiting increasing differences in collector type and in collector and household type. We discuss these results in Section 8.2.

FIGURE A14: TAX REVENUE FUNCTION – COLLECTORS' TYPE: COLLECTOR CHARACTERISTICS MODEL



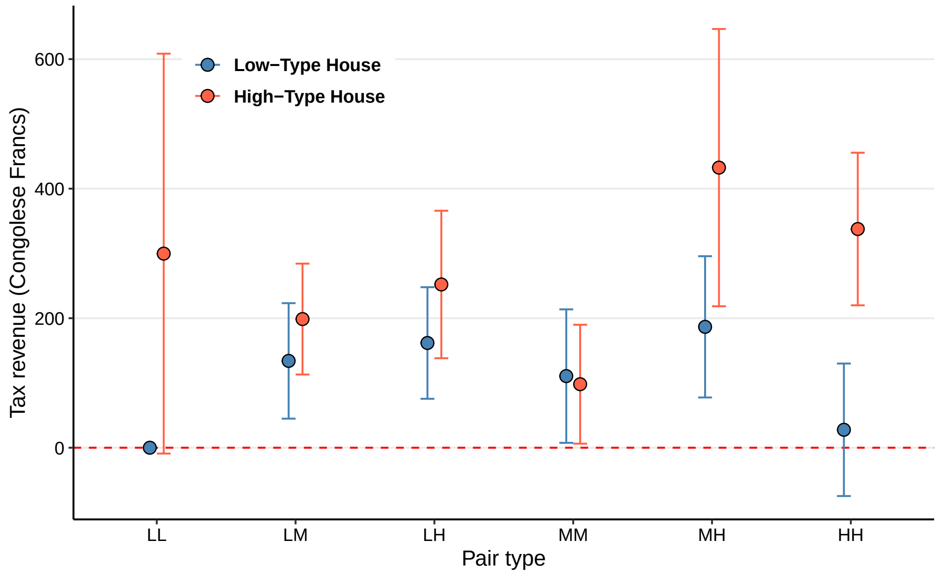
Notes: This figure shows the estimates of average tax revenue per property owner (in Congolese Francs) when assigned to different types of collector pairs (low-low or LL, low-high or LH, high-high or HH) by households' type (low or high). Collectors' types are estimated from tax collectors' characteristics 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 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. We report the p-value associated with a test for tax revenue exhibiting increasing differences in collectors' type and in collector and household type. We discuss these results in Section 8.2.

FIGURE A15: 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 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. We discuss these results in Section 8.2.

FIGURE A16: TAX REVENUE BY COLLECTOR AND HOUSEHOLD TYPES – THREE TYPES OF COLLECTORS



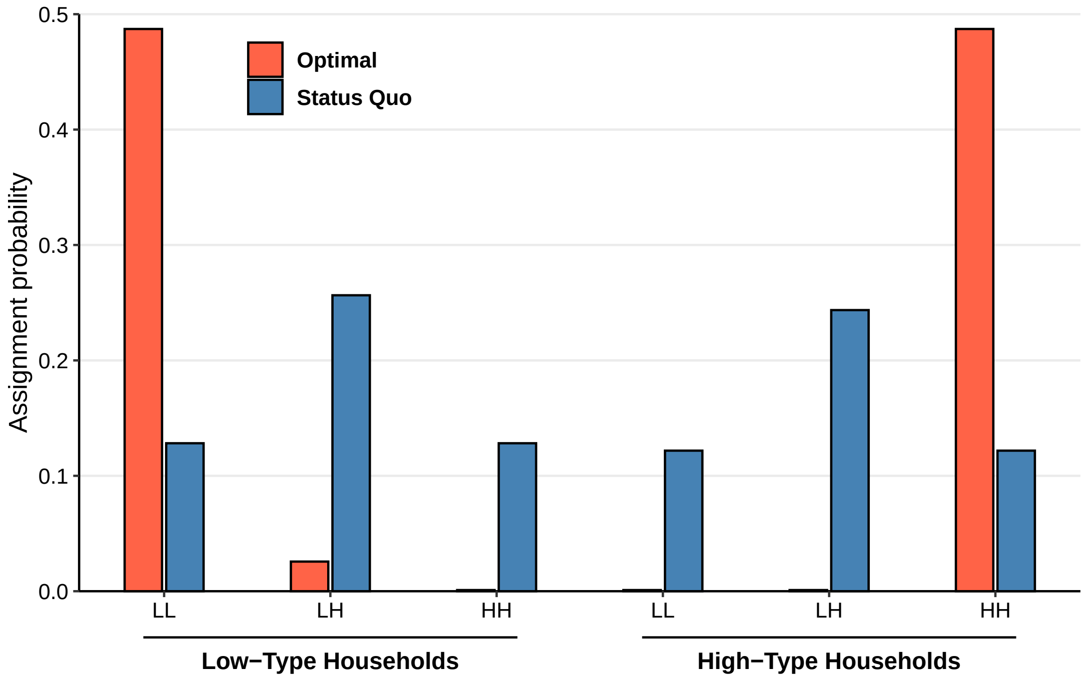
Notes: This figure shows the estimates of average tax revenue (in Congolese Francs) when assigned to different types of collector pairs (low-low or LL, low-medium or LM, low-high or LH, medium-medium or MM, medium-high or MH, high-high or HH) by households' type (low or high). The x-axis shows the six different types of collector pairs: LL, LM, LH, MM, MH, HH. The y-axis captures 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. We discuss these results in Section 8.2.

TABLE A7: 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

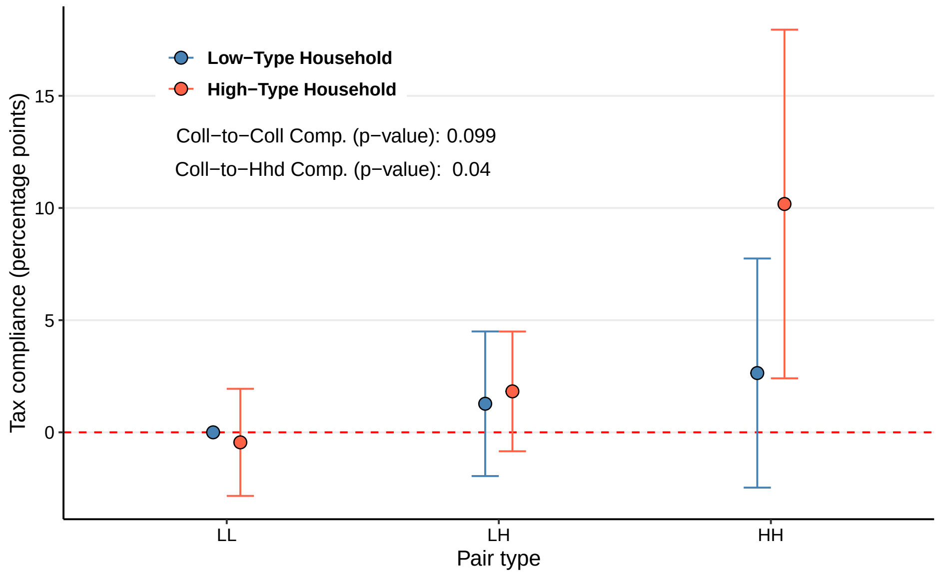
Notes: This table shows the impact of the counterfactual optimal assignment policy with three types of tax collectors (low or L, medium or M, high or H), relative to the status quo (random) assignment. Columns 1 and 3 show results for probability of compliance, i.e., a dummy indicating whether households 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. p-values are presented in brackets. 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 A17: OPTIMAL VS. STATUS QUO ASSIGNMENTS – HOUSEHOLDS' TYPE: HOUSEHOLDS CHARACTERISTICS MODEL



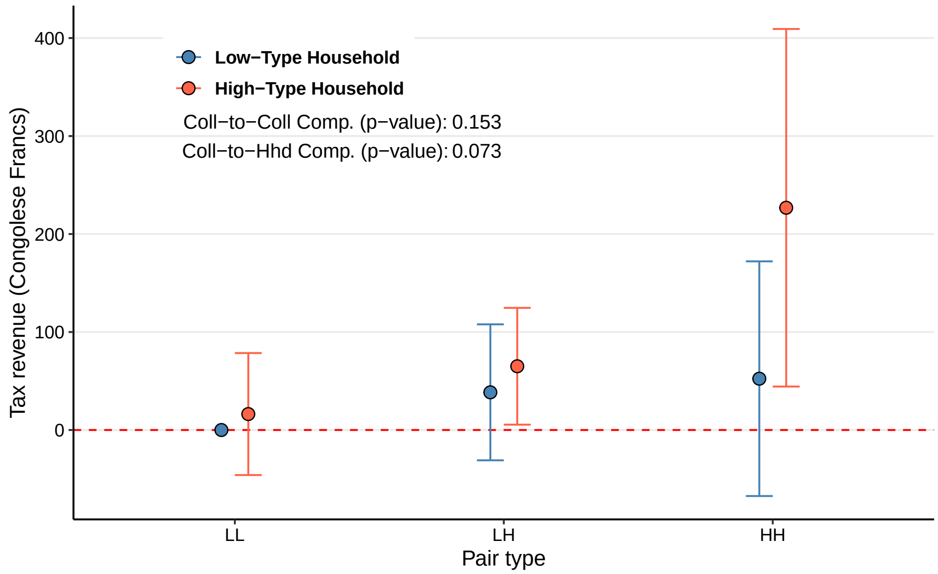
Notes: This figure shows the optimal and the status quo assignment functions. Each bar represents the probability of each match type under the optimal (red) and status quo (blue) assignment functions. The first 6 bars show the assignment functions with matches involving low-type households. The 6 subsequent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

FIGURE A18: TAX COMPLIANCE BY COLLECTOR AND HOUSEHOLD TYPES – HOUSEHOLDS' TYPE: HOUSEHOLDS CHARACTERISTICS MODEL



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 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 for compliance exhibiting increasing differences in collectors' type and in collector and household type. We discuss these results in Section 8.2.

FIGURE A19: TAX REVENUE FUNCTION – HOUSEHOLDS' TYPE: HOUSEHOLDS CHARACTERISTICS MODEL



Notes: This figure shows the estimates of 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). 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 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. We report the p-value associated with a test for tax revenue exhibiting increasing differences in collectors' type and in collector and household type. We discuss these results in Section 8.2.

TABLE A8: 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 2. 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. p-values are presented in brackets. 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 A9: 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. p-values are presented in brackets. 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 A10: 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]	[0.101]
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. p-values are presented in brackets. 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.

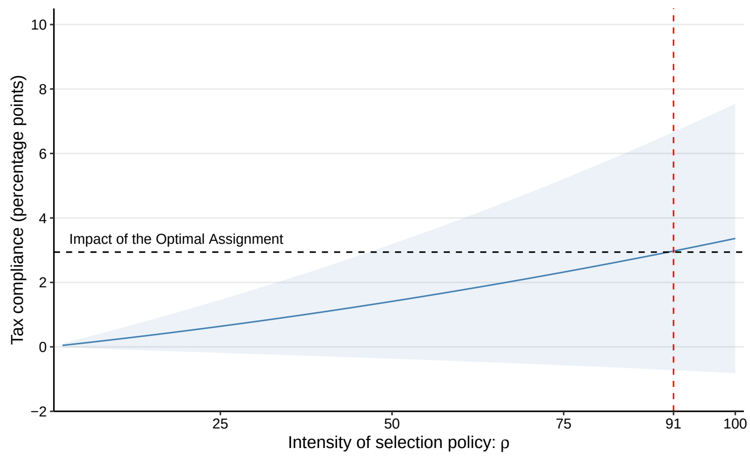
TABLE A11: EFFECTS OF THE OPTIMAL ASSIGNMENT ON TAX COMPLIANCE AND REVENUE – ROBUSTNESS: INFERENCE ON WINNERS

	Objective: Compliance Maximization		Objective: Revenue Maximization	
	Tax Compliance (in percentage points) (1)	Tax Revenue (in Congolese Francs) (2)	Tax Compliance (in percentage points) (3)	Tax Revenue (in Congolese Francs) (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

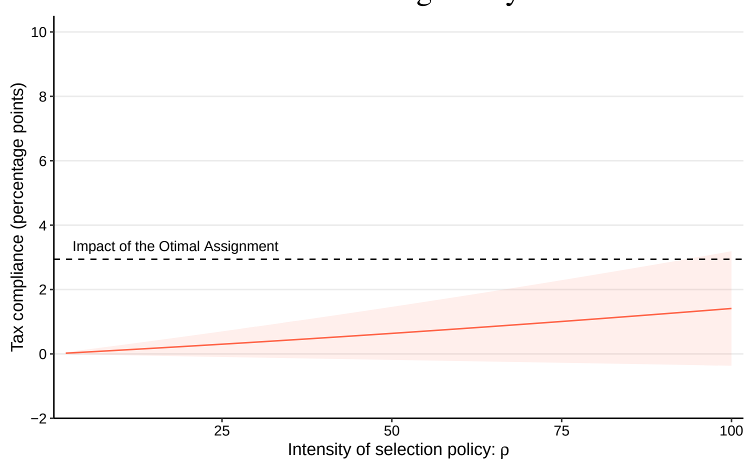
Notes: This table provides estimates and 90% confidence intervals for the impact of the optimal policy after accounting for possible over-fitting concerns associated with the "winner's curse" problem (Andrews et al., 2019). Section 8.2 describes how we adapt the methodology proposed by Andrews et al. (2019) to our context. Row 1 provides our baseline estimates from Table 2 and Table A9. Rows 2 and 3 provide the conditional and hybrid estimators suggested by Andrews et al. (2019). 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 A20: EFFECTS OF SELECTION POLICIES WHEN COLLECTOR TYPES ARE ESTIMATED USING COLLECTORS' CHARACTERISTICS

Panel A: Reallocation Policy



Panel B: Hiring Policy



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 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 2. We discuss these results in Section 8.4.1.

TABLE A12: 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

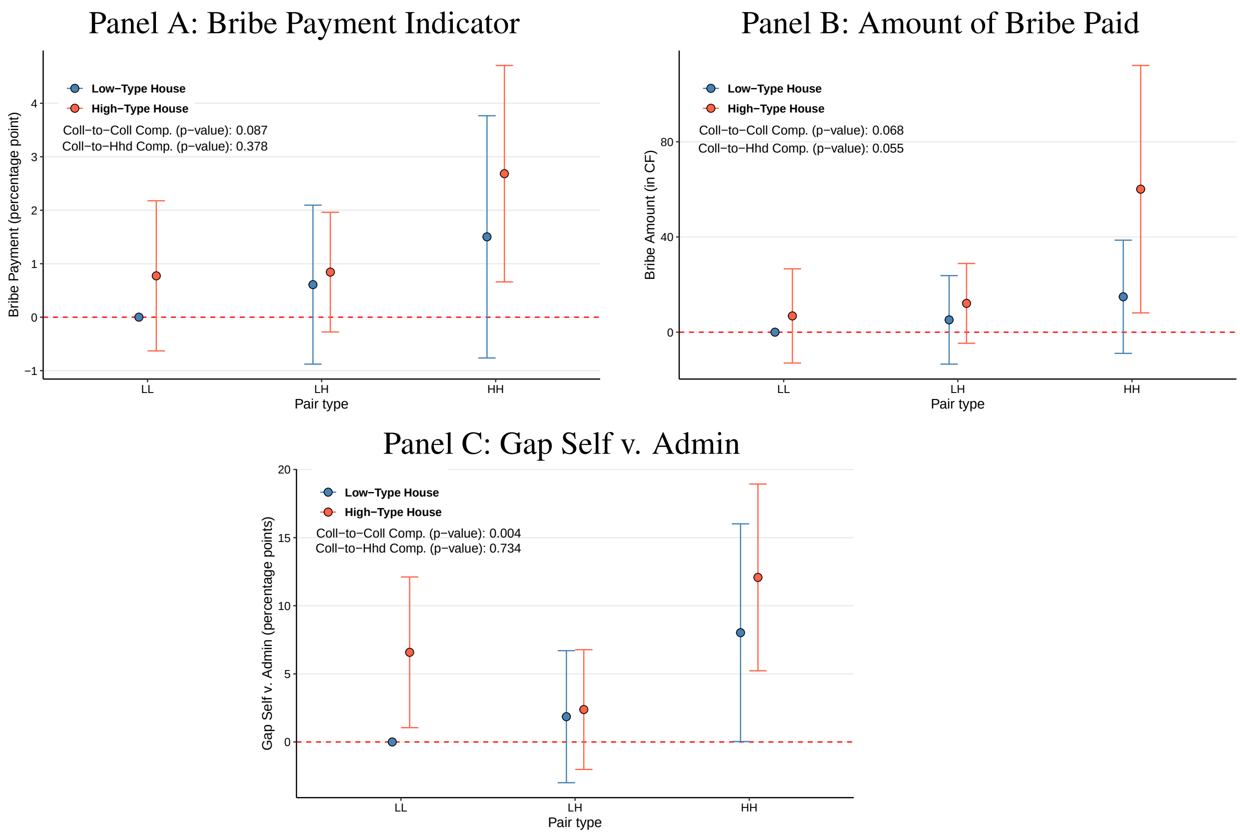
Notes: This table examines treatment effects of the collectors' piece-rate wage on tax compliance, tax revenues, tax visits, and bribe payments. It reports the results of regressions of the log of the piece-rate wage on tax compliance (Columns 1), tax revenue (Columns 2), a post-registration visit indicator (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 8.4.2.

TABLE A13: 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. (2020b) 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 8.4.2.

FIGURE A21: 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 from 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 discuss these results in Section 9.1.

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 et al., 1984). However, there might be more than one solution to the optimal assignment problem.[1] For simplicity we follow Bhattacharya (2009) and assume uniqueness of the optimal assignment (Assumption 1).

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

A2.2 Asymptotic Distribution

The importance of the uniqueness assumption lies in the asymptotic properties of the optimal assignment and ARE estimators (Bhattacharya, 2009).[2] We show that 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.

To prove these results, we need to show that β identifies the average compliance function up to a constant. This can be obtained by further assuming that the assignment is conditionally exogenous:

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

Where $D_h(c_1, c_2)$ is an indicator for observing the 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 $Y_h(c_1, c_2)$.[3] 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 1 and described in Section 3.

[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]The asymptotic distribution of $\widehat{\tau}(\rho, \lambda)$ is standard, being a weighted average of a (asymptotically) normally-distributed vector.

[3]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.

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 that Assumptions 1–2 hold. Then:*

1. $\widehat{f^*}$ *is consistent to* f^*.

2. $\widehat{ARE}$ *is consistent to* ARE.

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

The third result of Proposition 1 states that the sampling error of $\widehat{f^*}$ is asymptotically irrelevant for the estimation of ARE. This result relies on $\widehat{f^*} \xrightarrow{p} f^*$ at a rate faster than $\sqrt{n}$ (Bhattacharya, 2009).

A2.3 Proof of Asymptotic Distribution Properties

Item 1 of Proposition 1. It is exactly the same as Proposition 1 in Bhattacharya (2009), so we refer the reader to the details there.

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

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

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

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

Item 1 of Proposition 1 guarantees that $\widehat{\boldsymbol{f}}^* \xrightarrow{p} \boldsymbol{f}^*$ and $\widehat{\boldsymbol{\beta}}$ converges in probability to $\boldsymbol{\beta}$ by assumption. The limit of the multiplication of two objects is the multiplication of the limit (in probability) of these two objects, which gives us the desired result.

Item 3 of Proposition 1. 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

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

Where the first term can be written as

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

The second term in the last line, $\sqrt{n}\left(\widehat{\beta}-\beta\right)\widehat{f}^{*}1_{\left[\widehat{f}^{*}\neq f^{*}\right]}$, is $o_p(1)$ (i.e., converges in probability to zero) since $\widehat{f}^{*}$ is bounded (because it's 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)). The third term in the last line, $\sqrt{n}\beta\left(\widehat{f}^{*}-f^{*}\right)1_{\left[\widehat{f}^{*}\neq f^{*}\right]}$ 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 Estimation of the Average Compliance Function

The coefficients of interest when estimating the average compliance function in Equation 7:

$$y_{hnt} = \sum_{a_1 \in A} \sum_{a_2 \geq a_1} \sum_{v=0,1} \beta(a_1, a_2, v) \cdot 1_{[c(n)=(a_1,a_2)]} \cdot 1[v_h = v] + \lambda_t + \varepsilon_{hnt}$$

are the $\beta(a_1, a_2, v)$ coefficients. Absent the campaign month dummies, these coefficients are the average tax compliance function $Y(a_1, a_2, v)$. Because we include campaign month dummies, $\beta(a_1, a_2, v)$ should be interpreted as a convex combination of $Y(a_1, a_2, v, t) - Y(L, L, l, t)$ (Abadie and Cattaneo, 2018), where $Y(.)$ is a function of the campaign month t.[4] To avoid this complication in the notation, we make the additional assumption that the average compliance function is separable in campaign month.[5]

Assumption 3 (Time Period Separability)**.** *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.*

A4 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 (3)).

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*, $\lambda = \frac{1}{2}$.[6]

Under a selection policy characterized by ρ and λ, the number of assignments to high-

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

[5] The estimand could be interpreted as a convex combination of $Y(a_1, a_2, v, t) - Y(L, L, l, t)$ if this assumption was invalid.

[6] 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).

type collectors is given by:

$$N^{asgmt}(H;\rho,\lambda) = N^{asgmt}_{f^{SQ}}(H) + N^{asgmt}_{f^{SQ}}(L)\rho\lambda \tag{13}$$

$N^{asgmt}_{f^{SQ}}(H)$ 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:

$$f^S(a_1, a_2, v; \rho, \lambda) = f^S(a_1; \rho, \lambda) f^S(a_2; \rho, \lambda) f^{SQ}(v) \tag{14}$$

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:

$$\tau(\rho,\lambda) \equiv \sum_{v\in V} \sum_{a_1,a_2\in A^2} \left[f^S(a_1,a_2,v;\rho,\lambda) - f^{SQ}(a_1,a_2,v)\right] Y(a_1,a_2,v) \tag{15}$$

To estimate the impact of selection policies, $\tau(\rho,\lambda)$, we substitute the estimated average tax compliance function $\widehat{\beta}(a_1,a_2,v)$ in Equation (15), which gives:

$$\widehat{\tau}(\rho,\lambda) \equiv \sum_{v\in V} \sum_{a_1,a_2\in A^2} \left[f^S(a_1,a_2,v;\rho,\lambda) - f^{SQ}(a_1,a_2,v)\right] \widehat{\beta}(a_1,a_2,v) \tag{16}$$

where the distributions $f^S(\rho,\lambda)$ and f^{SQ} in $\widehat{\tau}(\rho,\lambda)$ are the theoretical distributions.[7]

A5 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 the complementarities in collector-to-collector and collector-to-household match type that we observe in the average

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

compliance function.

A5.1 Homophily

Another explanation for complementarity in collector types includes performance gains or losses arising from homophily that are more pronounced among high-type collectors. Tax collection could for example be enhanced if collectors from similar backgrounds can communicate more easily.[8] Generally, horizontal differentiation among collectors that impacts the tax collection production function will be orthogonal to the collector types we estimate. But it remains possible that certain collector traits associated with type could differentially boost compliance among high type collectors.

To be precise, for homophily to explain the complementarities we observe in the tax compliance function, we would need to observe that (*i*) similarity between collectors in certain traits is associated with higher tax compliance, and (*ii*) the benefits from homophily should be more pronounced among *H-H* teams and for high-type households. Regarding (*i*), regressing similarity in collector traits within pairs on tax compliance, we find relatively few traits for which similarity between tax collectors is associated with higher tax compliance. The only traits where homophily is associated with higher compliance among high-type households include collectors' wealth (number of possessions) and their redistributive preferences (Table A14).[9]

Turning to (*ii*), we find little evidence that these relationships between collector similarity and productivity are more pronounced for *H-H* teams (relative to *L-H* and *L-L* teams) among high-type households, as would be necessary to explain complementarity.[10] Similarity in wealth, redistributive preferences, or other traits do not appear to differentially boost compliance for *H-H* pairs when assigned to high-type households (Table A15).[11] Homophily is therefore unlikely to explain complementarity in the tax compliance func-

[8] For instance, Hjort (2014) and Marx et al. (2021) find that ethnically homogeneous teams in Kenya are more productive in flower factories and during voter registration campaigns, respectively.

[9] By contrast, similarity in traits typically associated with homophily — gender, age, and language ability (literacy) (Lang, 1986) — are not associated with higher team performance among high type households (Table A14, Panel A). There is marginally significant evidence that teams of mixed ethnicity collect more tax, which runs counter to evidence on team ethnic composition from Kenya (Hjort, 2014; Marx et al., 2021) (though Marx et al. (2021) do find similar results to ours when examining manager-worker ethnic matches). However, there is too little variation in ethnicity among collectors to put much stock in this result.

[10] We focus on high-type households since complementarities in the collector-to-collector dimension of the assignment are only present among high-type households (Figure 2).

[11] The 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 average complementarities in collector type we observe.

tion.

TABLE A14: TAX COMPLIANCE BY SIMILARITY IN COLLECTOR CHARACTERISTICS

	Col. Similarity				
	Coef.	SE	p-value	Mean Char.	Obs.
Outcome: Tax Compliance	(1)	(2)	(3)	(4)	(5)
Panel A: Demographics					
Female	0.000	0.007	0.949	0.063	19,587
Age	0.007	0.007	0.316	30.043	19,196
Main Tribe	-0.014*	0.008	0.083	0.197 4	19,587
Years of Education	0.003	0.006	0.614	3.588	19,196
Math Score	0.002	0.007	0.738	-0.181	19,196
Literacy (Tshiluba)	-0.008	0.007	0.252	-0.029	19,196
Literacy (French)	0.000	0.007	0.995	-0.004	19,196
Monthly Income	0.010	0.011	0.364	171.938	19,587
Possessions	0.012*	0.006	0.060	1.582	19,196
Born in Kananga	0.005	0.006	0.396	0.539	19,587
Panel B: Trust in the Government					
Trust Nat. Gov.	0.003	0.006	0.648	2.928	19,587
Trust Prov. Gov.	-0.006	0.005	0.172	2.953	19,587
Trust Tax Min.	0.002	0.005	0.687	3.495	19,587
Index	0.003	0.006	0.636	0.094	19,587
Panel C: Perceived Performance of Government					
Prov. Gov. Capacity	-0.006	0.006	0.349	0.386	19,587
Prov. Gov. Responsiveness	0.007	0.007	0.276	1.706	19,587
Prov. Gov. Performance	-0.006	0.005	0.221	4.476	19,587
Prov. Gov. use of Funds	-0.003	0.009	0.738	616.936	19,587
Index	-0.006	0.005	0.209	0.082	19,587
Panel D: Government Connections					
Job through Connections	-0.023***	0.006	0.001	0.275	15,609
Relative work for Prov. Gov.	0.001	0.005	0.794	0.236	19,587
Relative work for Tax Ministry	-0.002	0.008	0.819	0.297	19,587
Index	-0.006	0.007	0.387	0.042	19,196
Panel E: Tax Morale					
Taxes are Important	0.001	0.007	0.873	2.788	19,587
Work of Tax Min. is Important	0.005	0.009	0.579	3.805	19,587
Paid Taxes in the Past	0.010	0.006	0.105	2.090	19,587
Index	0.008	0.007	0.219	0.119	19,587
Panel F: Redistributive Preferences					
Imp. of Progressive Taxes	-0.002	0.006	0.808	1.612	19,587
Imp. of Progressive Prop. Taxes	0.016***	0.004	1.180	0.003	19,587
Imp. to Tax Employed	-0.007	0.007	0.306	3.335	19,587
Imp. to Tax Owners	0.004	0.008	0.642	3.088	19,587
Imp. to Tax Owners w. title	-0.004	0.006	0.496	3.354	19,587
Index	0.010*	0.005	0.060	-0.277	19,587

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 "Local Information" neighborhoods. Columns 1–3 report the correlation coefficient, standard error (clustered at the neighborhood level), and p-value on the similarity measure. 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. The variables come from surveys with tax collectors described in Section 4. We discuss these results in Section A5.1.

TABLE A15: TAX COMPLIANCE BY PAIR TYPE AND PROXIES FOR SOCIAL LINKS BY COLLECTOR TYPES (HIGH TYPE HOUSEHOLDS)

	Measure of Similarity in Collector Characteristics								
Outcome: Tax Compliance	Female (1)	Age (2)	Main Tribe (3)	Born 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. The sample used is only high type households in "Local Information" neighborhoods. 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 Sections A5.1 and A5.2.

A5.2 Social Incentives

A related but distinct explanation for complementarities in 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 productivity differentially among high-type collectors (Granovetter, 1973; Ashraf and Bandiera, 2018). For example, social incentives could generate convexity in collector type if pairing friends together among *H-H* teams triggers "contagious enthusiasm," while pairing friends together among *H-L* or *L-L* teams triggers an averaging of productivity (conformity) or even generates "contagious malaise" (Bandiera et al., 2010).[12] While homophily concerned the technology of collection — e.g., communication between collectors — this mechanism concerns collectors' incentives to exert effort.

Although we do not directly observe social links, we examine several proxies, including whether collectors hail from similar locations in the city, from the same cohort of collectors working on the tax campaign,[13] or from the same religious denomination.[14] There is marginally significant evidence that less distance between collectors' homes differentially leads *H*-*H* teams to exert more effort — measured by the post-registration visits — among high-type households but this does not translate into higher compliance (Table A16, Columns 1–2).[15] 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).[16] Thus, while we find evidence that social incentives matter in this context, they are unlikely to be the mechanism driving complementarities in the average tax compliance function.

[12] Social incentives could also arise in another form as discrimination against out-group teammates. For example, collectors might be willing to reduce their own payoffs to lower those of out-group teammates (Kranton et al., 2013), which would lower performance among mixed-type teams if ability types align with salient social divisions. However, for this to be the case social divisions would need to match with ability types, such that high type collectors would be more likely to punish their teammate by reducing their own performance when paired with a low-type collector (i.e., low-types would be more often members of the out-group). Though we do not directly observe the strength of social divisions among collectors, the most salient identity marker in our context — ethnicity (tribe) — does not differ across types.

[13] Most collectors began at the start of the tax campaign, but others joined in later months. We therefore define cohort as the first month in which a collector began working on the tax campaign.

[14] All collectors were Christian, the dominant religion in Kananga. Churches are a principal nexus of social activity, 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.).

[15] As noted, we study these patterns among high type households, where there are complementarities in the tax compliance function.

[16] As we note in Section A5.1, 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 A16: SOCIAL INCENTIVES: COLLECTOR HOME LOCATION, COHORT, AND CHURCH BY COLLECTOR TYPE (HIGH TYPE HOUSEHOLDS)

	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. The sample used is only high type households in "Local Information" neighborhoods. 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 A5.2.

A6 Neighborhood-Level Optimal Assignment

To obtain the neighborhood-level optimal assignment, we first estimate the average tax compliance in neighborhood n when assigned to collectors of type a_1 and a_2:

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

where $N_n(l)$ and $N_n(h)$ are the number of low-type and high-type households in neighborhood n, respectively.

The neighborhood-level optimal assignment f^* is the probability mass function that maps the probability of assigning a collector of type a_1 and a_2 to a neighborhood n and solves

$$f^* \equiv \arg\max_f \sum_{n\in N} \sum_{a_1,a_2\in\{L,H\}^2} f(a_1, a_2, n)\overline{Y}_n(a_1, a_2) \tag{17}$$

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

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

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 instead of over household types v.[17] The first constraint imposes that the probability that a neighborhood is assigned to one pair of collectors equals one. The second constraint imposes that tax collectors receive the same number of assignments as under the status quo assignment. These constraints are analogous to the constraints in Problem (1), but at the neighborhood instead of the household level.

In Problem 17, the neighborhood-level outcome of interest is the average compliance $\overline{Y}_n(a_1, a_2)$. An alternative outcome of interest would be the expected number of tax payers, $N_n\overline{Y}_n(a_1, a_2)$, where N_n is the number of households in neighborhood n. To obtain the

[17]For this exercise, we exclude neighborhood with less than 10 observations. We thus exclude 6 neighborhoods from this analysis for a total sample size of 74 neighborhoods.

optimal assignment, we then replace the objective function in Problem 17 with

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

While in Problem 17, the neighborhoods only differ by their relative share of high-type households, with this alternative definition of the neighborhood-level outcome of interest, neighborhoods also differ in their number of households. Thus, the government can 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 $a_1, a_2 \in \{L, H\}^2$.

A7 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,

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

where $w_f(z) = \dfrac{f(z)\Pr(Y_h = 1|z)}{\sum_{z'} f(z')\Pr(Y_h = 1|z')}$ 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 $\widehat{w}_{f^{SQ}}(z) = \dfrac{f^{SQ}(z)\widehat{\beta(z)}}{\sum_{z'} f^{SQ}(z')\widehat{\beta(z')}}$.

A8 Spillovers and the SUTVA Assumption

Throughout the analysis, we have assumed that potential outcomes are not affected by the assignment function. This assumption, known as the stable unit treatment value assumption (SUTVA) in the impact evaluation literature, is essential for the identification of average compliance under different assignment functions. To see this, we generalize the average compliance function so that it depends on the assignment function f and denote it $Y(a_1, a_2, v_h, f)$. Using this notation, the average compliance under f is given by

$$\overline{Y}(f) \equiv \sum_{a_1,a_2,v_h} f(a_1, a_2, v_h) Y(a_1, a_2, v_h, f)$$

Because we can only identify the average compliance function under the status quo assignment function f^{SQ}, $Y(a_1, a_2, v_h, f^{SQ})$, we can only identify[18]

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

which might be different from

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

unless $Y(a_1, a_2, v_h, f^{SQ}) = Y(a_1, a_2, v_h, f^*)$, which is implied by SUTVA. In other words, SUTVA assumes that the potential outcomes remain the same when types a_1, a_2, v_h are preserved but the assignment is modified.

In our context, changing the collector assignment function could impact potential outcomes by match type and thus constitute a SUTVA violation for two reasons. First, implementing the optimal assignment could impact collector effort, which is a key input to tax compliance given the door-to-door nature of tax collection. Second, changing the assignment function could impact potential outcomes if collectors learn tax collection skills over time and from one another. We explore both possibilities below.

[18] $\overline{Y}^P(f^*)$ can be interpreted as a partial equilibrium quantity (thus the p superscript).

A8.1 Endogenous Effort Provision

A8.1.1 Endogenous Effort due to Time Constraints

Throughout the analysis, we have implicitly assumed that changing the assignment function would not affect collector effort provision by match type. In practice, however, this assumption might not hold and endogenous effort responses to changes in the assignment function would mean that the impact of the optimal assignment policy would differ from the ones presented in Section 8. Endogenous effort might in particular affect the impact of the optimal assignment policy if (*i*) collectors target visits differently by household type — e.g., they visit high-type households more than low-type households — and (*ii*) collectors are constrained in the time they spend working in each neighborhood during the tax campaign. If both conditions are met, then implementing the optimal assignment could decrease the probability that high-type households are visited and thus impact potential outcomes by match type.

To see this, consider the following simplified example. Assume that there are four households in Kananga, two low-type households (v^L) and two high-type ones (v^H). Additionally assume that there are two collector teams, a low-type team (a^{L-L}) and a high-type one (a^{H-H}), each assigned to two households. 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 $e_{p,h}$ approximates collector effort and is a dummy for whether collector pair p visited household h after registration. Finally, assume that effort is constrained, i.e., after property registration each collector team can only visit one of the two households it is assigned to. This restriction captures potential time constraints tax collectors faced due to the government's need to complete the tax campaign in all neighborhoods of Kananga by the end of the fiscal year.

In this example, when collectors are time-constrained, the gains in tax compliance under the optimal assignment will be affected by changes in collectors' visits by household type. Under the status quo assignment, each collector pair is assigned to one low- and one high-type household. Because $v^H > v^L$, both collectors choose to visit the high-type household and not the low-type one.[19] The compliance function under the status quo assignment would thus be $v^H a^{H-H} + v^H a^{L-L}$. Because $a^{H-H} > a^{L-L}$, the optimal assignment function f^* would assign both high-type households to the *H-H* team and both low-type households to the *L-L* team. Due to time constraints, the *H-H* team would visit

[19] Collectors would likely do this if they are paid in proportion to tax compliance, as is the case in this setting, or if they face any kind of promotion incentive based on performance.

one of the high-type households, and the *L-L* team would visit one of the low-type households. Thus, the average compliance would be $v^H a^{H-H} + v^L a^{L-L}$ which is strictly lower than $v^H a^{H-H} + v^H a^{L-L}$. By contrast, if collectors are not time-constrained and effort by collector and household type is not endogenous to the assignment, compliance under the optimal assignment would be $2v^H a^{H-H} + 2v^L a^{L-L}$ which is strictly higher than the compliance under the status quo assignment $(v^H + v^L)a^{H-H} + (v^H + v^L)a^{L-L}$ when $v^H a^{H-H} + v^L a^{L-L} > v^L a^{H-H} + v^H a^{L-L}$.

There are two key assumptions that would generate such a SUTVA violation, each of which we examine in our context. First, tax collectors must choose to visit properties based on the type of the household. Although one might expect collectors to exert more effort in visiting high-type households, other factors like the shoe-leather costs of visiting households might be equally important. Examining heterogeneity in post-registration collector visits by household type, we do find evidence of effort targeting towards high-type households.[20]

The second condition is that collectors are time-constrained.[21] However, several pieces of evidence suggest that collectors did not face binding time constraints when working on the property tax campaign. First, we examine the distribution of tax payments over the month-long tax collection period in each neighborhood. If collectors were time-constrained, then the marginal value of an additional visit should be larger than its marginal cost at the end of the month. Correspondingly, we should expect a steady stream of tax payments until the end of the tax collection period. However, the data reveal that tax payments across neighborhoods are on average close to zero on the last few days of the tax collection period (Figure A22, 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. We investigate this empirically by estimating the relationship between post-registration visits and the number of households in each neighborhood. Because assignment of collectors to neighborhoods was randomized, unobservable collector characteristics were orthogonal to neighborhood size. However, the

[20]Effort targeting is more pronounced for *L-L* teams than *H-H* teams (Figure A8). Specifically, *L-L* teams are 8 percentage points more likely to visit high- than low-type households ($p = 0.045$), and *H-H* teams are 5 percentage points more likely to visit them ($p = 0.17$).

[21]If collectors are able to visit as many households as they want, then changing the assignment would not affect effort provision, even if collectors target their visits toward high-type households.

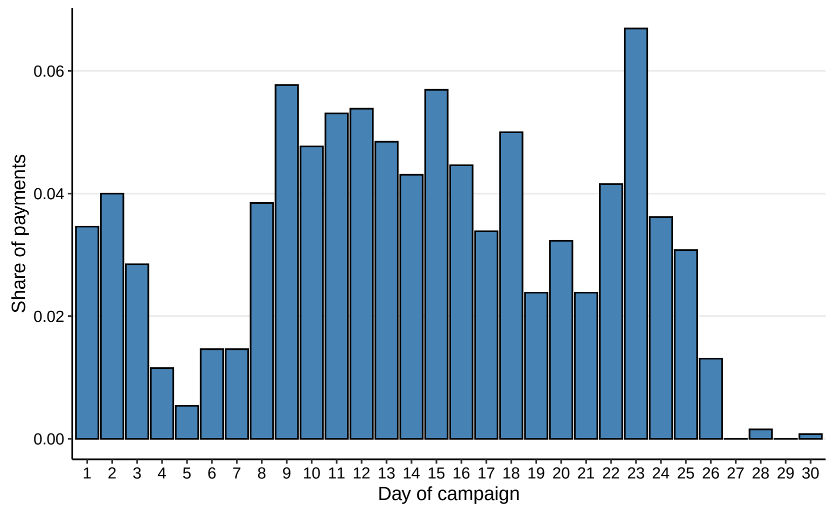
[22]This is unlikely to be explained by collector fatigue given that their activity jumps sharply immediately following the assignment to new neighborhoods in the next campaign month.

data show no significant relationship between neighborhood size and proportion of households visited (Figure A22 Panel B).[23] Taken together, these results suggest that a SUTVA violation is unlikely to arise from endogenous changes in collector effort by match type resulting from tax collectors' time constraints.

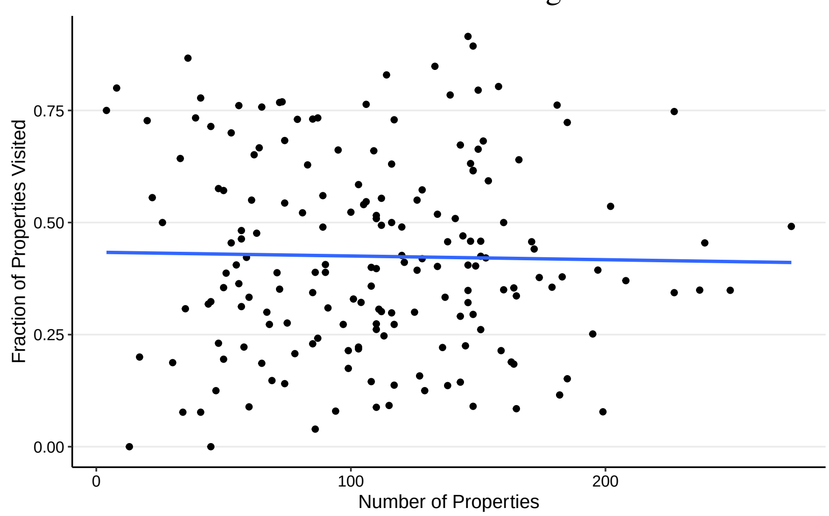
[23] A one standard deviation increase in the number of households (53 households) in a neighborhood has a small and insignificant effect on the likelihood of being visited (1.4pp, $p = 0.29$).

FIGURE A22: 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. 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. We discuss these results in Section 8.2

A8.1.2 Endogenous Effort due to Demoralization

An alternative endogenous effort concern is that assigning low-type collectors to low-type teammates and low-type households — as in the optimal assignment — would demoralize them and lead to lower effort levels by *L-L* collector teams under the optimal assignment than under the status quo assignment. While their individual incentives (piece-rate performance-based wages) would remain unchanged under the optimal assignment, it is possible that they might anticipate lower group productivity or lower future wages, which could result in lower levels of motivation when working on the tax campaign.

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' endline levels of motivation. We rely on measures of motivation from a survey with collectors after the tax campaign concluded. Drawing on the psychology literature (Tremblay et al., 2009), this 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 their type and, more importantly, with the share of low-type teammates they were assigned to during the tax campaign (Table A17) and, separately, the share of low-type households they were assigned to during the campaign (Table A18).

While we do find evidence that low-type collectors exhibited lower levels of motivation at endline (Table A17 and A18, Column 1), we find no evidence that being exogenously exposed to more low-type teammates or low-type households during the campaign undermined collectors' motivation, especially for low-type collectors (Table A17 and A18, Column 2). If anything, low-type collectors' motivation levels appear to have been *less* impacted than high-type collectors by assignment to low type teammates (Table A17, Column 3) and low-type households (Table A18, Column 3). Overall, theses results run counter to the low-type collector demoralization story.

We also investigate a more extreme form of demoralization, namely that low-type tax collectors might drop out of the campaign entirely under the optimal assignment and thus contribute zero revenue to the state (Table A19). We find no evidence that low-type collectors are more likely to drop out (Column 2) or that being exogenously exposed to more 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).[24] Thus, according to the available evidence from the 2018 campaign, it appears unlikely that assignment of low-type collectors to low-types teammates or households under the optimal assignment would trigger demoralization and reduce effort levels of low-type collector pairs (Tables A17-A19).

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. Specifically, we assume low-type collectors cease their work on the campaign immediately (on day 1) and permanently, thereby contributing zero revenue to the state. We then compare the impact of the optimal assignment with low-type dropout to the status quo assignment. As expected, the effect of the optimal assignment on tax compliance is decreasing in the fraction of low-type collectors who drop out of the campaign. That said, the estimated effect remains positive and significant at the 5% level for dropout rates below 25% and significant at the 10% level for dropout rates below 50% (Figure A23).[25] Thus, we find that the optimal assignment would outperform the status quo even for high rates of dropout. As a benchmark, only three tax collectors in our sample (8.82%) did not complete the full 2018 tax campaign.[26]

[24]If anything, Column 4 of Table A19 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.

[25]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.

[26]Moreover, Figure A23 assumes that low-type collectors drop out on day 1 before they collect any revenue. By contrast, in practice low-type collectors would likely work for an initial period before becoming demoralized and dropping out. For example, of the three collectors in our sample who did not complete the six-month campaign in 2018, two worked for two months and one worked for four months. If tax collectors who ultimately drop out of the tax campaign were to similarly work for a few months first, then Figure A18 would *under*estimate the effect of the optimal assignment policy when low-type collectors drop out.

TABLE A17: 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 collectors' 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 collectors' 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 collectors' teammates that were low-type during the tax campaign, and the interaction of both variables. We report robust standard errors. The sample size is reported at the bottom of the table. We discuss these results in Section 8.2.

TABLE A18: 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

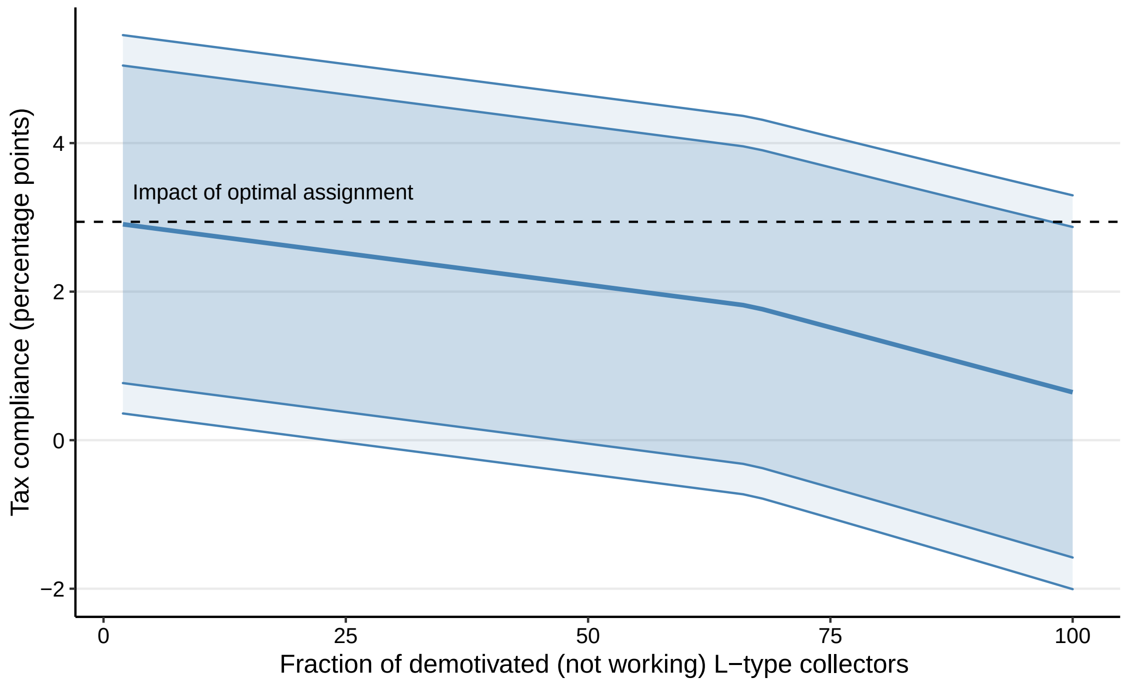
Notes: This table shows the impact of each collectors' own type (Column 1), of the household type they were assigned to (Column 2), and their interaction (Column 3) on endline measures of collectors' extrinsic motivation (Panel A), intrinsic motivation (Panel B), introjection (Panel C), and goal orientation (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 collector's 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 collectors' assignment that were low-type households during the tax campaign, and the interaction of both variables. We report robust standard errors. The sample size is reported at the bottom of the table. We discuss these results in Section 8.2.

TABLE A19: 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 collectors' 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 tax collector not completing the entire property tax campaign (i.e., "dropping out"). We report robust standard errors. The sample size is reported at the bottom of the table. We discuss these results in Section 8.2.

FIGURE A23: EFFECTS OF LOW-TYPE COLLECTOR DROPOUT UNDER THE OPTIMAL ASSIGNMENT



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 grey represent the 90% confidence interval while the one in light grey represents the 95% confidence interval. Standard errors use bootstrap re-sampling (100 samples) at the neighborhood level. The dashed red 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, as reported in Column 1 of Table 2. 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 8.2.

A8.2 Endogenous Learning Dynamics

Our analysis assumes that the potential outcomes and the assignment problem are static. However, according to two types of time dependence, changing the assignment function could impact potential outcomes by match type. First, collectors' ability could vary over time because of learning-by-doing. Second, collectors could learn differentially more from being assigned to certain types of teammates, which could also shape the impact of implementing the optimal assignment policy. We explore each of these possibilities in turn.

A8.2.1 Learning-by-doing

If collectors learn and improve over time, the government might want to first assign collectors to households from whom they will learn the most about tax collection and then deploy them to the other households.[27] 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:

$$y_{hnt} = \gamma \left(X_{c_1(n),t-1} + X_{c_2(n),t-1} \right) + \lambda_t + \varepsilon_{hnt} \tag{18}$$

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 in this context, more opportunities to learn (i.e., more past assignments) should be associated with better tax collector performance and we should find that $\gamma > 0$. The coefficient γ is unbiased given that collectors were randomly assigned to neighborhoods of different size, as described in Section 3.

We find no evidence of learning-by-doing in this context. If anything, increasing the the number of past assignments by 1 SD *decreases* tax compliance by 1.58 percentage points (Table A20, 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).[28] We find similar results when analyzing

[27] As an example, if collectors are more likely to learn when assigned to a high-type household and this especially true for low-type collectors, then the results presented in Section 8 would likely overestimate the effect of the optimal assignment as it would diminish opportunities for learning-by-doing among low-type collectors.

[28] 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. (2021), 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.

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 two pieces of evidence suggest a limited role for learning-by-doing in our setting.

TABLE A20: 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 (18) 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. p-values are presented in brackets. 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 8.2.

A8.2.2 Learning from Teammates

Collectors could also learn from their teammates. For instance, experienced or talented collectors might increase their teammates' performance by sharing skills and knowledge useful for tax collection, such as techniques for convincing households to pay.[29] Whether learning from teammates would create problems for our analysis depends on the functional form of such learning, as we discuss below. In particular, if there are differences in learning from teammates by collector types, then there are implications for our estimates of the impact of the optimal assignment policy.

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:[30]

$$y_{h,n,t} = \delta \cdot \mathrm{E}_{c_1(n),c_2(n),t} + \lambda_t + \varepsilon_{h,n,t} \tag{19}$$

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:

$$\mathrm{Exposure}_{c,t}(l) = \sum_{c' \in C} 1_{[a_{c'}=H]} \cdot 1_{[m_c(t-l)=c']} \tag{20}$$

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.

Second, we examine a cumulative measure that captures collector c's exposure to high-

[29] Such learning might be more pronounced when paired with high-type collectors 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).

[30] 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.

type teammates in all campaign months prior to month t. Formally, it is defined as:

$$\text{Exposure}_{c,t} = \frac{1}{t - t_c^0} \sum_{l=1}^{t-t_c^0} \text{Exposure}_{c,t}(l) \tag{21}$$

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

When estimating learning from teammates, one potential concern is that the estimation of collector c's type might result in systematically overestimating 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. To alleviate this issue, we estimate collector types in a holdout sample, and we perform the empirical analysis in a different sample. The holdout and analysis samples are described in Section 3.

We use these measures to estimate the OLS regression specifications given by Equation (19). 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:

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

$$\text{E}_{c_1(n),c_2(n),t} = \text{Exposure}_{c_1(n),t} + \text{Exposure}_{c_2(n),t} \tag{23}$$

depending on whether past exposure to high-type teammates is defined using $\text{Exposure}_{c,t}(l)$ or $\text{Exposure}_{c,t}$.[31]

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).

[31] 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 (19) and (26).

These results, an important empirical object in their own right, do not on their own constitute a source of bias in our estimation of the impact of the optimal policy. Whether learning will impact our counterfactual estimates depends on the functional form of learning in the tax compliance function. 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:

$$\mathbb{E}\left[y_{ht}|a_1,a_2\right] = m(a_1,a_2) + \left[l(a_1) + l(a_2)\right] \tag{24}$$

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.[32]

We define the learning function of a collector of type a as

$$l(a) = \sum_{a' \in A} g(a')f(a'|a) \tag{25}$$

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. Then, $l(a)$ 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 (24) and (25), 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 $\mathbb{E}\left[y_{ht}|a_1,a_2\right]$ takes the form defined in Equations (24) and (25). 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*

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

[32]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 (24), this restriction is harmless, since accounting for dynamics cannot improve over a static assignment.

Proof:

For a tax campaign month $t > 1$ (at $t = 1$ there is no inter-period learning), the average tax compliance for the assignment function f is given by

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

Let us focus on

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

Thus,

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

Then, 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 $2\sum_{a' \in A} g(a')f^1(a') = 2\sum_{a' \in A} g(a')f^2(a')$ for $f^1(a') = f^2(a')$ for all a' by assumption. $\square$

The main counterfactual assignment function in the paper, the optimal assignment f^*, satisfies the criterion laid out by Proposition 2 since it has the same marginal distribution of types as the status quo assignment f^{SQ}. A functional form that could invalidate Proposition 2 is collector learning that depends on collector type — i.e., if we replace $g(a')$ by $g(a', a)$ in Equation (25).[33] For example, if low-type collectors were better learners than

[33] Additionally, Proposition 2 would not hold if learning is not separable, i.e. if $[l(a_1) + l(a_2)]$ is replaced

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.

We provide evidence on the functional form of the learning function by estimating the following equation:

$$y_{h,n,t} = \gamma_1 \mathrm{E}_{c_1(n),c_2(n),t} \cdot HH_{c_1(n),c_2(n)} + \gamma_2 \mathrm{E}_{c_1(n),c_2(n),t} \cdot LH_{c_1(n),c_2(n)} + \delta \mathrm{E}_{c_1(n),c_2(n),t} + \omega_1 HH_{c_1(n),c_2(n)} + \omega_2 LH_{c_1(n),c_2(n)} + \lambda_t + \varepsilon_{h,n,t} \tag{26}$$

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, 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.

We do not find evidence that low-type collectors are better at learning tax collection skills when exposed to high-type collectors in past tax campaign months. If anything, there is weakly suggestive evidence of more pronounced learning among high-type collectors, i.e., $\gamma_1 > 0$, across measures of past exposure to high-type teammates (Table A21), especially when restricting the sample to high-type households, where all *H-H* pairs are assigned under the optimal assignment (Table A22). More pronounced learning among high-type collectors would mean that our main results underestimate the impact of the optimal assignment. Indeed, high-type collectors, who are better at learning if $\gamma_1 > 0$, would have more opportunities to learn from high-type teammates under the optimal assignment than under the status quo assignment. While the γ_1 coefficients estimated in Table A21 and A22 are consistently positive, they are not statistically significant at conventional levels, making this analysis only suggestive.

by $l(a_1, a_2)$ in equation Equation (24). Although we cannot directly test whether learning is separable, this is a standard assumption in the peer effects literature (e.g., Todd and Wolpin, 2003; Burke and Sass, 2013). We also believe it is likely to hold in the context of door-to-door tax collection where the main scope for learning involves mastering which messages/pitches are most persuasive in seeking to convince property owners to pay.

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 (19), 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 (26), 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. p-values are presented in brackets. The average tax compliance and the sample sizes are reported at the bottom of the table. We discuss these results in Section 8.2.

TABLE A22: LEARNING FROM HIGH-TYPE TEAMMATES (HIGH-TYPE HOUSEHOLDS)

	Tax Compliance					Tax Revenue				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cumulative High-Type Exposure	5.13			3.35		120.43			82.23	
	(2.16)			(1.60)		(49.32)			(42.36)	
	[0.02]			[0.04]		[0.01]			[0.05]	
High-Type Exposure Lag 1		3.53	4.41		1.84		74.44	91.13		16.35
		(2.02)	(2.43)		(1.92)		(48.26)	(60.79)		(40.80)
		[0.08]	[0.07]		[0.34]		[0.12]	[0.13]		[0.69]
High-Type Exposure Lag 2			1.31					43.32		
			(1.19)					(29.16)		
			[0.27]					[0.14]		
Cumulative High-Type Exposure × HH				8.06					240.73	
				(7.82)					(182.90)	
				[0.30]					[0.19]	
Cumulative High-Type Exposure × LH				-1.13					-36.53	
				(2.95)					(66.98)	
				[0.70]					[0.59]	
High-Type Exposure Lag 1 × HH					4.65					164.26
					(4.83)					(111.62)
					[0.34]					[0.14]
High-Type Exposure Lag 1 × LH					-1.15					-14.39
					(2.61)					(56.98)
					[0.66]					[0.80]
Mean	10.36	10.36	8.76	10.36	10.36	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
Observations (Analysis Sample)	4,480	4480	3003	4480	4480	4480	4480	3003	4480	4480

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. The sample is restricted to high-type households. Columns 1–3 and 6–7 report estimates from equation (19), 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 (26), 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. p-values are presented in brackets. The average tax compliance and the sample sizes are reported at the bottom of the table. We discuss these results in Section 8.2.

A9 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.

A9.1 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. *Other Payments.* This is a variable providing the respondent's self-reported informal payments to officials. The underlying exact midline and endline survey question is as follows: 'Now, I'd like to talk about small payments made to officials such as small amounts paid for transport, water, tea, etc. In the past 6 months, did you make any such payment?'

21. *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)

22. *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?'

23. *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?'

24. *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?'

25. *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?'

26. *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?'

27. *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.

28. *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.

29. *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.

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)

A9.2 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 per-

formance 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.'