NBER WORKING PAPER SERIES

TECHNOLOGY AND LOCAL STATE CAPACITY: EVIDENCE FROM GHANA

James Dzansi
Anders Jensen
David Lagakos
Henry Telli

Working Paper 29923
http://www.nber.org/papers/w29923

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2022

For helpful comments and guidance on this project we thank Ray Fisman, Andrew Foster, Gordon Hanson, Asim Khwaja, Craig McIntosh, Karthik Muralidharan, Nii Sowa, Chris Udry, Silvia Vannutelli and seminar participants at Brown, BU, Harvard, NYU Abu Dhabi, Peking HSBC Business School and Williams. For outstanding research assistance we thank Manon Delvaux, Radhika Goyal, Mary Nyarkpoh, Isaac Otoo and Cynthia Zindam. For help implementing a pilot version of the experiment we thank IPA Ghana. For financial support we thank the International Growth Centre, J-PAL and the Harvard Data Science Initiative. All potential errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Technology and Local State Capacity: Evidence from Ghana
James Dzansi, Anders Jensen, David Lagakos, and Henry Telli
NBER Working Paper No. 29923
April 2022
JEL No. H2,H71,O12,O33

ABSTRACT

This paper studies the role of technology in local-government tax collection capacity in the developing world. We first conduct a new census of all local governments in Ghana to document a strong association between technology use and property tax billing, collection and enforcement. We then randomize the use of a new revenue collection technology within one large municipal government. Revenue collectors using the new technology delivered 27 percent more bills and collected 103 percent more tax revenues than control collectors. Collectors using the new technology learned faster about which households in their assigned areas were willing and able to make payments. We reconcile these experimental findings in a simple Beckerian time-use model in which technology allows revenue collectors to better allocate their time towards households that are the most likely to comply with taxpaying duties. The model's predictions are consistent with experimental evidence showing that treatment collectors are more likely to target households with greater liquidity, income, awareness of taxpaying duties, and satisfaction with local public goods provision.

James Dzansi
International Growth Centre
c/o Institute of Statistical, Social
& Economic Research
University of Ghana
Legon, Accra
Ghana
james.dzansi@theigc.org

Anders Jensen
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
and NBER
anders_jensen@hks.harvard.edu

David Lagakos
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
lagakos@bu.edu

Henry Telli
International Growth Centre
c/o Institute of Statistical, Social
& Economic Research
University of Ghana
Legon, Accra
Ghana
henry.telli@theigc.org

1 Introduction

A common feature of most low-income countries is a government that collects little tax revenue and provides few public goods. The literature on state capacity and development argues that the inability to collect taxes efficiently is at the heart of why low-income countries are as poor as they are (e.g. Besley and Persson, 2009, 2013; Besley, Ilzetzki, and Persson, 2013; Dincecco and Katz, 2016; Mayshar, Moav, and Pascali, 2021). This research suggests that the path to economic growth for low-income countries may begin with investing in the government's capacity to collect tax revenues, so as to provide productivity-enhancing public goods.

This paper studies the role of technology in improving government tax capacity. The setting is local governments in Ghana, which are in charge of collecting property taxes but collect very little in practice (Government of Ghana, 2014). As we detail below, the technology in question consists of a geospatial database of properties embedded into an electronic tablet with GPS capabilities. Similar technologies have seen a significant increase in adoption in developing countries over the past decade (Fish and Prichard, 2017). Though to our knowledge, our paper is the first to randomize the presence of a new technology for tax collection to study its impacts.[1]

We set the stage by describing the results of a new census we conducted of tax collection capacity in every local government in Ghana. The census data highlight how poor infrastructure for collecting taxes – with limited street naming and property addressing – shapes collection practices in most areas. Nearly all bills are hand-delivered by collectors to taxpayers, and collectors typically visit individual taxpayers multiple times before collecting (if they collect at all). The majority of tax payments are made in cash and paid directly to revenue collectors. Not surprisingly, government officials cite "leakages" by tax collectors as a significant constraint on their revenues. We show that a minority of local governments have adopted revenue management software and electronic databases of properties, and these governments have significantly better outcomes at every stage in the tax collection process. In particular, they deliver more bills, collect more revenues and have lower non-payment rates than do governments without technology.

The tight empirical link between technology use and tax collection outcomes in the cross-section of local governments naturally invites questions about the direction of causality. To address this issue, we partnered with one large municipal government

[1]Our work complements several prior studies that have leveraged policy reforms to create non-experimental variation in technology usage focusing on technologies which digitize third-party transactions between taxpayers, including Eissa and Zeitlin (2014); Brockmeyer and Somarriba (2022) and Fan, Liu, Qian, and Wen (2021); see Okunogbe and Santoro (2021) for a review of other related studies.

in Ghana and a private technology firm to randomize the use of its technology within the government's jurisdiction. In particular, we randomized the use of a new revenue collection software and geospatial database of properties at the level of a revenue collector. In the experiment, both treatment and control collectors were given a stack of around 135 bills of similar value in a randomly assigned area and tasked with collecting as much revenue as possible in six weeks. The treatment group was given an electronic tablet that uses the geospatial data to make locating households easier. Otherwise the two groups of collectors, and their assigned areas, were observationally similar.

Revenue collectors that use the new technology delivered 27 percent more bills than the control collectors by the end of the study. We view this result as reflecting the mechanical advantage that the technology provides in locating taxpayers more efficiently in an environment with scant property addressing. The time series of cumulative bills delivered in both groups exhibits a concave pattern, as collectors shift emphasis over time from delivering bills to following up with the households that were already served a bill in order to collect payment from them. Revenue collections were 103 percent higher among the collectors assigned to the technology group, on average, implying a much larger effect on revenue collections than on bills delivered. Moreover, we find that the treatment effect on collections grows over time, leading to a rising average effect on the amount collected per bill delivered through the course of the experiment.

We explore several potential hypotheses for why the treatment effect on collections is so much larger than the treatment effect on bill deliveries. One simple story is that households have different attitudes toward payment when visited by collectors who show up with the technology than by "status quo" collectors without the technology. Yet households in treatment and control areas surveyed right after the experiment report statistically similar levels of perceived integrity and ability to enforce tax payments among local government officials. A second hypothesis is that the technology helps reduce leakages, e.g. in the form of payments made by households but diverted by revenue collectors before reaching the local government's coffers. However, several types of household survey questions about the preponderance of bribe payments point to modestly more – rather than less – bribe activity in treatment areas than in control areas.

We argue that the most likely mechanism is that the technology allows collectors to learn about – and focus their scarce time on – households that are most likely to make tax payments. Using surveys of collector behavior and strategies, we show that treatment collectors over time report having better knowledge of individual households' propensity to pay and focus more on collecting from households that are more more able to pay, better aware of taxpaying duties, and more satisfied with local public goods. Importantly, none of these household characteristics would have been known to the collectors at the start of the study period. The implication is that technology allowed collectors to learn, through repeat visits (or longer visits), about hard-to-observe household characteristics. Consistent with this idea, we document that households with greater liquidity and higher income – which are unobservable to the collectors ex ante – are more likely to be targeted by the treatment group than control group.

We formalize this differential learning mechanism in a simple dynamic Beckerian time use model in which forward-looking revenue collectors maximize cumulative revenue collections subject to a time constraint each period. The time constraint is that the total amount of time spent delivering bills and attempting to collect revenue equals a fixed time endowment. Households have either a high probability of payment or a low probability, and the household type is initially unknown to the collectors. Treatment collectors have exogenously higher probabilities of delivering a bill and learning about a household's type, which is meant to capture the mechanical advantage in locating households provided by the technology. The probability of collecting from a given household of each type is assumed to be identical across treatment and control collectors.

When calibrated to match the treatment effects on bill delivery and revenue collection by the end of the experiment, the model largely reproduces the concave time series pattern of treatment effects on bill deliveries – with the largest effects occurring in the middle of the experiment – and the ever-increasing treatment effect on collections over time. Counterfactual simulations show that without faster learning about household types in the treatment group, the treatment effect on bill deliveries and collections would be similar in magnitude. In addition, without behavioral responses by treatment collectors which shift emphasis to collections sooner, the treatment effect on bill deliveries would be counterfactually highest at the end of the period, rather than the middle.

Improved learning through technology has important distributional impacts. The increased information about household income gathered by the treatment collectors, and the subsequent targeting of high-income households, makes the local tax system more progressive. Specifically, technology increases tax payments as a share of taxes due in the top quartiles of the income-asset distribution, but leaves tax payments unchanged in the bottom quartile. However, increased information appears to be a double-edged sword, as technology also increases the incidence of bribes – with effects concentrated in the bottom quartile. Additional analyses suggest treatment collectors also learn about, and subsequently target, those households that are more willing to engage in bribes. This is consistent with our preferred explanation about how technology facilitates learning about which households are more likely to make tax payments.

Our experimental findings on technology investment shed light on the promises and pitfalls of using technology to build tax capacity, and the societal desirability must balance the positive and progressive tax effects against the regressive bribe effects. Our work complements studies that indirectly highlight technology's value by providing taxpayers with incentives or information made available due to its presence (Carillo, Pomeranz, and Singhal, 2017; Okunogbe, 2021; Okunogbe and Pouliquen, 2022). Experimental evidence on technology exists in other governance areas, including social transfers (Muralidharan, Niehaus, and Sukhtankar, 2016) and monitoring (Callen, Gulzar, Hasanain, Khan, and Rezaee, 2020; Dal Bo, Finan, Li, and Schechter, 2021; Vannutelli, 2022). More generally, our results are related to the theories of government which argue that technology investments are central to growth in government size, including due to efficiency improvements (Brennan and Buchanan, 1980; Becker and Mulligan, 2003; Margetts, 2012; Cowen, 2021). At the same time, our bribe results are consistent with historical accounts in the United States, United Kingdom and China, where initial expansions of government have been found to be be associated with increased corruption and bribes (Daunton, 2001; Carpenter, 2020; Cui, 2022). Our results on collector strategies are complementary to other experiments with tax collectors, including on performance-based postings and financial incentives (Khan, Khwaja, and Olken, 2015, 2019), group-work assignments (Bergeron, Bessone, Kabeya, Tourek, and Weigel, 2021), and local leaders (Balan, Bergeron, Tourek, and Weigel, 2020).

Our results suggest that the positive effects of technology on tax outcomes are only partly due to the presence itself of electronic devices embedded with geo-spatial data. Technology allowed collectors to overcome learning constraints in the field (in this case, stemming from navigational challenges) which limited their ability to build information about taxpayers' propensity to pay. Our findings therefore relate to papers which show how *pre-existing* information sources from third-parties can be leveraged to improve collection (Kleven, Knudsen, Kreiner, Pedersen, and Saez, 2011; Pomeranz, 2015; Naritomi, 2019; Balan, Bergeron, Tourek, and Weigel, 2020). Most prior studies place third-party information at the heart of governments' informational capacity (Gordon and Li, 2009; Kleven, Kreiner, and Saez, 2016); our work shows how, in settings where such information-sources are largely non-existent, the state can still strengthen its informational capacity by *directly* building information about taxpayers' propensity to pay.[2]

[2] It is precisely in settings where third-party information coverage and enforcement are constrained, such as sub-national taxation in most developing countries, that information on taxpayer types, including propensity to pay, will be most relevant for tax collection (Luttmer and Singhal, 2014; Dwenger, Kleven, Rasul, and Rincke, 2016) and state legibility (Scott, 1998; Lee and Zhang, 2017).

2 Census of Tax Collection Capacity in Local Government

In order to better understand the determinants of property tax collections, we conducted a census of all local governments in Ghana in 2017. In this section we summarize our findings from this census.

2.1 Design of 2017 Census of Local Governments

We conducted the census of local governments in fall 2017 in collaboration with several national ministries and all 216 local governments in the country. The aim of the census was to collect data on every relevant dimension of the local tax collection system. Three main sets of respondents were interviewed: local government officials; locally elected assembly members; and, citizens. Within the first set, survey responses were collected from every official that participated in the tax collection process, including: the chief executive, the coordinating director, the finance officer, the budget officer, the physical planner, the revenue accountants, and the revenue collectors. Survey modules for officials and assembly members captured information on the tax collection process and demographics and experience. Modules administered to the citizens measured tax morale, knowledge about local taxes and demand for public goods.

The census contains 5,375 citizen responses (approximately 25 per district) and 2,785 local government officials and assembly members (13 per district). In addition to the survey data, we digitized and harmonized administrative records to measure all sources of local tax collection and all types of public expenditure in each district.

2.2 Local Governments have Limited Tax Collection and Information

Our census data allows us to document facts on local tax capacity and its constraints, which are reported in Table 1. In Panel A, we calculate that the average local taxes collected per person is only 4.15 GHC (0.67$ USD) and there is substantial variation across local governments (10^{th} percentile= 0.84; 90^{th} percentile = 6.90). It is useful to consider that taxes collected are determined both by the probability of bill delivery (*delivery margin*) and by the average amount paid conditional on delivery (*payment margin*). On the first margin, we find that, in the average district, only 43 percent of bills are estimated (by local officials) to be delivered (Panel A, Table 1). The delivery margin is thus an important constraint on tax capacity in Ghana, whereas most studies in public finance and development have focused on the payment margin (e.g. Gordon and Li, 2009). On the payment margin, there are significant challenges to extracting tax payments from property owners, conditional on bill delivery. The average tax paid per bill delivered is 11.45 GHC; since delivery is constrained, this is substantially higher than the average tax payment per citizen.[3] The payment margin is both determined by the value of the tax bill, as well as the likelihood of paying the tax bill. According to local officials' estimates, the likelihood of paying the bill is 30.2 percent in the typical district (Panel A, Table 1).

Limited information is an important determinant of constrained tax collection. Indeed, in Panel A of Figure 1, we find that the absence of data on residential and commercial property owners is the most frequently cited constraint on tax collection amongst local bureaucrats, centrally appointed district executives, and locally elected assembly members. The second and third most cited constraints are absence of valuations and enforcement difficulties. The two least-cited constraints are absence of incentives due to central government transfers and resistance by business and residents' unions that deliberate on tax rates with the local assembly. In relation to the delivery margin, absence of information on property owners and their location is cited as the most important constraint on bill deliveries (Panel B, Figure 1). The lack of information starts with the simple absence of precise street addressing: Panel B of Table 1 shows that, in the average district, only 26.8 percent of properties have an official address. Only 28 percent are located on a street with an official name. In the absence of precise information, collectors find it challenging to deliver bills: 37 percent report it being common not to be able to find a property, and 74 percent say that it is common not be able to locate the property owner (Panel B, Table 1).

2.3 Technology and its Relevance for Tax Capacity

Technology has the potential to alleviate collection constraints, including by expanding the information available to the government. In the local tax context, technology can contain two components: a geospatial database of properties; and, the integration of the property information into a software that assists in bill delivery and enforcement. Recent developments in geographic information systems (GIS) reduce the time and cost to integrate new geospatial information sources, including high resolution satellite data, into pre-existing property registries. In addition to the geo-referenced information, the database also contains property characteristics that are used as inputs in the revenue management software to calculate property taxes. Often, the revenue software is accom-

[3] The census information does not directly allow us to measure this variable, so we derive it as follows: $(\text{Amount paid}|\text{Bill delivered})_d = \frac{(\text{Tax per capita})_d}{P(\text{bill delivered})_d}$. This variable is an imperfect measure of taxes collected per bill delivered, since we only have data on the total number of citizens (rather than taxpayers) in each district d. The extent of understatement is related to the ratio of citizens to taxpayers in each district.

panied by the use of point-of-sale devices (PoS), which assist collectors to navigate in the field by providing the geo-location of the property (similar to navigation applications such as Google Maps). In a context with inadequate local property numbering and maps, technology can therefore both expand the initial information available to governments on the existence of properties and improve the ability of collectors to navigate in the field and find properties – thus alleviating the delivery margin.

Technology may also improve the payment margin, through several channels. The existence of technology can act as a signal that the government is more capable, which may both raise households' perception of government enforcement as well as improve their willingness to contribute to local development by paying taxes (Luttmer and Singhal, 2014). By automating the collection process, technology may also reduce collectors' opportunities to privately capture household payments. Finally, technology may help collectors locate property owners more easily in the field and enable better learning about their propensity to pay. We study these mechanisms in detail in Section 6.

In our setting, we proxy technology use by whether the local government in question has a digital database of properties (of any kind) or a revenue software to assist with billing (of any kind). Using this definition, technology is only used by 17 percent of local governments in the country. Of these, 12 percent have both a digital database of properties and a revenue software, while 5 percent have one but not the other (Panel B, Table 1). Adoption of technology is at the discretion of each local government, and the variation in adoption across the country reflects individual governments' choices. Appendix Table A1 provides correlates of adoption choices at the district-level. We find that local governments are more likely to adopt technology in districts where a larger share of properties have official street addresses and property valuations and where legal capacity to enforce taxes is reported to be stronger. Adoption is also positively correlated with overall population size and urban share of the population. We find no correlation between adoption and proxies for citizen tax morale, knowledge of the tax system or trust in local government officials. One interpretation of these results is that technology is complementary to other characteristics that permit higher tax collection (as emphasized by Besley and Persson, 2009).

These patterns of selection contextualize the impacts of technology adoption on tax capacity. While we provide experimental evidence on these impacts within one local government in Section 3 onward, here we leverage the variation in adoption across districts to investigate the association between technology and tax outcomes at the level of entire local governments. When comparing outcomes between local governments with and without technology, the key identification concern is that adoption of the technology may be correlated with other district characteristics which also determine tax collection. We make some headway on this concern through three sets of controls. First, we include the district-covariates which are found to statistically predict adoption (Table A1). Second, we include the (district-specific) share of geographically adjacent districts that have adopted technology, to capture the influence of neighboring policies on local governments' decisions. Third, we include 10 region fixed effects to narrow the comparison between adopters and non-adopters withn each region.

The results are presented in Table 2. Technology is associated with positive outcomes at each step of the collection process: tax collection per capita, share of bills delivered (the *delivery margin*), and taxes per bill delivered (the *payment margin*). The effect of technology on taxes collected per capita ranges from 97.6 percent without covariates to 74.5 percent with district controls (for population, per capita income, urbanization rate and the share of properties with addresses and valuations, and the share of neighboring districts using technology) and region fixed effects. With these same controls, technology adoption is associated with a 22 percent increase in the share of bills delivered and a 38 percent increase in taxes paid per bill delivered.

2.4 Other Constraints and Features of Property Tax Collection

The results in this section suggest that technology may improve tax collection, at least in part by alleviating information constraints. At the same time, local tax capacity remains limited by other factors, which are described in Panel C of Table 1. Only 17.1 percent of properties have official valuations in a typical district. In the absence of official valuations, local governments are forced to tax properties according to a presumptive schedule. In a presumptive schedule, the tax liability of a property is based on a formula that incorporates coarse but easily observable proxies, such as number of floors and windows, quality of the roof and broad geographical location.[4] The presumptive tax schedule caps the tax amount that can be levied on higher property values, thus reducing the payment margin, and curbing the progressivity of property tax rates.

Most payments for property taxes in Ghana are made in cash, and directly to the revenue collector. In the average local government, an estimated 72.1 percent of property tax payments are made in cash directly to collectors, rather than by check or via elec-

[4]Reliance on presumptive tax schedules, rather than more direct capital-valuation based methods, is itself a consequence of the government's limited information. To feasibly implement direct methods requires accurate and continuously updated property value information, including from third-party institutions. Capital-valuation based property taxes are more common in developed countries, while presumptive tax schedules are common in developing countries with administrative and information constraints (Casanegra and Tanzi, 1987; Tadesse and Taube, 1996).

tronic transfers directly to the local government finance office. Cash payments provide collectors with discretion to capture some of the household payments, due to the absence of paper trail coverage that can otherwise help detect this type of unlawful behavior.

Costs of collecting property taxes are generally quite high. As a crude but simple proxy for the collection cost, we take the average monthly salary of revenue collectors as a percent of average monthly revenue collections. In the average local government, the cost of collection is 64.1 percent of taxes collected. In other words, for every 100 Ghanaian GHC in taxes collected, the local government retains only 35.9 GHC. As a frame of reference, the Internal Revenue Service estimates that for every 100 dollars it collects, it retains 99.7 dollars!

One factor that appears *not* to be a strong constraint on tax collection is human capital levels of local government officials per se. Indeed, in the average district, the share of officials (including collectors) with post-secondary education is 67 percent; this contrasts to 9.4 percent amongst citizens. In addition, officials are experienced, having worked an average of 11.7 years in local government.

Local governments do seem to face significant constraints to enforcing tax payments. In principle, properties that are delinquent can be summoned to court. The court summons can result in the confiscation of the owner's property or the shutdown of the business. Thus, court action is in principle the only enforcement tool that effectively has monetary consequences for tax delinquents. Yet in practice, only 22 percent of local governments report taking any tax defaulters to court in the previous year. The reasons for limited court action lie outside the tax administration's immediate scope, and are mainly due to legal constraints or political costs (Panel C, Table 1).

Finally, the intrinsic motivation of citizens to comply with taxation appears to be low. In the average district, 70 percent of surveyed residents think that taxes should only be paid if citizens believe that the payments will be usefully spent by government. In a national context where public use of tax funds is often found to be wasteful (Williams, 2017), the strong norm for conditional compliance likely implies significant resistance by households to make tax payments. Moreover, citizens may be reluctant to comply with their taxpayer duties if they distrust officials. We find that, in the average district, citizens have almost exactly the same level of trust in their local government officials as they have in complete strangers. The low tax morale amongst households is likely to negatively impact compliance along the payment margin.

3 Experiment: Setting and Design

The census results of the previous section showed a strong association between technology use and tax outcomes in the cross-section of local governments in Ghana. This association suggests a potentially important role for technology in alleviating some capacity constraints. Though there is also clear evidence that richer and more urban districts – which on average have more potential for collection – are more likely to adopt technology in the first place. Moreover, there are other stated constraints on tax collections that may or may not be relaxed through technology use, such as political will to collect or legal constraints on enforcement. In this section we describe an experiment conducted with a large local government in urban Ghana. The goal of the experiment is to causally estimate the impacts of technology on tax outcomes and investigate mechanisms.

3.1 Setting

The experiment was embedded in the 2021 property tax campaign in La Nkwantanang Madina Municipal Assembly (henceforth, Madina). Madina is part of the Greater Accra region, and is more affluent and urban than the average district. Madina's local government worked in collaboration with a domestic private firm named Melchia Investments that developed a technology to assist in local tax collection. The technology features the two components described in Section 2.3: a geospatial database of commercial and residential properties and a revenue management software. The geospatial database was created by combining high-resolution aerial photographs with digital registry maps.[5]

Supplemental information was collected from in-person visits to properties. The geospatial information serves as an input to the revenue management software which automates the creation of bills and follow-up notices and records tax payments. At the 'last mile' of the collection process, technology consists of a tablet that assists in the field-work of collectors who deliver bills and collect payments. The tablet provides navigational assistance to help the collector go from an initial point to the location of a designated property (as illustrated in Appendix Figure A1). What varies across treatment and control groups is the presence of the tablet.

Before the campaign, collectors received training from both municipal officers and employees of the private firm. The main sessions, common to all collectors, described the rules for property tax collection in Madina and the protocols to follow during in-

[5] Aerial imagery, including satellite data, is increasingly used for research in economics (see Donaldson and Storeygard (2016) and Michalopoulos and Papaioannou (2018) for reviews). Closely related to our paper is the technology studied in Casaburi and Troiano (2016) which detected property tax evasion by overlaying aerial photographs and property registry data.

teractions with property owners. In addition, the collectors assigned to the treatment group received training in how to use the handheld tablets.

During a fiscal year, the local government assigns collectors to work in designated geographic areas, which we refer to as 'collection units', for around six weeks at a time (see Appendix Figure A2 for illustration). Property owners are legally required to pay within four weeks of receiving the tax bill. While tax payments can be made at designated pay-stations beyond the four-week deadline, most payments in practice are made to collectors during the campaign cycle. Each collection unit is defined with geographical boundaries and creates a cluster of physically adjacent properties (see Figure A2 for an illustration). During a collection cycle, each collector is responsible for delivering and collecting a specific number of bills in one collection unit. At the end of a cycle, the collector is assigned to a new collection unit. Each area of the municipality is only covered once during a fiscal year, as a consequence of the large number of properties in Madina relative to the number of available collectors. Our experiment was specifically embedded in the six-week cycle between March 15^{th} and April 25^{th} in 2021.

At the time of our experiment, all collectors in Madina received an 8 percent commission rate on taxes collected on their assigned properties. Collectors also receive a daily allowance, meant to cover transportation costs, and a monthly salary. The compensation scheme is constant across treatment and control groups.

3.2 Experimental Design

We trained 56 collectors and randomly assigned 28 to the treatment group and 28 to the control group. Of the 56 collectors, 39 had previously worked with the private firm and 17 were newly hired shortly before the experiment began. Of the 39 collectors with previous experience, 11 were designated as 'high performing' by the private firm. In the treatment group, all collectors were given tablet available for use during the six-week tax campaign. Other than the use of the tablet, the treatment group was not provided with any additional advantages.

Collectors work individually in their assigned collection unit, where they are responsible for approximately 135 bills. Each collector has a supervisor available to them during the campaign, who can assist with challenges in the field.[6] All supervisors were randomly assigned to be in charge of both treatment and control collectors. At the be-

[6]One potential concern is that the supervisors provide more assistance to the control group, due to greater navigational challenges, or to the treatment group, in order to improve the perceived performance of the technology. However, we find no differences by treatment status in the amount of supervisor support and monitoring reported by collectors; see Appendix Table A2. Reported monitoring by the supervisor was also not significantly different by treatment status.

ginning of the campaign, each collector was provided with their set of physical bills (see Figure A1 for illustration); in addition, the treatment group was provided with a tablet. As described above, the geospatial data is embedded in the tablet and helps collectors navigate in the field to locate properties. The tablets are also loaded with the tax information contained on the physical bills (Figure A1) – thus, apart from the electronic map, the information provided to collectors was constant across groups.

Randomization and Balance Our randomization proceeds in two steps. First, we randomly assign each collector to a collection unit. Second, we randomly assign the collector-unit pair to the treatment or control group. We stratified on the share of properties in the collection unit that were businesses (rather than residential). To avoid chance imbalances, we run the full randomization 100 times, selecting the run with the minimum *t*-statistics from a series of balance checks on six variables (as in Banerjee, Chassang, Montero, and Snowberg, 2020). Two of these variables are at the collector level: a dummy for previous work experience at the private firm, and a dummy for being a high-performing collector (as assessed by the private firm). The other four variables are at the collection-unit level: total bills to deliver; total taxes (current due and arrears); average current amount due per bill; and average previous pay status per bill (unpaid, partially paid, fully paid).

Table 3 summarizes a series of balance checks. In Panel A, we consider a set of characteristics at the tax bill level, based on administrative registry data. In Panel B we consider characteristics at the collector-unit level. None of the variables are statistically significantly different between groups (using randomization inference) at the 10 percent level or lower. At the bottom of each panel, we also report the F-test from the null hypothesis that the difference in characteristics across variables are all zero. We fail to reject the null at either the tax bill level ($F = 0.71, p = 0.66$) or the collector-unit level ($F = 0.16, p = 0.95$). In Appendix Table A3 we also compare characteristics of households in the treatment areas and control areas. We fail to reject the null that the difference in household characteristics are all zero as well ($F = 1.07, p = 0.38$).

4 Data and Estimation

4.1 Data

For our main analysis, we rely mostly on daily reports by revenue collectors, survey responses from households, and surveys of collector behavior and strategy. Our bills and households are generated from underlying administrative data managed by the

Madina Municipal government and the private firm. In this section we describe these data sources in more detail.

Administrative Data Administrative data at the property level, covering 7,560 residential and business properties, contain information on owner names, property classification, location, current tax due, and tax arrears from previous years. This data set served to create the geographical units for all collectors and to issue all the bills that were to be delivered during the tax campaign.

Daily Collector Reports Our research team collected daily data from each collector on the number of bills delivered and the amount of revenue collected. These data allow us to study the activity of revenue collectors in the treatment and control groups at a high frequency. Compliance with daily reporting was quite high overall, though our small sample of 56 collectors raises concerns about the role of idiosyncratic measurement error in these daily reports. For this reason, our main results with these data winsorize the administrative outcomes at the 95^{th} percentile, separately by group-day.[7]

Collector Surveys Enumerators working for the research team conducted three rounds of surveys with all 56 collectors – at the beginning, middle, and end of the tax campaign. The first round was conducted during the initial week of the campaign; the mid-line during the third and fourth weeks; and, the end-line at the end of the sixth week. The surveys covered challenges in the field, strategies used for bill delivery and collection, and self-assessed knowledge about households, among other topics.

Household Surveys The team of enumerators administered end-line surveys with 4,353 randomly selected households in April and May of 2021. A random sample of equal size was drawn from within each of the 56 collection units. Whenever an initially selected property could not be located or contacted, the enumerator would randomly pick an adjacent property within the same collection unit. The end-line survey covered household characteristics, interactions with and views of collectors, taxation, and beliefs about enforcement and governance.

[7]In principle the tablets record daily data that may be informative, such as date-stamped reported payments by collectors, though this information is only available for the treatment group, making it ill suited for comparisons between the two groups.

4.2 Estimation

Given the random treatment assignment, we use OLS to estimate the causal impacts of technology. The econometric model varies slightly depending on the unit of observation. For outcomes which vary at the day and collector level, we estimate:

$$y_{cd} = \beta_d \cdot \mathbf{1}(Tech)_c + \theta_d + \Omega \cdot X_c + \epsilon_{cd}, \tag{1}$$

where y_{cd} is the outcome for collector c on day d, θ_d are campaign-day fixed effects, and X_c is a vector of time-invariant controls. In the main analysis, X_c only includes strata fixed effects, which are dummy variables for ten deciles of the share of businesses in total properties in each collection unit. In robustness checks, we include additional controls for previous work experience in Madina; a dummy for high quality rating of the collector; total number of bills to deliver; and, average tax due per bill. The dummy $\mathbf{1}(Tech)_c$ takes a value of 1 for all collector-units randomly assigned to the technology treatment. The treatment coefficient, β_d, is indexed by day because the panel-nature of the daily collector reports allows us to estimate dynamic treatment effects for every day of the campaign; we do this by interacting the treatment dummy with individual day fixed effects. Standard errors are clustered at the collection-unit level (56 in total).

For outcomes at the household level, we estimate:

$$y_{hc} = \beta \cdot \mathbf{1}(Tech)_c + \Omega \cdot X_{hc} + \epsilon_{hc}, \tag{2}$$

where h indexes households and c collection units. Standard errors are clustered by collection unit. X_{hc} always includes strata fixed effects; in robustness checks, we also include control variables. At the collection unit, these are the same controls as in equation 1. At the household level, the controls are: previous pay status and property category. The previous pay status measures if the property tax bill in the past year was: paid in full; partly paid; not paid at all. We measure this variable using the administrative data.[8]

5 Experimental Effects on Tax Outcomes

We begin by studying the impacts of technology on bill delivery and tax collection using the collector daily reports. In Figure 2, we show the impacts on bills delivered. Panel A shows the averages by group and day, while Panel B reports the daily treatment coefficients β_d based on estimating equation (1). The treatment group delivers more bills than

[8] In Appendix Table A4, we include more extensive household controls, namely the fixed household characteristics shown in Appendix Table A3.

the control group. This difference initially builds up and peaks by the 24^{th} day, where treatment collectors have delivered 34 more bills than the control group, representing a 58% increase. The gap narrows in the second half of the campaign, where the stock of bills delivered in the treatment group steadies while control collectors continue to hand out bills. The confidence interval around the treatment coefficients is meaningfully wide, likely owing to the limited sample size and number of clusters; notwithstanding, the effect is statistically significant at the 5 percent level in all campaign-days beyond the 10^{th} day. At the end of the campaign, the treatment collectors have delivered 21.5 more bills on average, representing a 27 percent increase over the 80.7 bills in the control group.

In Figure 3, we find that technology causes a large increase in total taxes collected. There are no differences in tax performance during the first week, in which most collectors focus on bill delivery. However, from the second week onward, the treatment group collects at a higher rate; the treatment effect is statistically significant at the 5 percent level on all days and grows over time. At the end of the campaign, the treatment group has collected an additional 856 GHC on average, representing a 103 percent increase over the 829 GHC collected on average in the control group.

Figure 4 shows that the treatment group manages to collect more taxes per bill delivered.[9] This performance measure also grows over time; at the end of the campaign, the treatment group collects 7.8 GHC more per bill delivered, which represents a 117.7 percent increase (control group mean is 6.6 GHC). This result implies that the tax collection impact in Figure 3 is not only driven mechanically by the increase in bills delivered in Figure 2, but by a higher collection rate from each delivered bill. This higher collection rate motivates our investigation of mechanisms in Section 6.

In Figures A3, A4 and A5, we explore the robustness of our experimental estimates from Equation 1 for the main outcomes (bills delivered, taxes collected per collector and tax collections per bill). First, we find that the estimates are quite similar when using non-winsorized outcomes but, as expected, less precisely estimated.[10] Second, the results are also similar, but more precisely estimated, upon including additional covariates. Third, the panel-structure of the data permits the inclusion of collector fixed

[9]This variable is only when at least one bill has been delivered and is therefore endogenous to delivery effort. Results are robust to assigning a value of zero to the remaining collector-days.

[10]Winsorizing outcomes is partly motivated by the small sample size of collectors. In Figure A6, we show an additional robustness check related to the sample size. For bills delivered and taxes collected, we re-estimate equation (1) but in sub-samples which leave out one collector at a time. We retain the daily treatment coefficients β_d from all sub-samples and plot them in Figure A6. The dynamic effects in all sub-samples are strongly comparable to the effects in the full sample; this further guards against the concern that outlier performances by any individual collector drives the average effects.

effects. The inclusion of collector fixed effects means that the estimation of β_d in equation (1) relies on the existence of a time-varying component in the treatment effect. In other words, β_d will reflect the treatment effect based on changes within collector over time. The presence of this type of dynamic treatment effect within collector is consistent with our mechanism evidence on learning in the field over the course of the tax campaign (Section 6.3). The results are robust to the inclusion of collector fixed effects; tellingly, the dynamic effects estimated at the end of the campaign are comparable in magnitude to those obtained in the models without fixed effects.[11]

Complementary – and independent – evidence about the effects of the technology on tax outcomes are available from our households surveys. Table 4 reports the treatment effects on key tax outcomes based on estimating equation (2). Households in the treatment group are more likely to get a visit from a tax collector, get more total visits from tax collectors, and are more likely to have a bill delivered. The impacts on visit probabilities and total visits are statistically significant from zero at the five percent level, whereas the impacts on receiving a bill are positive but insignificant. One potential explanation is the households making excuses for their lack of payment (just 16 percent of these households actually make a payment). A second is that the lack of bill delivery in spite of successful visits reflect the outcome of a collusive bribe, as we discuss in Section 6.2. In terms of magnitudes, the impact of technology on bill delivery is smaller in the household surveys than in the daily collector reports, though we cannot reject the null hypothesis that the two effects have the same effect in percentage terms (p-value 0.30).

Treatment households are more likely to report making a tax payment and report higher payments than control areas. Treatment areas also exhibit higher reported payments conditional on bill delivery. The magnitudes of the effects on collections are again smaller in the household surveys than in the collector reports, though we fail to reject the null hypothesis that the effects are similar across sources (p-value 0.27).[12]

Another important concern is whether the context of the COVID-19 pandemic impacted the results. We conducted a pilot experiment in the spring of 2019 in the same location, using the same technology and the same research protocol (though with a smaller sample of collectors). In that pilot we found qualitatively similar effects as in the main experiment. Qualitatively, both the pilot and main experiment produce an effect on bill delivery which is larger in the middle-periods of the intervention rather than at

[11] Including collector fixed effects also alleviates a bias-concern that is inherently unsolvable in cross-sectional RCT analyses: that treatment and control groups are unbalanced on some unobservable collector-specific (and time-invariant) characteristic which determines the outcome.

[12] In Appendix Table A4, we show that the household-level results are also robust to both the removal of controls and to the inclusion of even more extensive controls than in Table 4.

the end. On the quantitative side, at the end of the interventions, the impact on bills delivered was 32 percent in the pilot versus 27 percent in the main experiment; the impact on taxes collected was 74 percent in the pilot versus 103 percent in the main experiment (see Appendix Figure A7). This suggests that the results of the main experiment were not somehow an artifact of abnormal conditions during the pandemic.

6 Mechanisms Behind Experimental Effects on Taxes

The experimental results show that technology caused an increase in taxes both by improving the bill delivery margin and by improving the payment margin. How did technology allow collectors to improve so much on the payment margin relative to the delivery margin? In this section, we investigate three potential mechanisms: (i) the presence of technology improves citizens' tax morale or strengthens the perceived enforcement capacity; (ii) technology reduces collectors' ability to pocket payments made by households (in the form of collusive or coercive bribes); or, (iii) technology permits learning in the field and improves targeting of households with higher propensity to pay.

6.1 Tax Morale and Perceived Enforcement Capabilities

In a first set of mechanisms, it is possible that technology impacted households' beliefs or tax morale, conditional on having been delivered a tax bill. There are two main 'sub-mechanisms'. First, the presence of technology may stimulate households' tax morale (Luttmer and Singhal, 2014). Tax morale is broadly defined as the non-pecuniary motivations for tax compliance. For instance, the presence of technology may improve households' views that the government is making efforts to collect taxes in more efficient and equitable ways. Technology may also act as a signal that the local government is broadly improving integrity and governance and cares about service delivery. Technology may also stimulate households' sense of reciprocity and increase their perception that they will receive useful public goods in return for making tax payments. Through stimulating tax morale in these different ways, technology may increase households' willingness to comply with taxes. Second, the presence of technology may increase the household's perception of the government's informational capacity and/or its ability to enforce taxes. The increased enforcement perception may raise the household's estimated pecuniary costs of non-compliance, and make them more likely to pay taxes.

We use our household survey to create three indices for tax morale: government efforts to collect taxes in equitable and efficient ways; satisfaction with government services; and, government governance capacity and integrity. Each index is based on several individual questions. We also create an index for information-enforcement, which tracks households' perception of government informational capacity as well as the perceived likelihood that tax delinquents will be subject to enforcement and eventually comply.

In Table 5, we estimate the impact of technology on these different indices, by estimating equation (2). We find null effects on all outcomes.[13] In Appendix Table A5, we find null effects on 15 of the 16 individual underlying questions used to build the indices.[14] Per example, there are null effects on enforcement-related questions such as "Next time the tax collectors come to collect, what percent of households do you think will pay their taxes?" and "Imagine someone refuses to pay taxes – how likely do you think it is that the local government will pursue and enforce sanctions?". There are also null effects on questions about satisfaction with government, including "In your opinion, what has been the overall quality of services offered by the local tax department of Madina?" and "Overall, how would you rate the competency of the local government of Madina?". In Appendix Figure A8, we investigate the possibility that the average null effects mask heterogeneity along the asset-income distribution. For example, it is possible that the presence of technology stimulates tax morale but only amongst more well-off households that are more likely to have paid taxes in the past. We find null effects across the income-wealth distribution.

These null effects do not necessarily imply that technology investments cannot increase households' tax morale or perceived enforcement capacity. Such views are likely shaped in the longer-run (Luttmer and Singhal, 2014) and sustained use of technology in the field may eventually shift beliefs and thereby further increase tax collection. Our experiment, in contrast, is only able to capture short-run effects. It is also possible that the tablet which is randomized in this experiment is not a sufficiently 'large' or salient signal about technology investments; perhaps dissemination efforts to emphasize the government's transition to an electronic property registry integrated with a revenue management system that includes tablets would be sufficiently salient and large-scale to impact morale and enforcement beliefs.

6.2 Bribes

In a second set of mechanisms, technology may have impacted official tax collection on the intensive margin (conditional on bill delivery) by changing the collectors' scope for

[13]Several prior studies have also found that exogenous increases in tax collection effectiveness increase tax payments and bribes but do not impact households' morale or beliefs about government, including Khan et al. (2015) and Balan et al. (2020).

[14]Out of the 16 individual outcomes, the only one that is statistically impacted shows a decrease in the perception that everyone pays their fair share of taxes. If anything, this effect should lower morale.

private capture of household payments. Capture can take the form of a "collusive bribe," where the household and collector agree on a payment made to the collector in exchange for a cessation of follow-up visits. Capture can also take the form of a "coercive bribe," in which the collector pockets tax payments made by the household in combination with a strong threat of retaliation against whistle-blowing.[15]

The impact of technology on private capture is ex ante ambiguous. On one hand, technology may increase collectors' perceived monitoring of their activities in the field, which reduces their ability to take bribes. On the other hand, technology may increase households' perception of collectors' enforcement capacity and raise collectors' bargaining power under collusive bribes or ability to impose coercive bribes. Moreover, some studies on private capture find that the incidence is higher in settings where officials have repeated interactions with citizens; technology may create scope for collectors to learn about households' willingness to engage in bribes or their ability to report predatory behavior. Yet another possibility is that the time savings associated with better navigation frees up more time for the collectors to do all of their previous activities, including attempting to take bribes.[16]

To investigate how the technology affects bribes in our setting, we estimate equation (2) with various outcome measures of collector capture from the household survey. Importantly, due to the illegal and culturally sensitive nature of bribes, these measures came from indirect questions to households about bribe activity, and questions about the household's own bribe payments. For example we ask whether it is likely that collectors *in the household's area* are likely to ask for bribes.[17] We find positive and statistically significant effects of the treatment on whether there was any indication of a bribe payment and the total bribe amounts. The treatment effect on the collusive bribe amount was around 1 percent, and the effective on the coercive bribe amount was around 4 percent. So these are modest effects overall relative to the treatment effects on collections. Though caution is in order when interpreting these magnitudes, since they are not directly about a household's bribe propensity but perceived bribe activity in their area.

Overall, the fact that in all specifications we find positive, rather than negative, effects of the technology on bribe activity suggests that the technology's substantial impact on revenue collections does not work through a decrease in leakage by collectors. To the

[15]Other studies have found that coercive and collusive forms of private capture often co-exist within the same setting (Djankov and Sequeira, 2014; Okunogbe and Pouliquen, 2022).

[16]Other technologies, such as electronic filing of tax returns, can reduce private capture by limiting the extent of in-person interactions between officials and taxpayers (Okunogbe and Pouliquen, 2022).

[17]More detail on each bribe variables is in Data Appendix B.3. The results on bribes from this section are robust to different specifications (Table A4) and different measures of bribe incidence (Figure A13).

contrary: the experiment highlights how there is serious potential downside to technology in tax collection efforts that local governments should be aware of.

6.3 Learning and Differential Targeting

The third mechanism we consider is that collectors leverage the time savings from better navigation to learn more about which households are most likely to make a tax payment. This greater knowledge about the taxpayers then allows the collectors to better target their collection efforts to the households with the highest return on the collector's effort.

Three sets of observations help to motivate this mechanism. First, delivery and collection in the field is characterized by significant challenges, which means that the time constraint on collectors' ability to complete tasks during a campaign appears to be binding fairly strongly.[18] In the first campaign week, for example, 71 percent of control collectors reported finding it challenging or very challenging to locate taxpayers in their collection units. Second, propensity to pay is quite heterogeneous across households due to differences in income, liquidity, knowledge of taxpaying duties, satisfaction with local government or other factors.[19] Third, collectors have limited knowledge ex-ante about which households have high propensity to pay taxes. Indeed, our survey data reveal that 75 percent of collectors at baseline report not having a good understanding of which households are more able and willing to pay (Panel A of Figure 6).[20]

Collector Behavior and Strategy To investigate this mechanism, we start by examining technology's impact on challenges in the field using the collector surveys. From the outset of the campaign, treatment collectors report less navigational challenges and less challenges in locating taxpayers (Figure 5). These gaps in reported challenges are statistically significant in all survey rounds, despite the small sample size. The gaps

[18] In the first survey round, control collectors report that: the average weekly time devoted to work in the field is 19.5 hours; the average time required to deliver a single bill is 1.5 hours. Thus, to deliver the assigned 135 bills would in principle require 10.4 weeks ((135 1.5)/19.5 = 10.4) – while the collectors only have 6 weeks. Moreover, collectors most also devote some time for repeat visits to collect payments.

[19] We find evidence consistent with this intuition in Appendix Table A6. Based on discussions with local officials, we build measures of both willingness to pay, proxied by taxpayer knowledge of taxation, and ability to pay, proxied by income and liquidity, using survey data (Appendix B.4). We create a propensity to pay index by combining the proxies for taxpayer knowledge, liquidity and income. Appendix Table A6 shows that households' propensity to pay strongly predict actual compliance outside of the experiment.

[20] Household propensity to pay is hard to know in this setting for several reasons. First, ability to pay depends on income and liquidity, which are transitory. Second, propensity to pay is only weakly correlated with more easily observable characteristics, including property taxes owed. This is because the property tax in Madina is calculated on a presumptive schedule (Section 2.4), which relies on coarse proxies for capital value (e.g. number of floors and rooms) and therefore breaks the link between property tax owed and income. Consistent with this, in Figure A9 we find that the value of the tax bill accounts for less than 1 percent of the variation in household income.

do decrease at the end of the campaign, suggesting that control collectors also improve their navigation over time. In Appendix Table A2, we find no strong evidence that other challenges in the field significantly differ between groups (e.g. resistance by property owners, wrong bill information, and lack of supervisor support). Due to the improved ability to navigate and locate, treatment collectors spend 48 percent less time per bill delivered; at the same time, there are no differences in total weekly hours worked between groups (Appendix Table A7).

What do treatment collectors spend their freed-up time on? We find that they make more return-visits to property owners (Table 4, column 2). In the context of these repeated interactions, the collector surveys show that treatment collectors increase their knowledge about the types of households that have higher propensity to pay. Indeed, Panel A of Figure 6 shows that in the initial survey round there were no differences in collectors' knowledge about household types. Over time, a positive knowledge gap opens up, as treatment collectors gather more information about households types while doing their return-visits to property owners. This difference in knowledge is statistically significant at the 5-percent level at the end of the campaign. Part of the knowledge gap in the middle and final survey rounds is also due to a decrease in reported knowledge amongst control collectors: this could reflect collectors' updating about knowledge accuracy when confronted with the actual work in the field.

The collector surveys reveal that the treatment group uses this additional information to target those households with high propensity to pay. Mirroring the result on knowledge, Panel B of Figure 6 shows that there were no differences in collection strategies in the initial survey round, but treatment collectors over time increasingly make use of the strategy to visit areas on specific days where property owners are more likely to be able to pay. Figure 7 shows that treatment collectors also increasingly make use of collection strategies which target property owners that are more willing to pay – by visiting households that have a stronger awareness of their duty to pay taxes (Panel A; p-value = 0.07) and that are more satisfied with public goods (Panel B; p-value = 0.06).

It is useful to divide collection strategies into two broad types: those that focus on hard-to-observe household characteristics, and those that focus on easy-to-observe characteristics. We define the former to be focusing on households that are: willing and able to pay taxes, aware of taxpaying duties, or satisfied with public goods delivery. We define the latter to be focusing on areas with: more previous payments; higher bill values; or greater proximity to the main road, one's house, or the company headquarters.

Table 7 reports the use of each of these two broad strategy types at the beginning, middle and end of the experiment. The table shows that differences in strategies are insignificant at the beginning of the experiment. Over time, however, treatment collectors make disproportionately more use of the hard-to-observe strategies, consistent with learning. In fact, the disproportionate reliance on hard-to-observe strategies is even more pronounced, and statistically significant, when we include collector fixed effects. Since the inclusion of such fixed effects isolates the part of the treatment effect which varies within collector over time, this result is strongly consistent with learning over the course of the experimental period.

Which Household Types Get Targeted with the Technology? As a second piece of evidence related to this mechanism, we investigate how targeted households (those that make payment) differ from non-targeted households within a collection unit and, importantly, how technology causes this to differ between treatment and control areas.

The mechanism predicts positive selection under technology on proxies for propensity to pay. We estimate the following selection model using the household survey

$$y_{hc} = \theta \cdot \mathbf{1}(Pay)_h + \beta \cdot [\mathbf{1}(Pay)_h * \mathbf{1}(Tech)_c] + \Omega \cdot X_h + \mu_c + \epsilon_{hc} \tag{3}$$

y_{hc} is a fixed household or property characteristic and where $\mathbf{1}(Pay)_h$ is a dummy which takes a value of 1 if the household made any positive tax payment. In interpreting equation (3), it is important to note that $\mathbf{1}(Pay)_h$ is an endogenous outcome. The coefficient θ indicates whether there is a difference in a fixed household characteristics between targeted and non-targeted households in the control group. The treatment coefficient β shows how the difference in characteristic between targeted and non-targeted households changes in treatment versus control areas; any non-zero β would indicate differential selection caused by technology. Note that, because we are focusing on differences in characteristics between targeted and non-targeted households within collection areas, we can include collection area fixed effects (μ_c).

We focus on three fixed household characteristics of propensity to pay: income, liquidity, and taxpayer knowledge. As argued above, these are hard-to-observe characteristics which local officials identified as determinants of compliance.[21] We also consider characteristics that are more easily observable: tax bill value, previous tax payment, and observable assets. The first two are directly observable on the tax bill. The third charac-

[21]The construction of these variables is described in detail in Data Appendix B.4. Even though these proxies are based on end-line household surveys, we think they are plausibly not impacted by the treatment. It is unlikely that technology-induced payment of taxes affects households' earnings choices within the six-week span of the tax campaign. The questions on liquidity refer to a 'typical' month rather than the specific past month during the campaign. Finally, no property owner from the areas of the experiment was neither summoned to court nor had their property confiscated during the tax campaign.

teristic is derived from the household survey and measures assets that can be observed outside the property (e.g. car, truck, electric generator). Targeting households with these observable characteristics may be useful in general, and specifically for collectors that have less knowledge about households' willingness and ability to pay.

The results from estimating equation (3) are presented in Figure 8. The figure presents both the level of selection in the control group (θ) and the treatment group ($\theta + \beta$); the differential selection (β) can visually be inferred as the difference in levels across groups. In Panel A, we observe that, in the treatment group, targeted households have higher liquidity, income, and taxpayer knowledge than non-targeted households; all differences are significant at the 5 percent level. In contrast, targeted and non-targeted households in the control group are precisely estimated to have no differences in liquidity and income; targeted households have more knowledge than non-targeted, but this difference is not significant at the 5 percent level. A summary index for these proxies of propensity to pay reveals strongly positive and precisely estimated selection in the treatment group; and, almost null and statistically insignificant selection in the control group. Panel B studies selection on more easily observable characteristics. Interestingly, there is little targeting on value of tax bill in both groups – suggesting that such information is not a useful predictor for compliance. We observe positive selection on previous tax payment and assets in both treatment and control groups.[22]

Taken together, the results here paint a picture consistent with the proposed mechanism: technology reduces navigational challenges and frees up scarce time, which the collectors use to learn about households' ability and willingness to pay; based on the improved knowledge, technology induces a shift in strategy away from focusing on more easily observable characteristics and towards targeting of hard-to-observe dimensions of propensity to pay. We formalize this mechanism in Section 7, and show that it can account for the main results – including the dynamic treatment path of higher collection per bill delivered (Figure 4), which grows over time despite more bills delivered.[23]

[22] One concern is that these payment patterns are driven by heterogeneous effects of technology on willingness to pay (per example, by level of taxpayer knowledge). However, Appendix Table A8 finds no differential impacts of technology on tax morale and perceived enforcement capacities by income, liquidity or taxpayer knowledge. Moreover, Appendix Figure A10 shows that selection on bill delivery is symmetric to selection on tax payment. Bill delivery may be a dimension of targeting if collectors learn during their very first encounter with property owners and, given limited time to conduct return-visits, selectively choose who to deliver a bill to. These two results suggest that heterogeneous effects of technology on willingness to pay are unlikely to confound the targeting interpretation of the patterns in Figure 8.

[23] Our results are not consistent with a strategy where propensity to pay is in fact observable and collectors focus initially on those with highest propensity and then move 'down the curve'. Since the treatment collectors deliver more bills, this strategy would generate negative selection on proxies for propensity to pay such as income and liquidity (while we find positive selection). Moreover, such a strategy would generate decreasing collection per bill delivered over time (while we find it increases).

Finally, suggestive evidence shows that learning in the field may also have driven the positive impact of technology on bribes (Section 6.2). In Appendix Figure A11, we find that households exposed to bribes in the treatment group have higher taxpayer knowledge than those not exposed to bribes. This taxpayer characteristic was also targeted by treatment collectors for bill delivery. Within this characteristic, the households targeted are specifically more likely to have witnessed or heard about court actions and property confiscations by the local government. To have seen or heard about effective enforcement may raise the perceived credibility of collectors' threat of retaliation in a setting of collusive or coercive bribes. Thus, this descriptive result is consistent with the learning mechanism: technology provides more time for collectors to learn about household types, including those that are more amenable to collectors' private rent capture.

7 Model

The experimental results of the previous sections uncover larger effects of technology on collections than on bill deliveries. We argued that the data best support an explanation based on faster learning and differential targeting of households by the treatment group. In this section we provide a model that formalizes this mechanism. We then use the model to simulate several counterfactual scenarios to illustrate the importance of learning in driving the experimental results.

7.1 Environment

The experiment lasts D periods. Collectors are endowed with one unit of time each day and split time between delivering bills and trying to collect revenues. On day one each collector is endowed with a large number of bills. Each bill has a face value of one local currency unit. Neither the number of bills or face value amount of the bill is important for the model so we leave them off.

Collectors come in two types: treatment (T) and control (C). The two types differ exogenously in the number of bills they can distribute in a fixed amount of time. In one unit of time a treatment collector can distribute θ_T bills, and a control collector can distribute θ_C bills, where $\theta_T \geq \theta_C$. This is supposed to capture the extra efficiency in searching for households provided by the technology used by the treatment group.

Households also come in two types: "high-types," with a high probability of payment after each visit, and "low types," with a *zero* probability of payment after each visit. The household type is not known to the collectors until after they deliver a bill to that household. In other words, collectors learn about household types only after bill delivery. The treatment collectors have an advantage in learning about whether households are high-types or not. For each bill delivered, treatment collectors have a probability η_T of discovering that the household is a high type, and a probability $1 - \eta_T$ of learning that the household is a low type. Control collectors have probabilities η_C and $1 - \eta_C$ of learning that the household is high type and low type. We assume that $\eta_T \geq \eta_C$, which captures the better opportunities for learning afforded by the technology, which helps the collector navigate the district easier.

The collection technology is exactly the same for treatment and control collectors. Each period, collectors devote time, c, to collecting from each high-type household they have learned about. We assume that each unit of time has diminishing marginal value in collecting from each household in a given day. This could be because repeat visits in the same day reduce the household's willingness to comply, or because the constraints keeping the household from paying earlier in a day are still likely to be binding later in the day (e.g. the household lacks the funds to pay).

We model diminishing returns to collection activity at the daily level as follows. Spending c units of time per bill trying to collect from h bills yields the following probability of collection per bill: λc^{μ}, where $\lambda > 0$ and $0 < \mu < 1$. Hence, the total collections are the following: $\lambda c^{\mu} h$. As a simple example, suppose the collector has identified measure $h = 5$ high-type bills. If they spend $c = 1$ on measure one of bills then they can collect from each of those bills with probability λ. They then have measure $5 - \lambda$ bills left over in the next period.

The collector's choice variables are time spent distributing bills, b, and time spent on each bill trying to collect, c. Note that a collector would never spend a different amount of time on different bills because of the concavity of collection probability in time spent trying to collect. Since $\mu < 1$, the highest returns are for the first minutes spent trying to collect. So optimality implies that all bills should get equal time trying to collect.

Collector's Problem. The goal of a collector is to maximize tax revenues. Collectors have the following state variables each day: h, the number of bills that have been delivered to a household identified as a high type, and d, the day of the experiment. The collectors' choice variables for each period are b, the time spent trying to deliver bills, and c, the time spent trying to collect from each high-type bill. The time constraint for the collector is that $b + ch = 1$ in each day.

The dynamic trade-off for a collector is that more time trying to collect from households today means less time delivering bills that can be collected from tomorrow. Intuitively, in the later periods, collection is a larger priority, while in the earlier periods, delivering bills is more important.

State variables evolve each period according to the collectors' time allocation choices. After spending b units of time delivering bills, the stock of high types learned about increases by $\theta_j \eta_j b$ for collector type $j \in \{C, T\}$. After spending c units of time per bill trying to collect, a fraction λc^μ get collected from, leaving $h(1 - \lambda c^\mu)$ remaining high-type bills for the next period. So the law of motion for known high types with a bill delivered becomes $h' = \theta_j \eta_j b + h(1 - \lambda c^\mu)$.

Let $V(h, d)$ be the present discounted value of having h high-type bills delivered by day d. The collector's dynamic problem is therefore:

$$V(h,d) = \max_{\{b,c\}} \left\{ \lambda c^\mu h + E\Big[V(h', d') \Big] \right\} \tag{4}$$

where $d' = d + 1$, subject to the time constraint, $b + ch = 1$, and the law of motion for high-type bills discovered, $h' = \theta_j \eta_j b + h(1 - \lambda c^\mu)$.

In the last period, collectors spend the maximum time trying to collect. This means that $b = 0$ and c is the maximum amount of time spent on each bill that uses up the collector's full time endowment. If the collector has measure h high-type bills, then she can spend $c = 1/h$ units of time collecting from each bill. The result is $\lambda(1/h)^\mu h$ revenues collected, meaning that $V(h, D) = \lambda(1/h)^\mu h$.

The remaining periods can be solved by backwards iteration. The optimal dynamic program is to allocate time so that the marginal benefit of trying to collect today equals the marginal benefit of delivering more bills to collect on tomorrow. If a collector spends too much time collecting now, the benefit of collecting today will be very low at the margin relative to the value of having more bills in hand for tomorrow. If the collector spends too little time collecting now, the marginal benefit of collecting now will be very high compared to the value of having more bills delivered.

7.2 Quantitative Counterfactuals

We now parameterize the model and use it to simulate the effects of several types of counterfactual changes to the environment. We begin by setting the number of periods to be $D = 6$, so that each period represents a week. We then set $\theta_C = 1$, which is a normalization on bill delivery efficiency in the control group. We set $\eta_C = 0.20$ meaning that 20 percent of households are found by the control group to be high-types. We set the parameters of the collection function to be $\lambda = \mu = 0.5$ The former controls the average level of collections, which is not central to any of our analysis. The latter controls the degree of curvature in collection efforts, and we have found that the results are not very sensitive to other values of μ.

The free parameters to choose are θ_T and η_T, which largely govern the productivity advantage in delivering bills and in finding high-type households. Both are meant to capture the navigational advantages that the technology makes available to the treatment group. Our strategy is to choose values of these two parameters so as to minimize the distance between the model's estimated treatment effect on deliveries and collections and the actual treatment effects in the data. In particular, we match a treatment effect on deliveries by the end of the study of 27 percent, and a treatment effect on collections of 103 percent. This ultimately requires setting $\theta_T = 1.40$ and $\eta_T = 0.35$.

Figure 9 plots the model's predictions for bills delivered, taxes collected, the stock of high-type bills discovered, and the fraction of time spent attempting to collect on bills already delivered. The model does well in reproducing the concave time pattern of bill deliveries, with the largest differences between treatment and control occurring in the model of the time period. The model also gets the convex pattern of collections with the largest treatment effects coming at the end of the experiment period. As the bottom panel shows, the treatment group's faster delivery and learning leads to a larger stock of bills delivered to known high-types during the first half of the experiment, when the collectors are focused mostly on deliveries. This is consistent with the empirical observations that the treatment group is better aware of which households are willing and able to pay during the experiment (Panel A of Figure 6). The bottom right panel shows that the treatment group changes their behavior to focus more on collections than deliveries through the whole study period. Intuitively, the treatment collectors leverage their advantage in delivering and identifying high types to put more time into collection each period, allowing them more chances to collect taxes from each household. This is consistent with the empirical finding that treatment collectors focus more on strategies that target those more likely to make tax payments (Figures 6-7).

Figure 10 plots the model's prediction when the learning advantage is counterfactually shut off (i.e. by setting $\eta_C = \eta_T$.) Note this leaves in place the treatment group's advantage in delivering bills, but gives them no advantage afterwards in visiting those households again to learn whether they are a high type or not. In this counterfactual, the treatment effect on bill deliveries is similar to before, at 31 percent. Though the treatment effect on collections is counterfactually small relative to the data, at 28 percent compared to 103 percent in the data. The bottom panels illustrates why. The stock of high-type bills is still larger than the control group but not by much, meaning that the treatment group does not have much of an advantage in collection relative to the control group. Not surprisingly, the treatment group spends only slightly more time each period delivering bills as the control group (bottom right panel). The result is a treatment effect on collections that is similar in magnitude to the treatment effect on bills delivered.[24]

In a second counterfactual, which we do not plot for brevity, we simulate the effects of counterfactually removing the advantage in delivering bills from the treatment group ($\theta_T = \theta_C$), but restoring the advantage in learning about household type. In this counterfactual the model predicts a treatment effect on bill delivery of -10 percent. How could this be true if the treatment and control groups are identical in their bill delivering probabilities? The answer is that the treatment collectors endogenously shift their time allocations toward collecting rather than delivering. This leads to faster learning, and higher collections, relative to the treatment group, at the expense of fewer deliveries. The treatment effect on collections ends up being substantial, at 47 percent, illustrating that the advantage in identifying high-types drives a lot of the treatment effect on collection relative to the treatment effect on bill delivery.

In summary, the empirical patterns of deliveries and collections that arise from the experiment in Madina are consistent with a simple model of revenue-maximizing and rational tax collectors with binding time constraints each period. The model predicts that the delivery advantage offered by the technology by itself is not enough to match the much larger treatment effects on collections than deliveries. The model requires, in addition, an advantage in identifying high-type households, who are more likely to pay taxes than the average tax payer. In the model, as in the data, collectors using the new technology build up informational advantages about household types, and spend more of their scarce time trying to collect taxes from those most likely to pay.

8 Distributional Implications of Technology

Our analysis has revealed that households with higher income are more likely under technology to receive a bill and make tax payments. At the same time, households with more assets are slightly less likely to be targeted. These targeting results have important implications for the distributional impacts of technology.

We create four quartiles of the joint income-wealth distribution. This distribution is an aggregate index which captures information on both the household's income and total assets. We estimate distributional effects in the household survey by allowing the technology treatment to vary across the income-wealth quartiles (q):

[24]In Appendix C.1 we conduct a complementary exercise to investigate the importance of the 'mechanical' advantage whereby treatment reduces time-constraints. We leverage additional data-sources to measure the days between the delivery date and the end of the tax campaign at the bill-level. We then control for the 'mechanical' time-advantage and study how much of the treatment effect on tax collection remains; while imperfect, this exercise suggests that the mechanical time-advantage can account for 10-15% of the overall treatment effect on tax collection.

$$y_{hqc} = \beta_q \cdot \mathbf{1}(Tech)_c \cdot \mathbf{1}[Quartile = q] + \Omega \cdot X_{hc} + \epsilon_{hc} \quad (5)$$

The results from estimating equation (5) on the extensive margin of tax compliance are presented in Panel A of Figure 11. Due to the targeting of higher income, we find that technology strongly improves the equity of the local tax system: treated households in the top two quartiles are 9.8 to 10.1 percentage points more likely to comply than control households, representing a 65 percent increase. In contrast, technology has a somewhat precise null effect on tax payments in the bottom income-wealth quartile. In Appendix Figure A12, we show that these distributional results are robust to additional measures of compliance – importantly, technology raises taxes paid as a percent of taxes due in the top quartiles. Since technology does not plausibly impact household income, this result implies that technology improves the progressivity of the tax system (it makes taxes paid, as a share of income, more positively correlated with household income).

Our mechanism analysis also revealed that bribes increased in the treatment group (Section 6.2). What is the distributional impact of technology on bribe-incidence, and how does it compare to the progressive impact on formal tax payments? By estimating the distributional model (equation 5), we find that technology has a positive impact on bribes (on the extensive margin) which is entirely concentrated amongst households with lower income-wealth. This result is presented in Panel B of Figure 11 which, by visual comparison with Panel A, highlights that technology's impacts on formal tax payments and informal collector payments affect different segments of households.[25]. In Appendix Figure A13, we show that the distributional result is robust to using various measures of bribes. Importantly, the percent increase in bribes is also concentrated in the bottom quartiles. This result implies that technology makes the bribe system more regressive (it makes bribes paid, as a share of income, more negatively correlated with income).[26]

Thus, technology increases both formal tax payments and informal bribe capture payments, but with starkly contrasting incidence impacts: formal payments are progressive and concentrated amongst more well-off households, while informal payments are made more regressive and are concentrated amongst less well-off households.

[25] Khan et al. (2015) and Gauthier and Goyette (2014) also find that formal and informal payments are substitutes at the household level

[26] Our finding that bribes are regressive is consistent with previous work which also finds that 'street-level' bribes affect poorer households to a larger extent (Khan et al., 2015; Emran et al., 2013; Peiffer and Rose, 2018). Fried et al. (2010) show that government officials are more likely to demand bribes from poorer individuals because they associate wealth with the capacity to retribute.

Discussion Taken together, the results and model show how technology investments increase the informational capacity of the state. Our findings suggest that the technology provided only part of the information that was ultimately key to unlocking extra revenue collection. The geo-spatial data embedded in the tablet allowed collectors to overcome navigational challenges, which in turn freed up effective time that collectors spent on gaining the ultimately relevant information about households' propensity to pay. Other information at the property-level (value of tax bill and past tax payment) was common in treatment and control groups; in fact, treatment collectors ended up relying no more on this initial information yet performed better than control collectors.

This finding resonates with Scott (1998), who argues that a policy to construct information on citizens is more likely to strengthen state capacity if it builds local, practically relevant knowledge (e.g. propensity to pay) rather than if it mainly establishes formal, generic information (e.g. value of tax bill).[27] One implication is that additional expansions of digital registries with generic property-information may not lead to additional tax revenue if collectors face sufficiently severe time-constraints that make them unable to use this initial information to build locally relevant knowledge.[28]

9 Conclusions

This paper has studied the role of technology in improving local state capacity in Ghana. Both cross-sectional census data and experimental evidence point to a close connection between use of technology and outcomes at every stage of the tax collection process: technology increases the share of bills that are delivered as well as the amount of taxes collected per bill delivered and renders the local tax system more progressive. Additional results suggest that technology helps collectors to learn in the field about households' propensity to pay and focus their collection efforts on those more likely to pay. We formalize this learning mechanism in a dynamic time-use model of revenue-maximizing collectors. The model predicts that much of the gains from technology come from better opportunities to learn about which household are most likely to comply with taxpaying duties. Several types of experimental evidence support this prediction, including the finding that collectors equipped with the technology are more likely to target households with higher income, greater liquidity and more awareness of taxpaying duties.

[27]What constitutes locally appropriate information will naturally vary across settings and depend on the broader government capacity as well as social and economic structures.

[28]In Madina's context, the time-constraints arise due to limited collector staffing and a large set of properties to visit. Limited staffing is itself a consequence of the scarce revenue available to the local government – thus highlighting how limited revenue collection reinforces constraints to building tax capacity.

The literature has argued that state capacity depends on 'legibility,' or the breadth and depth of the state's knowledge about its citizens and their activities (Lee and Zhang, 2017). This paper provides direct evidence that investments in technology alleviate information constraints and thereby increase tax collection. Interestingly, our results suggest that technology did not directly provide the information on property owners that was ultimately key to unlocking extra tax collection. Rather, our evidence suggests that the geo-spatial information embedded in the technology allowed collectors to overcome navigational challenges, which freed up effective time that collectors spent on gaining information about households' propensity to pay. Prior work has shown that pre-existing information can be leveraged to improve collection, including from local leaders (Balan et al., 2020) and third-party sources (Kleven et al., 2011; Pomeranz, 2015; Naritomi, 2019). Our findings complement these studies by showing how governments can directly build information through realistically feasible policy investments. The need to directly build information is most relevant in developing countries' contexts where the limited presence of reliable third-party sources create the strongest information constraints.

Our results caution that improving the state's knowledge is a double-edged sword. Indeed, while formal tax payments become more progressive as a result of increased targeting of high-income households, technology also allows collectors to increase bribes, which disproportionately affected the less well-off households. The societal desirability of technology investments is thus unclear. Reflecting this ambiguity, 88% of treatment households self-report a preference for the technology-based collection system over the manual status-quo system but treated households also report a statistically significant dis-interest in engaging with the state – concentrated among households that were subject to increased bribes. Investigating which complementary policies can limit the adverse bribe impacts while maintaining the positive revenue effects of technology investments is an important area of future research.

Use of technology for local taxation is limited but growing in Africa, and many policymakers are looking for ways to stimulate faster adoption. Future research can fruitfully study the impacts and barriers to tax collection technologies outside of urban or peri-urban areas, where income levels are lower and internet and computer access may be more limited. The impacts of new collection technologies on public expenditures is also a key topic for future research, since the ultimate goal is of course the provision of useful public goods, and not revenues per se. More work is required to rigorously establish the extent to which these public goods improvements occur in practice following the adoption of new tax collection technologies.

References

Balan, P., A. Bergeron, G. Tourek, and J. L. Weigel (2020): "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the D.R. Congo," Working paper.

Banerjee, A. V., S. Chassang, S. Montero, and E. Snowberg (2020): "A Theory of Experimenters: Robustness, Randomization, and Balance," *American Economic Review*, 110, 1206–30.

Becker, G. and C. Mulligan (2003): "Deadweight Costs and the Size of Government," *The Journal of Law and Economics*, 46.

Bergeron, A., P. Bessone, J. Kabeya, G. Tourek, and J. Weigel (2021): "Optimal Assignment of Bureaucrats: Evidence from Randomly Assigned Tax Collectors in the DRC," Unpublished Working Paper.

Besley, T., E. Ilzetzki, and T. Persson (2013): "Weak States and Steady States: The Dynamics of Fiscal Capacity," *American Economic Journal: Macroeconomics*, 5, 203–235.

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

——— (2013): "Taxation and Development," *Handbook of Public Economics*, 5, 51–110.

Brennan, G. and J. Buchanan (1980): *The Power to Tax*, Cambridge University Press.

Brockmeyer, A. and M. S. Somarriba (2022): "Electronic Payment Technology and Tax Compliance: Evidence from Uruguay's Financial Inclusion Reform," *World Bank Policy Research Working Paper*.

Callen, M., S. Gulzar, S. A. Hasanain, M. Y. Khan, and A. Rezaee (2020): "Data and Policy Decisions: Experimental Evidence from Pakistan," *NBER Working Paper Series*.

Carillo, P., D. Pomeranz, and M. Singhal (2017): "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement," *American Economic Journal: Applied Economics*, 9, 144–64.

Carpenter, D. (2020): *The Forging of Bureaucratic Autonomy*, Princeton University Press.

Casaburi, L. and U. Troiano (2016): "Ghost-House Busters: The Electoral Response to a Large Anti-Tax Evasion Program," *Quarterly Journal of Economics*, 131, 273–314.

Casanegra, M. and V. Tanzi (1987): "Presumptive Income Taxation: Administrative, Efficiency, and Equity Aspects," *IMF Working Paper Series*, 1987.

Cowen, T. (2021): "Does Technology Drive the Growth of Government?" in *Essays on Government Growth*, Springer, 51–65.

Cui, W. (2022): *The Administrative Foundations of the Chinese Fiscal State*, Cambridge University Press.

Dal Bo, E., F. Finan, N. Li, and L. Schechter (2021): "Information Technology and Government Decentralization: Experimental Evidence from Paraguay," *Econometrica*, 89, 677–701.

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

Dincecco, M. and G. Katz (2016): "State Capacity and Long-run Economic Performance," *The Economic Journal*, 126, 189–218.

Djankov, S. and S. Sequeira (2014): "Corruption and Firm Behavior: Evidence from African Ports," *Journal of International Economics*, 94, 277–94.

Donaldson, D. and A. Storeygard (2016): "The View from Above: Applications of Satellite Data in Economics," *Journal of Economic Perspectives*, 30, 171–198.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2016): "Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany," *American Economic Journal: Economic Policy*, 8, 203–232.

Eissa, N. and A. Zeitlin (2014): "Using Mobile Technologies to Increase VAT Compliance in Rwanda," Unpublished Working Paper.

Emran, S., A. Islam, and F. Shilpi (2013): "Admission is free only if your dad is rich," *World Bank Policy Research Working Paper*, 6671, 1–48.

Fan, H., Y. Liu, N. Qian, and J. Wen (2021): "Computerizing VAT Invoices in China," *NBER Working Paper Series*.

Fish, P. and W. Prichard (2017): "Strengthening IT Systems for Property Tax Reform," *African Property Tax Initiative Brief*, 1–8.

Fried, B., P. Lagunes, and A. V. and (2010): "Corruption and Inequality at the Crossroad," *Latin American Research Review*, 45, 76–97.

Gauthier, B. and J. Goyette (2014): "Taxation and corruption: Theory and firm-level evidence from Uganda," *Applied Economics*, 45, 2755–2765.

Gordon, R. and W. Li (2009): "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93, 855–866.

Government of Ghana (2014): "Internally Generated Revenue Strategy and Guidelines: Maximizing Internally Generated Revenue Potentials for Improved Local Level Service Delivery," Ghana Ministry of Finance.

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

Khan, A. Q., A. I. Khwaja, and B. A. Olken (2015): "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *Quarterly Journal of Economics*, 131, 219–271.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011): "Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark," *Econometrica*, 79, 651–692.

Kleven, H. J., C. T. Kreiner, and E. Saez (2016): "Why can Modern Governments Tax so Much? An Agency Model of Firms as Fiscal Intermediaries," *Economica*, 83, 219–246.

Lee, M. and N. Zhang (2017): "Legibility and the Informational Foundations of State Capacity," *Journal of Politics*, 79, 118–32.

Luttmer, E. F. P. and M. Singhal (2014): "Tax Morale," *Journal of Economic Perspectives*, 28, 149–68.

Margetts, H. (2012): *Information and Technology in Government*, Routledge.

Mayshar, J., O. Moav, and L. Pascali (2021): "The Origin of the State: Land Productivity or Appropriability?" Unpublished Working Paper.

Michalopoulos, S. and E. Papaioannou (2018): "Spatial Patterns of Development: A Meso Approach," *Annual Review of Economics*, 10, 383–410.

Muralidharan, K., P. Niehaus, and S. Sukhtankar (2016): "Building State Capacity: Evidence from Biometric Smartcards in India," *American Economic Review*, 106, 2895–2929.

Naritomi, J. (2019): "Consumers as Tax Auditors," *American Economic Review*, 109, 3031–72.

Okunogbe, O. (2021): "Becoming Legible to the State: The Role of Detection and Enforcement Capacity in Tax Compliance," *World Bank Policy Research Working Paper Series*.

Okunogbe, O. and V. Pouliquen (2022): "Technology, Taxation and Corruption: Evidence from the Introduction of Electronic Tax Filing," *American Economic Journal: Economic Policy*, Forthcoming.

Okunogbe, O. and F. Santoro (2021): "The Promise and Limitations of Information Technology for Tax Mobilization," *World Bank Policy Research Working Paper Series*.

Peiffer, C. and R. Rose (2018): "Why Are the Poor More Vulnerable to Bribery in Africa? The Institutional Effects of Services," *Journal of Development Studies*, 54, 18–29.

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

Scott, J. (1998): *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*, Yale University Press.

Tadesse, H. and G. Taube (1996): "Presumptive Taxation in Sub-Saharan Africa: Experiences and Prospects," *IMF Working Paper Series*, 1996.

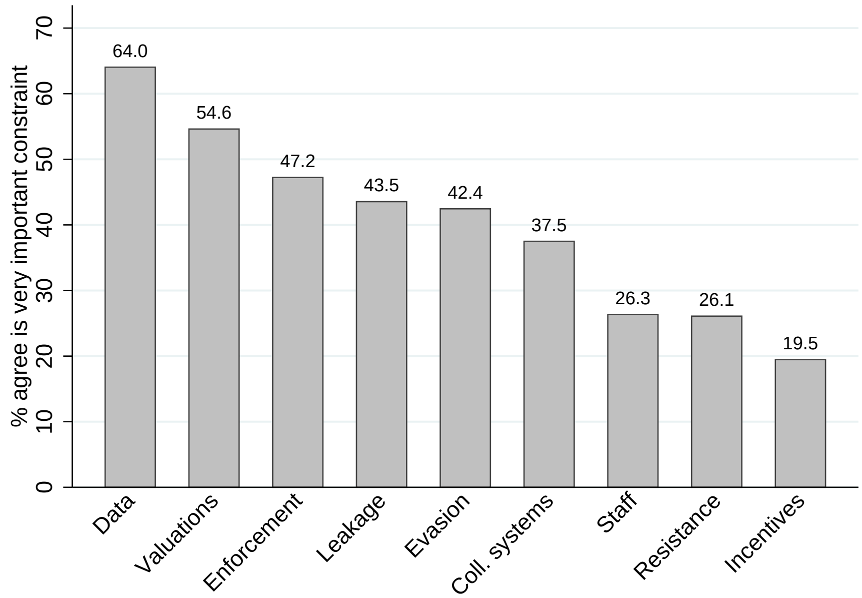
Vannutelli, S. (2022): "From Lapdogs to Watchdogs: Random Auditor Assignment and

Municipal Fiscal Performance," Unpublished Working Paper.

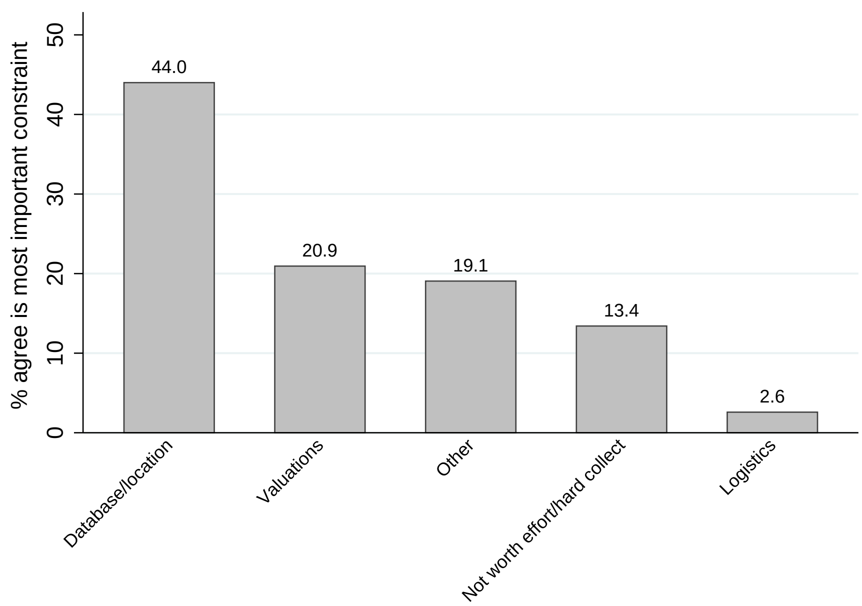
Williams, M. (2017): "The Political Economy of Unfinished Development Projects: Corruption, Clientelism, or Collective Choice?" *American Political Science Review*, 111, 705–723.

Figure 1: Constraints on Tax Collection and Bill Delivery

(a) Perceived Importance of Different Constraints on Tax Collection



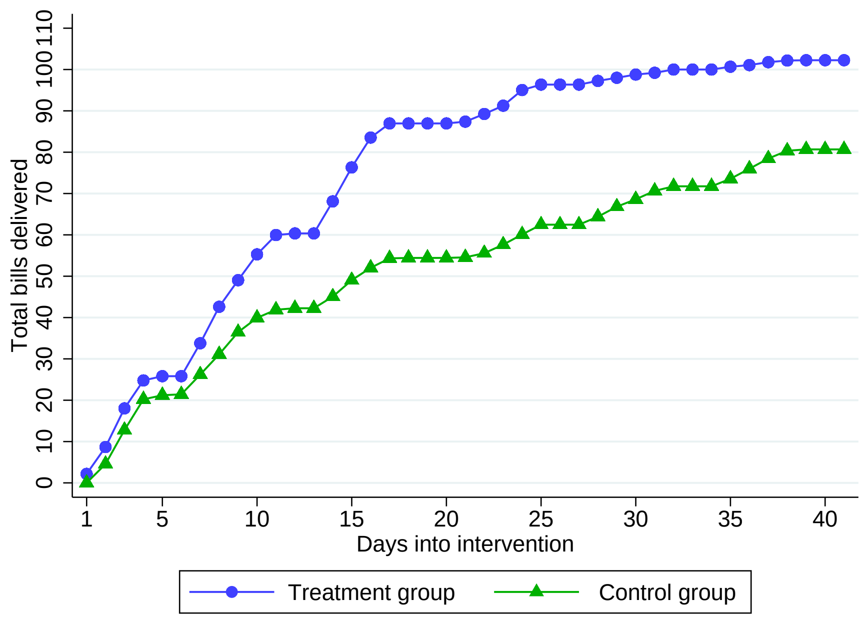
(b) Most Important Perceived Constraint on Bill Delivery



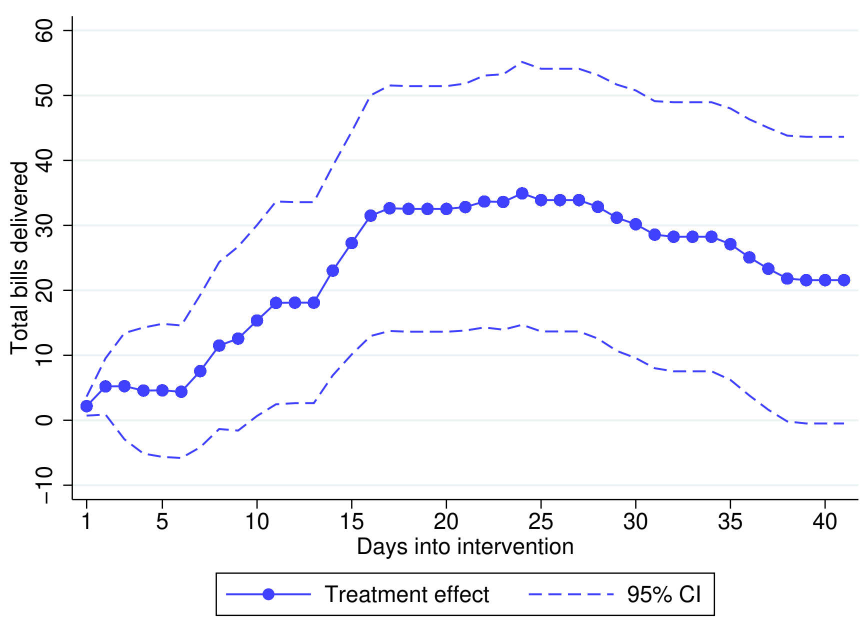
Notes: These panels show the perceived constraints on tax collection and bill delivery as reported by local government officials and politicians. In Panel A, the bars show the percent of all respondents that consider a particular constraint to be 'most important', on a five-choice scale from 'least important' to 'most important'. In Panel B, the bars show the percent of all respondents who consider a particular constraint to be the most important constraint (mutually exclusive choices). Responses are pooled across all 216 local governments.

Figure 2: Impact of Technology on Bills Delivered

(a) Bills Delivered per Collector By Group



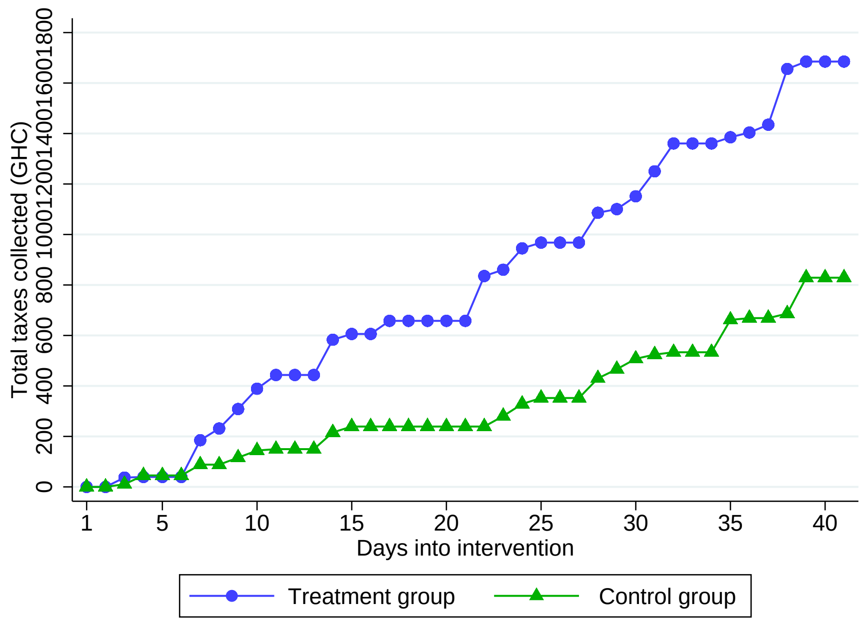
(b) Treatment Effect



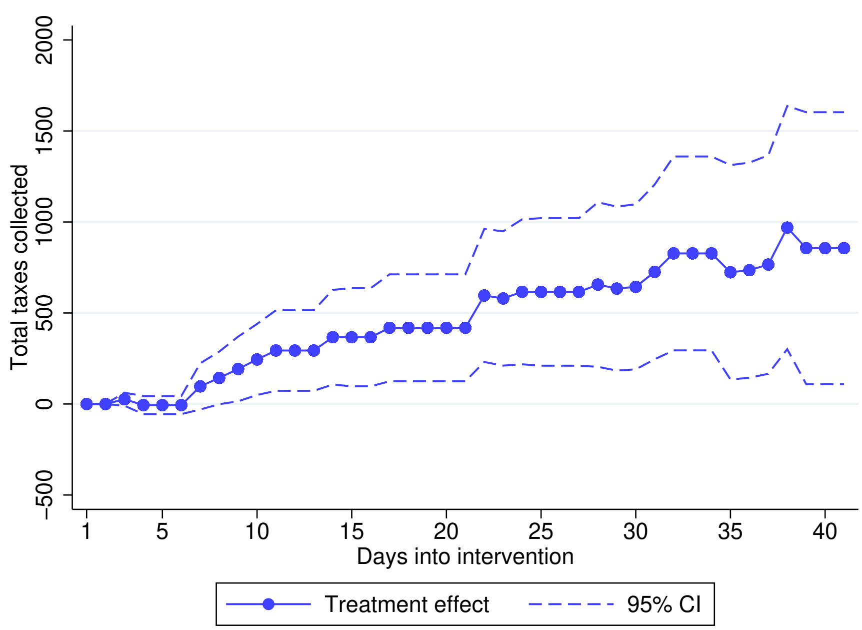
Notes: These panels show the impact of technology on the number of property tax bills delivered. Panel A shows the average number of bills delivered by group and by day of the intervention. Panel B displays the treatment effect coefficients on technology, separately by day, based on estimating equation (1). The analysis is based on the daily administrative data, described in Section 4.1.

Figure 3: Impact of Technology on Taxes Collected

(a) Taxes Collected per Collector by Group



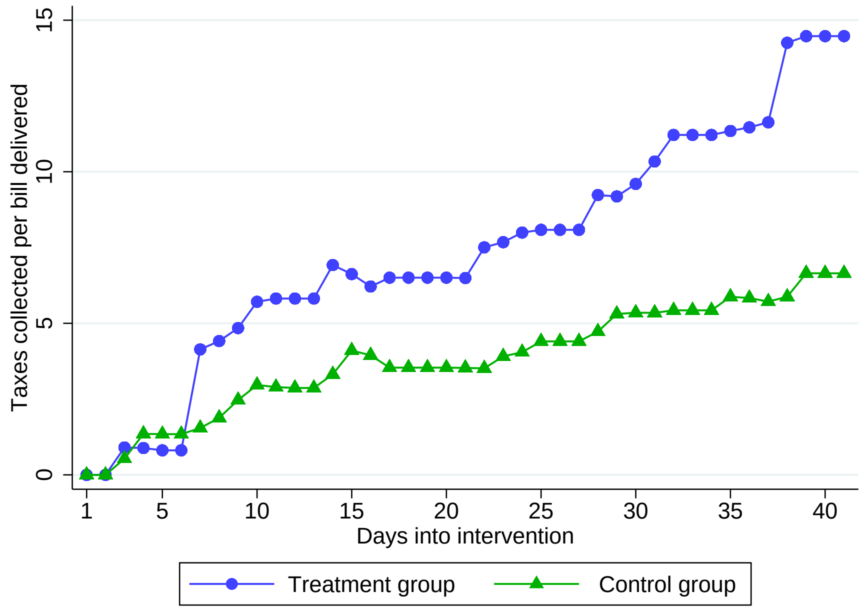
(b) Treatment Effect



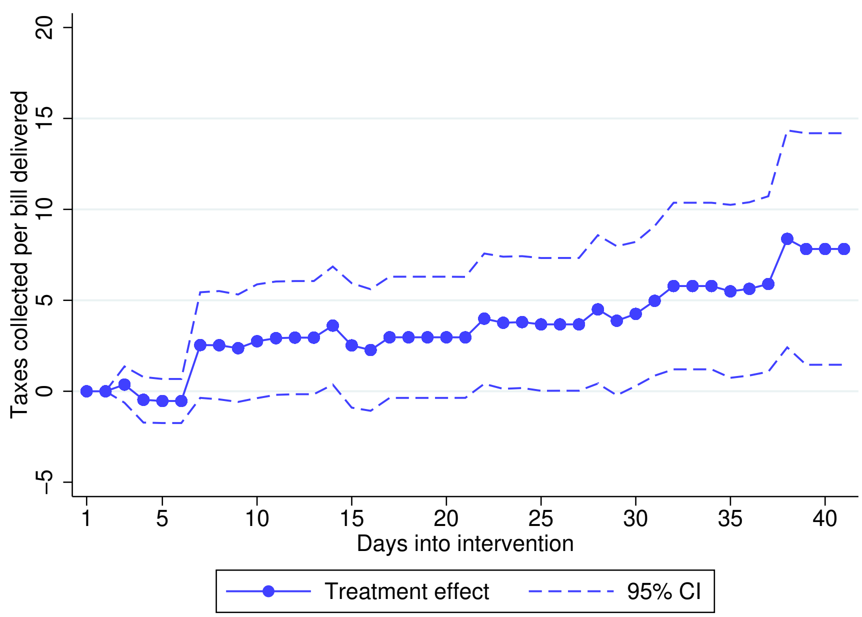
Notes: These panels show the impact of technology on the amount of property taxes collected. Panel A shows the average total amount of taxes collected by group (treatment, control) and by day of the intervention. Panel B displays the treatment effect coefficients on technology, separately by day, based on estimating equation (1). The analysis is based on the daily administrative data, described in Section 4.1.

Figure 4: Impacts of Technology on Taxes Collected per Bill Delivered

(a) Taxes Collected per Bill Delivered by Group



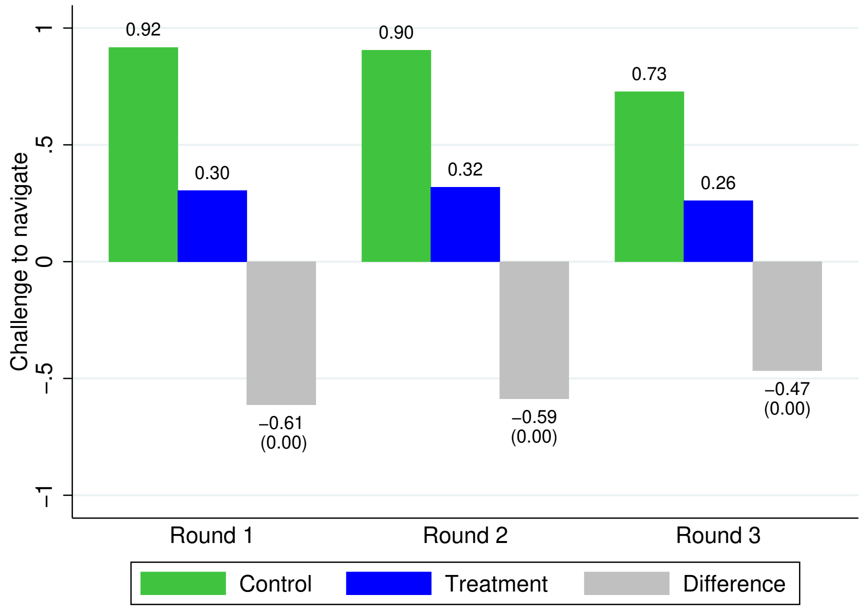
(b) Treatment Effect



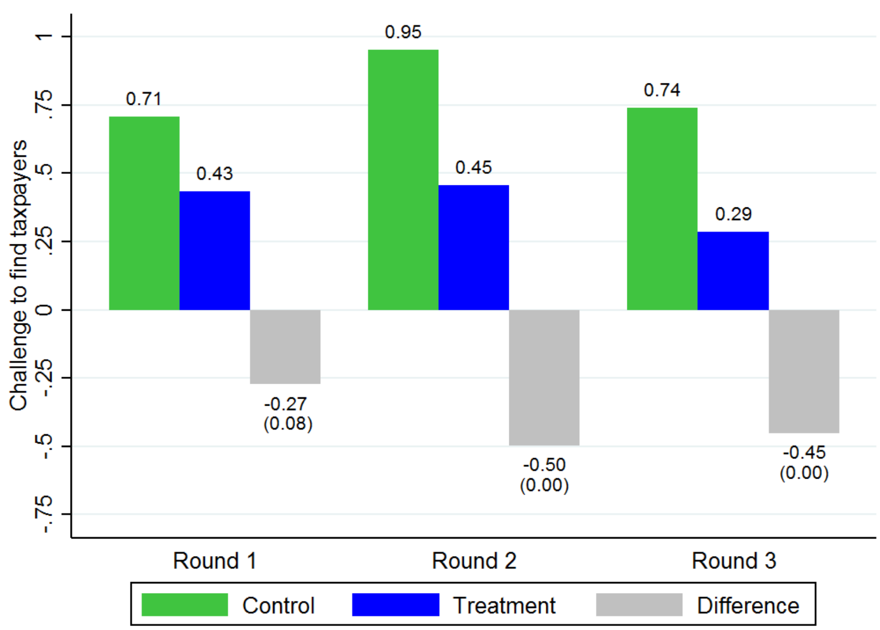
Notes: These panels show the impact of technology on the amount of taxes collected per bill delivered. Panel A shows the average amount of taxes collected per bill delivered by group (treatment, control) and by day of the intervention. Panel B displays the treatment effect coefficients on technology, separately by day, based on estimating equation (1). The analysis is based on the daily administrative data, described in Section 4.1.

Figure 5: Impact of Technology on Search Challenges in the Field

(a) Challenging to Navigate in the Field?



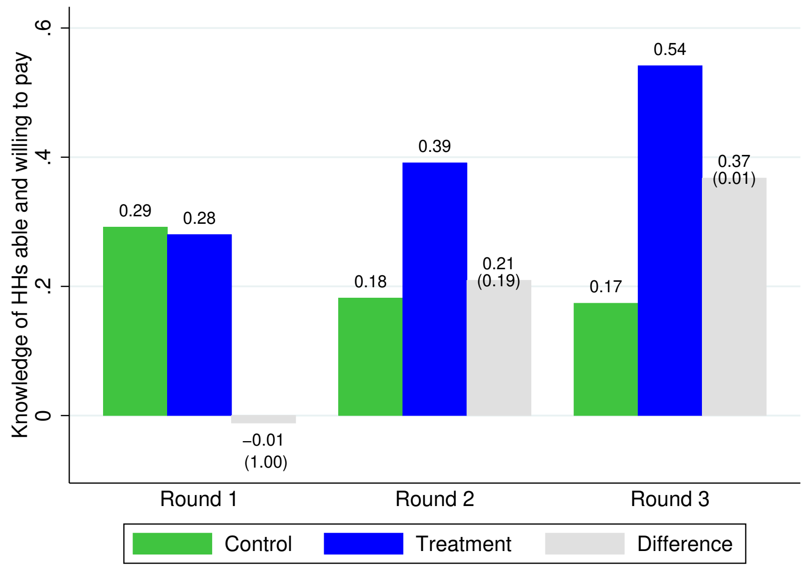
(b) Challenging to Locate Taxpayers?



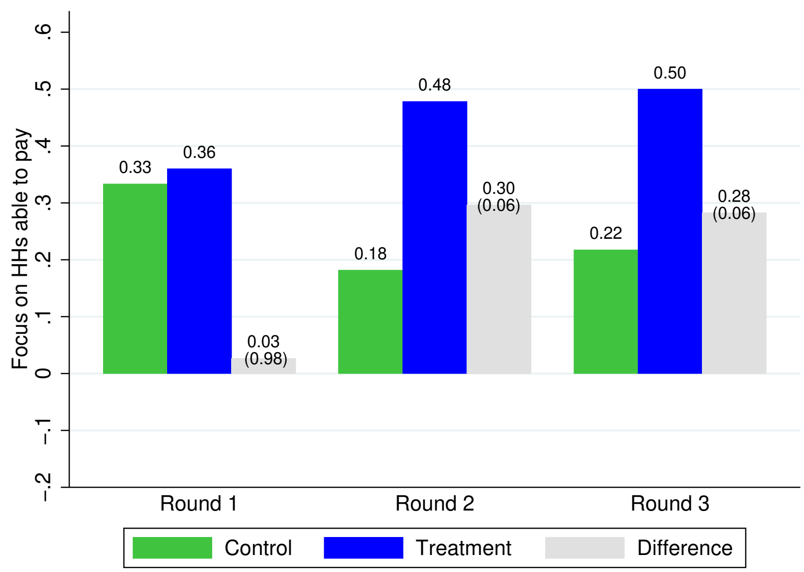
Notes: These panels show the impact of technology on search challenges in the field. Panel A shows the extent to which collectors find it challenging to navigate in the field, while Panel B shows the extent to which collectors struggle to locate intended taxpayers. The grey bar measures the difference in outcome between the treatment and control groups; the number in parentheses is the randomization inference-based p-value on the statistical significance of the difference. For a detailed description of the challenge measures, see Data Appendix B.5. The analysis is based on the collector surveys, described in Section 4.1.

Figure 6: Collector Knowledge of and Focus on Households that are Able to Pay

(a) Knowledgeable about Which Households are Able to Pay?



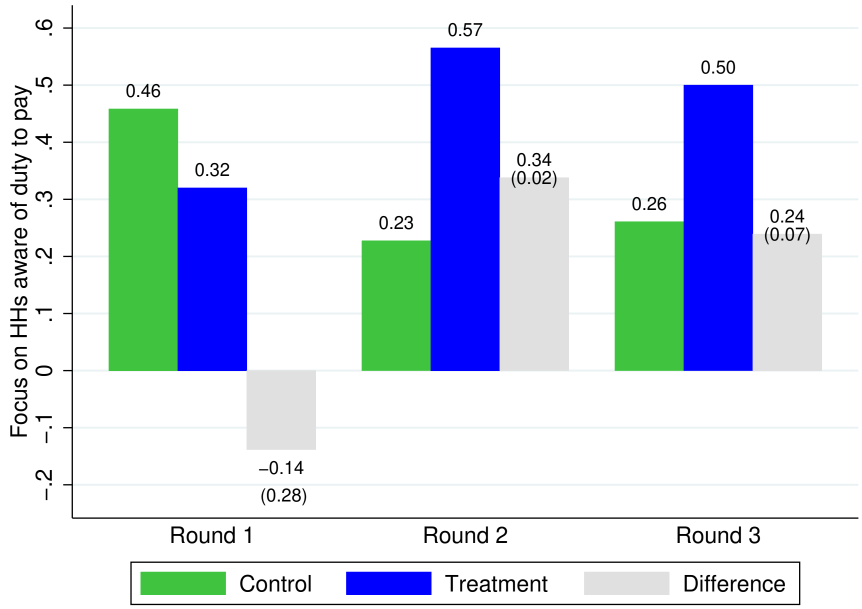
(b) Focus on Households that are Able to Pay?



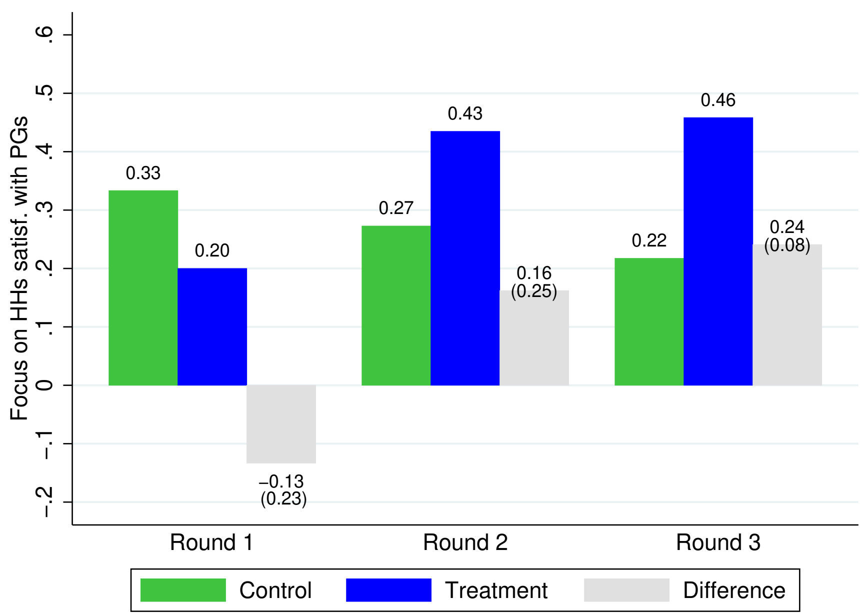
Notes: Panel A shows the extent to which collectors know where the households are located that are more able to pay taxes. Panel B shows the extent to which collectors rely on a collection strategy which targets these households that are more able to pay. The grey bar measures the difference in outcome between the treatment and control groups; the number in parentheses is the randomization inference-based p-value on the statistical significance of the difference. For a detailed description of the knowledge and strategy measures, see Data Appendix B.5. The analysis is based on the collector surveys, described in Section 4.1.

Figure 7: Collector Focus on Those Aware of Tax Duties and Satisfied with Public Goods

(a) Focus on Households that are Aware of Tax Payment Duty?



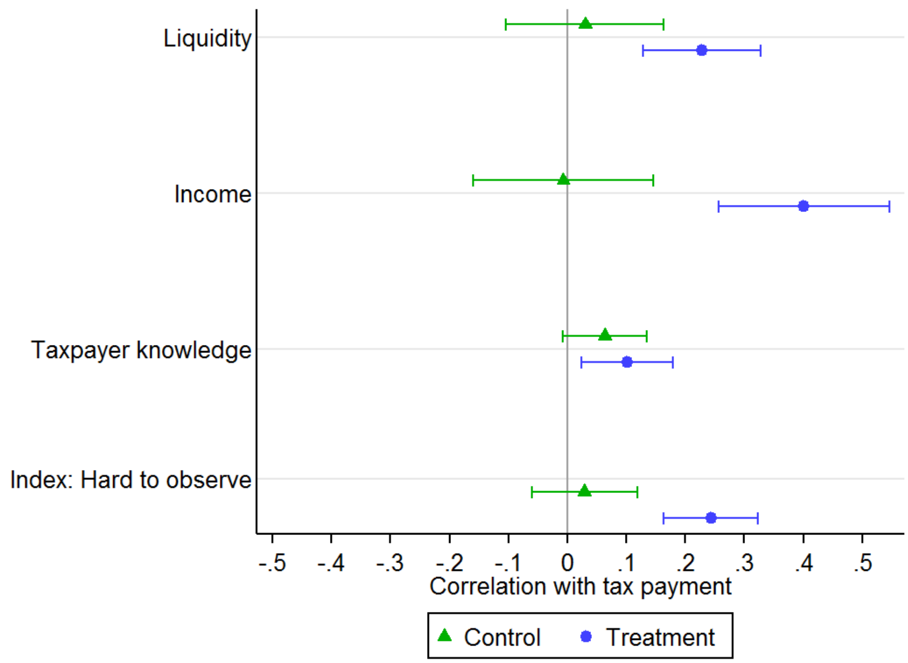
(b) Focus on Households that are Satisfied with Public Goods?



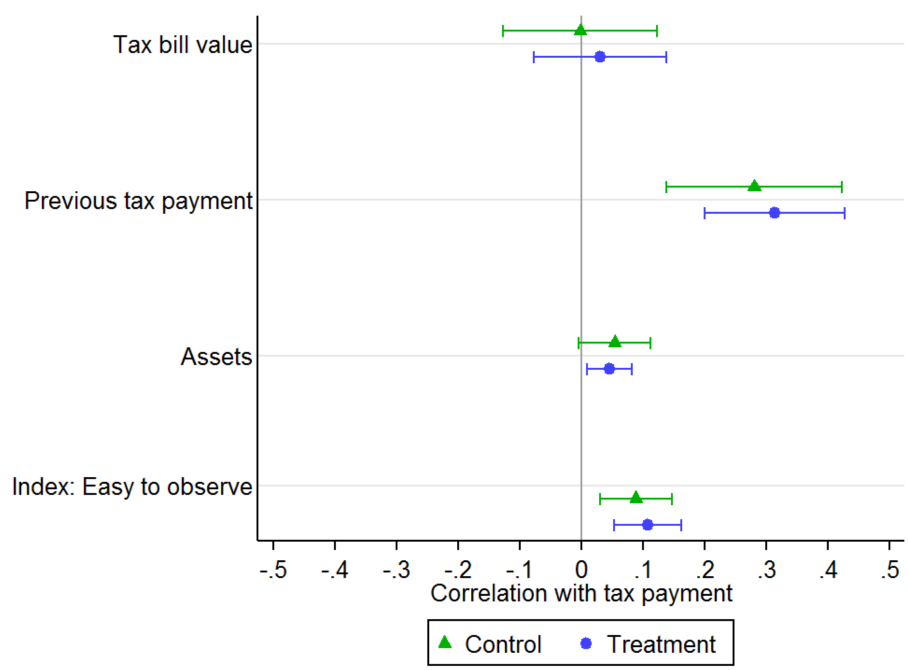
Notes: These panels show the extent to which collectors rely on different collection strategies – targeting households that are aware of tax-paying duties (Panel A) and households that are satisfied with local public service delivery (Panel B). The grey bar measures the difference in reliance on these strategies between the treatment and control groups; the number in parentheses is the randomization inference-based p-value on the statistical significance of the difference. For a detailed description of the strategies, see Data Appendix B.5. The analysis is based on the collector surveys, described in Section 4.1.

Figure 8: Characteristics of Households that Made a Tax Payment by Treatment Status

(a) Hard-to-Observe Characteristics

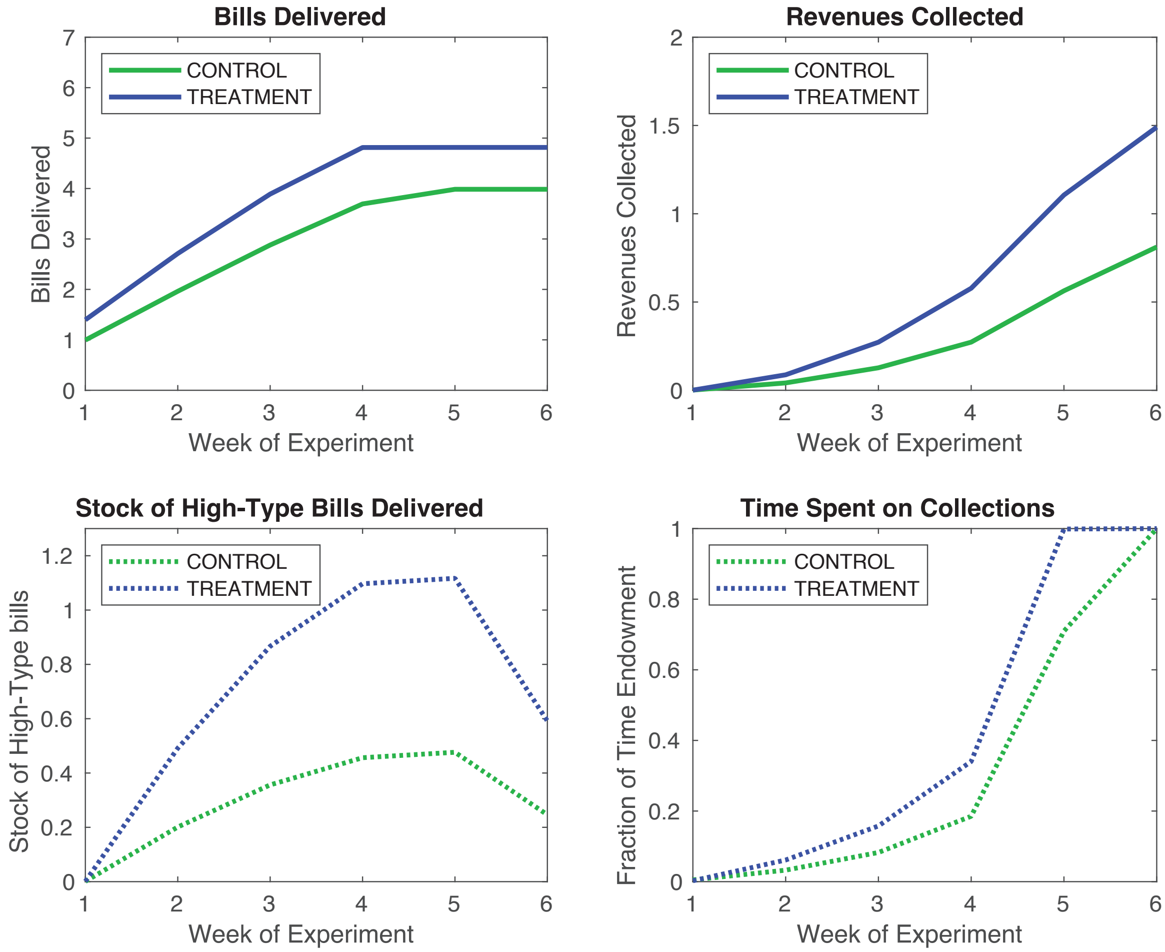


(b) Easy-to-Observe Characteristics



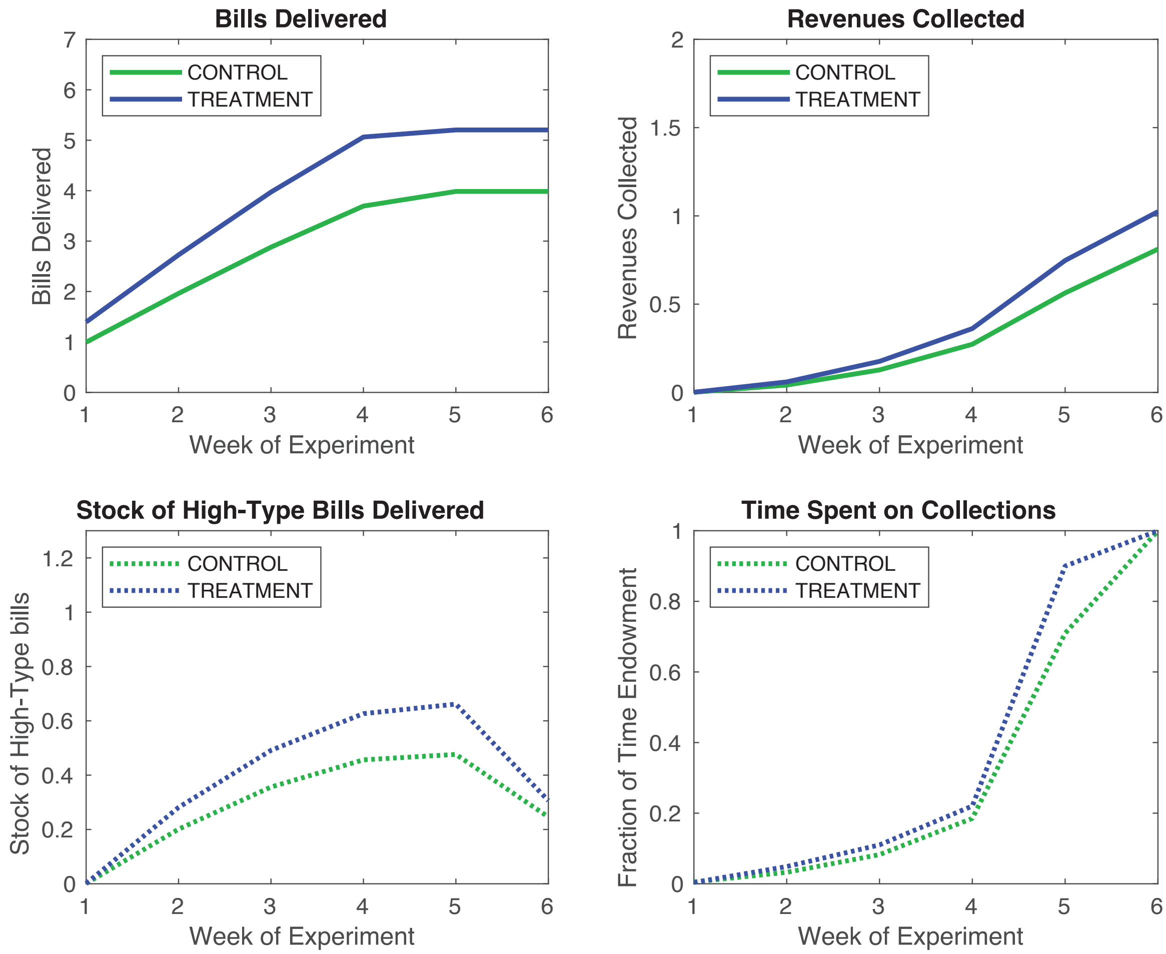
Notes: These figures show targeting of property owners for tax payment, based on estimating equation (3). The characteristics in Panel A are harder to observe, while the characteristics in Panel B are easier to observe. See Data Appendix B.4 for details.

Figure 9: Predictions of Benchmark Model



Notes: This figure shows the predictions of the benchmark model for the control and treatment groups. The upper panels plot the model's predictions for bills delivered and revenues collected over the course of the experiment. The bottom panels plot the stock of high-type bills delivered and the fraction of time spent on collections, rather than bill delivery.

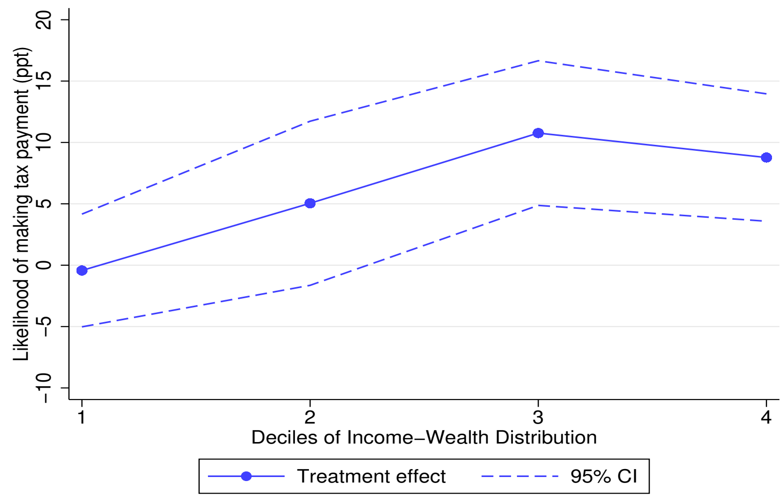
Figure 10: Predictions of Model with No Learning Advantage from Technology



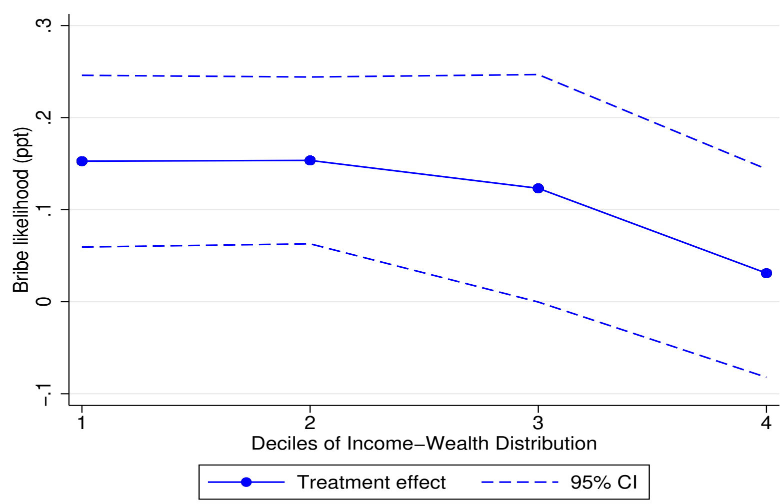
Notes: This figure shows the predictions of the model when the learning probability is set to be the same in the treatment and control groups. The upper panels plot the model's predictions for bills delivered and revenues collected over the course of the experiment. The bottom panels plot the stock of high-type bills delivered and the fraction of time spent on collections, rather than bill delivery.

Figure 11: Distributional Effects of Technology on Taxes and Bribes

(a) Treatment Effect on Tax Payment



(b) Treatment Effect on Bribe



Notes: These panels show the impact of technology on the likelihood of making any positive tax payment (Panel A) and on the likelihood of bribes (Panel B). Both panels display the treatment effect coefficient on technology, separately by quartile of the income-wealth distribution, based on estimating equation (5). The income-wealth distribution is calculated as the unweighted average, by household, of the income index and the assets index. For more detail on the index measures and the bribe measure, see Data Appendix B.3-B.4.

Table 1: Characteristics of Local Tax Capacity in Ghana

	Mean	p10	p50	p90
Panel A: Tax outcomes				
Taxes collected per capita (GHC)	4.2	0.8	2.6	6.9
Share of bills delivered (%)	43.0	2.7	43.3	80.0
Taxes collected per bill delivered (GHC)	11.45	2.09	6.72	21.50
Share of bills that are paid (%)	30.2	0.00	29.3	62.3
Panel B: Information and technology				
Share of properties with address (%)	26.7	0.0	20.0	70.0
Common not to locate property	0.37	0	0	1
Common not to locate owner	0.74	0	1	1
Technology: database and software	0.12	0	0	1
Technology: database or software	0.17	0	0	1
Panel C: Other capacity dimensions				
Share of properties with valuation (%)	17.1	0	0	55.0
Share of tax payments made in cash (%)	72.1	41.9	76.6	96.5
Cost of collection (% taxes collected)	64.1	15	47.5	91.1
Officials with post-secondary education (%)	0.67	0.53	0.66	0.82
Officials' avg. years work experience	11.7	7.8	11.8	15.1
1(Take tax defaulters to court)	0.22	0	0	0.66
1(Main reason for no court: Legal)	0.42	0	0.5	1
1(Main reason for no court: Political)	0.34	0	0.33	1
1(Heard about local tax code)	0.07	0	0.06	0.20
1(Agree that tax compliance is conditional)	0.70	0.40	0.73	1
Trust officials relative to stranger [0.25,4]	1.3	0.6	1	2
Number of local governments	216	216	216	216

All variables are calculated at the district level (N=216), using unweighted averages. In Panel A, the variables measure tax outcomes. In Panel B, the variables relate to information constraints and technology usage. In Panel C, the variables relate to additional constraints on tax capacity. For a detailed description of all variables, see Data Appendix B.1.

Table 2: Associations with Technology at the Local Government Level

	(1)	(2)	(3)	(4)	(5)
Panel A: Taxes per Capita					
1(Technology)	6.32***	3.71***	4.06***	3.08**	3.24***
	(1.62)	(0.88)	(0.60)	(1.16)	(0.94)
Mean outcome variable	4.15	4.15	4.15	4.15	4.15
Panel B: Share of Bills Delivered (%)					
1(Technology)	0.26***	0.08**	0.18***	0.14***	0.09**
	(0.06)	(0.03)	(0.04)	(0.03)	(0.03)
Mean outcome variable	0.43	0.43	0.43	0.43	0.43
Panel C: Taxes per Bill Delivered					
1(Technology)	6.9*	5.1**	4.8**	3.1***	4.2***
	(3.6)	(2.0)	(1.8)	(0.9)	(1.1)
Mean outcome variable	11.5	11.5	11.5	11.5	11.5
District controls		x			x
Share neighbors with tech			x		x
Region FE				x	x
Observations	216	216	216	216	216
Clusters	10	10	10	10	10

This regression table shows the correlation between technology adoption and tax-related outcomes. The model is a cross-sectional regression of all 216 MMDAs in Ghana. The variable **1**(Technology) is a dummy variable taking a value of 1 if the MMDA has: an electronic database of properties; and, a revenue management software, that assists with bill printing, payment recording, and follow-up enforcement. The results are similar if instead we code a district to have technology if it has either of the two components. Between panels, the outcome is: local taxes collected per capita (Panel A); the share of bills that are delivered to property owners (Panel B); taxes collected per bill delivered (Panel C). The outcome in Panel C is the ratio of the outcomes in Panel A and Panel B. Across columns, the specification varies: no controls in column 1; log per capita income, log population, urban share of population, share of properties with valuations, share of properties on official addresses are included in column 2; the share of each district's geographically adjacent neighbors is included in column 3; region fixed effects are included in column 4; all three sets of controls are included in column 5. Standard errors are clustered at the region level. See Data Appendix B.1 for more detail on the variables.

Table 3: Randomization Balance

	N (1)	Control mean (2)	Treatment (3)
Panel A: Tax bill characteristics			
Current tax amount	7843	322.8	-9.0
			(16.4)
Total tax amount	7843	692.5	-5.5
			(29.1)
Previous pay status	7843	1.2	0.0
			(0.0)
Previous tax payment	7843	59.7	-6.6
			(9.4)
Residential	7843	0.5	0.0
			(0.0)
Property quality	7843	0.5	0.0
			(0.1)
F-test joint significance [F, p]			[0.7,0.66]
Panel B: Collector characteristics			
Experience in Madina	56	0.7	-0.1
			(0.1)
			[0.74]
Performance rating	56	0.2	-0.1
			(0.1)
			[0.51]
Total bills to deliver	56	135.2	1.7
			(4.7)
			[0.73]
Average amount per tax bill	56	322.6	-7.4
			(16.5)
			[0.64]
F-test joint significance [F, p]			[0.2,0.95]

The construction of the variables used in this table are provided in: Section 4.1 (Panel A). The Treatment coefficient in column (3) is the coefficient on technology in a cross-sectional regression with strata fixed effects. Standard errors are clustered at the collection unit level. At the bottom of each panel, the F-test on the joint significance of all characteristics is reported, along with the p-value. In Panel B, the bracketed values indicate the p-value calculated with randomization inference.

Table 4: Impacts of Technology on Visits, Compliance and Revenues

	Any visit by tax collector (1)	Total visits (in %) (2)	Bill delivered by tax collector (3)	Any positive tax payment (4)	Total payment amount (in GHC) (5)	Payment amount per bill delivered (6)
Technology	0.087**	0.094*	0.054	0.043**	25.9**	47.3**
	(0.033)	(0.050)	(0.036)	(0.021)	(10.9)	(19.6)
Household controls	X	X	X	X	X	X
Collector controls	X	X	X	X	X	X
Strata FE	X	X	X	X	X	X
Mean in CG	0.55	0.67	0.51	0.16	41.0	80.9
Observations	4334	4334	4334	4334	4334	2276
Clusters	56	56	56	56	56	56

This table presents the impacts of technology on main outcomes of interest. All coefficients are based on estimating equation (2), and using the household sample (Section 4.1). Across columns, the outcome is: a dummy for any visit received by a tax collector; the total number of visits (expressed in %, using the inverse hyperbolic sine transformation; a dummy for any bill delivered; a dummy for any tax payment made; the total amount paid (in GHC); the payment made, restricted to households that received a bill. For a description of household controls and collector controls, please refer to Section 4.2. The robustness of these results to the removal of control variables, or the inclusion of more extensive controls, is presented in Appendix Table A4.

Table 5: Impacts of Technology on Citizen Beliefs and Tax Morale

	Satisfaction with government services (1)	Integrity of government (2)	Tax equity & efficiency efforts by government (3)	Enforcement information capacity of government (4)
Technology	-0.00771	0.0629	-0.0143	-0.0536
	(0.0701)	(0.0728)	(0.0604)	(0.0572)
Household controls	X	X	X	X
Collector controls	X	X	X	X
Strata FE	X	X	X	X
Mean in CG	0.05	-0.04	-0.03	0.00
Observations	4224	4226	4226	4226
Clusters	56	56	56	56

This table presents the impacts of technology on beliefs and tax morale. All coefficients are based on estimating equation (2), and using the household sample (Section 4.1). Each column is an index variable, which averages over multiple (standardized) household survey questions. For the description of each underlying question that is used in each index, please see Data Appendix B.2. Across columns, the index measures: the extent of satisfaction with government's delivery of services; the perceived integrity and competency of the local government; the local government's efforts to collect taxes in an equitable and efficient manner; the perceived enforcement capacity of the government and the informational knowledge that the government possesses about its citizens.

Table 6: Impacts of Technology on Collusive and Coercive Bribes

	Any bribe (coercive or collusive) (1)	Total bribe amount (in %) (2)	Collusive bribe amount (% of tax due) (3)	Coercive bribe amount (% of payment) (4)	Collusive bribe amount (in GHC) (5)
Technology	0.12***	0.03**	0.01**	0.04*	6.16**
	(0.04)	(0.01)	(0.00)	(0.02)	(3.07)
Household controls	X	X	X	X	X
Collector controls	X	X	X	X	X
Strata FE	X	X	X	X	X
Mean in CG	0.14	0.11	0.02	0.19	11.6
Observations	4334	4334	4334	4331	4334
Clusters	56	56	56	56	56

This table presents the impacts of technology on measures of payment capture by collectors. All coefficients are based on estimating equation (2), and using the household sample (Section 4.1). The bribe variables are described in detail in Data Appendix B.3. For a description of the household controls and collector controls, please refer to Section 4.2. The robustness of these results to the removal of control variables, or the inclusion of more extensive controls, is presented in Appendix Table A4.

Table 7: Impact of Technology on Collector Strategies in the Field

	Focus on collections, hard-to-observe household characteristics		Focus on collections, easy-to-observe household characteristics		Difference in strategies: Hard versus easy to observe	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Average effect						
Technology	0.140**	0.667**	0.088*	0.445*	0.051	0.223***
	(0.067)	(0.252)	(0.048)	(0.225)	(0.040)	(0.079)
	[0.046]	[0.000]	[0.086]	[0.001]	[0.178]	[0.001]
Panel B: Dynamic effect						
Technology × Round 1	-0.082	0.428	0.008	0.355	-0.091	0.073
	(0.103)	(0.271)	(0.068)	(0.230)	(0.077)	(0.078)
Technology × Round 2	0.264**	0.793***	0.123	0.492**	0.140**	0.300***
	(0.125)	(0.253)	(0.091)	(0.216)	(0.058)	(0.074)
Technology × Round 3	0.253**	0.778***	0.138	0.485**	0.114*	0.292***
	(0.117)	(0.259)	(0.088)	(0.227)	(0.062)	(0.071)
Collector controls	X		X		X	
Survey round FE	X	X	X	X	X	X
Collector FE		X		X		X
Mean in CG	0.280	0.280	0.239	0.239	0.041	0.041
Observations	141	141	141	141	141	141

This table presents impacts on technology on collector strategies. In column (1)-(2), the outcome is the average reliance on strategies which target hard-to-observe characteristics. In column (3)-(4), the outcome is the average reliance on strategies which focus on more easily observable characteristics. In column (5)-(6), the outcome is the difference between the reliance on hard-to-observe versus easy-to-observe strategies. For a detailed description of these collector strategies, see Data Appendix B.5. All regressions use the panel of three survey rounds, and include survey round fixed effects. Odd columns include collector controls, while even columns include collector fixed effects. Panel A reports the average effect of technology, while Panel B reports the round-by-round treatment effect (based on interacting round fixed effects with the technology variable. Standard errors clustered at the collector-level are reported in parentheses. In Panel A, the randomization inference based p-value is reported in brackets.

Online Appendix

Technology and Local Tax Capacity: Evidence from Ghana

A Additional Figures and Tables

Table A1: Associations with Technology Adoption

	1(Technology exists)	
	(1)	(2)
Total population	0.114***	0.084**
	(0.025)	(0.027)
Income per capita	0.073**	0.009
	(0.027)	(0.021)
Urban share of population	0.096***	0.048**
	(0.024)	(0.020)
Share of properties with address	0.112*	0.088*
	(0.053)	(0.044)
Share of properties with valuation	0.174***	0.130***
	(0.029)	(0.026)
Legal capacity to enforce taxes	0.084*	0.053**
	(0.042)	(0.022)
Tax-delinquents taken to court	-.001	-.001
	(0.016)	(0.011)
Officials' years of work experience	0.058**	0.058*
	(0.021)	(0.030)
Officials' years of education	0.014	-0.007
	(0.025)	(0.016)
Trust in officials	0.004	-0.002
	(0.013)	(0.013)
Citizen tax-knowledge	-0.030	-0.026
	(0.015)	(0.026)
Citizen compliance-attitude	-0.006	0.013
	(0.016)	(0.010)
Region FE		X
Observations	216	216
Clusters	10	10

Each cell represents the β coefficient from a separate regression, based on the model

$$\mathbf{1}(Technology)_{dr} = \beta \cdot X_d + \mu_r + \epsilon_{dr}$$

where $1(Technology)_{dr}$ is a dummy equal to 1 if the local government in district d in region r uses technology for tax collection (see Section 2). X_d is the district-characteristic which varies between rows; across columns, region fixed effects, μ_r, are included. All explanatory variables are standardized, for ease of comparison. Standard errors are clustered at the regional level. For a description of the different district-characteristics, see Data Appendix B.1.

Table A2: Challenges Reported in the Field by Collectors

	Unable to locate properties owners (1)	Wrong information printed on bills (2)	Resistance from property to accept bill (3)	Supervisors do not monitor activities in the field (4)	Supervisors do not check mistakes made in the field (5)	Supervisors are unavailable for for support when needed (6)
Technology	-1.040***	-0.266*	-0.100	-0.154	-0.065	0.181
	(0.130)	(0.145)	(0.124)	(0.164)	(0.167)	(0.153)
	[0.00]	[0.07]	[0.43]	[0.35]	[0.70]	[0.26]
Collector controls	X	X	X	X	X	X
Survey round FE	X	X	X	X	X	X
Mean in CG	0.51	0.13	0.04	0.07	0.03	-0.09
Observations	136	135	135	139	139	140

This table presents impacts on technology on the extent of challenges encountered in the field. All regressions pool the collector survey responses across the survey rounds, and include survey round fixed effects. All regressions also include the collector-level controls described in Section 4.2. The outcomes measure the extent to which collectors agree that a particular challenge characterized their weekly work in the field. Robust standard errors are reported in parenthesis; the randomization inference-based p-value is reported in brackets. For a detailed description of the outcomes, see Data Appendix B.5

Table A3: Balance on Household Characteristics

	N (1)	Control mean (2)	Treatment (3)
Income index	4353	-0.014	0.003
			(0.106)
Liquidity index	4353	0.051	-0.177
			(0.119)
Taxpayer knowledge index	4353	0.011	-0.01
			(0.039)
Asset index	4353	0.012	-0.031
			(0.034)
F-test joint significance $[F, p]$			[1.07,0.38]

The construction of the index variables is detailed in Data Appendix B.4. The Treatment coefficient in column (3) is the coefficient on technology in a cross-sectional regression with strata fixed effects. Standard errors are clustered at the collection unit level. At the bottom of each panel, the F-test on the joint significance of all characteristics is reported, along with the p-value.

Table A4: Robustness Checks for Technology Impacts on Tax and Bribe outcomes

	Any visit by tax collector (1)	Total visits (%) (2)	Bill delivered (3)	Any positive tax payment (4)	Total payment amount (in GHC) (5)	Any bribe (coercive or collusive) (6)	Total bribe amount (in %) (7)	Coercive bribe amount (in %) (8)	Collusive bribe amount (in %) (9)
Panel A: No Controls									
Technology	0.082**	0.085*	0.049	0.039**	24.93**	0.120***	0.027**	0.012**	0.042*
	(0.034)	(0.04)	(0.036)	(0.017)	(10.89)	(0.038)	(0.012)	(0.005)	(0.023)
Panel B: Extensive Controls									
Technology	0.086**	0.087*	0.055	0.047**	27.21**	0.113***	0.025**	0.014**	0.036*
	(0.032)	(0.049)	(0.034)	(0.020)	(11.18)	(0.037)	(0.010)	(0.005)	(0.019)
Strata FE	X	X	X	X	X	X	X	X	X
Mean in CG	0.55	0.67	0.51	0.16	40.95	0.14	0.11	0.02	0.19
Observations	4353	4353	4353	4353	4353	4353	4353	4353	4350
Clusters	56	56	56	56	56	56	56	56	56

This table presents technology impacts on the same set of outcomes as in Table 4 and Table 6. The estimation model is the same, except that: in Panel A, all household and collector controls are removed; in Panel B, additional controls are added. The additional controls in Panel B are the set of 6 fixed characteristics used in the targeting analysis – see Section 6.3 and Figure A10. For a description of the bribe variables, see Data Appendix B.3.

Table A5: Beliefs about Enforcement and Tax Morale

	Technology coefficient ($\hat{\beta}$) (1)	Mean in CG (2)	N (3)
Panel A: Enforcement and Information			
Share of HHs that comply with taxes	0.80	60.32	4330
	(2.38)		
Likelihood non-complier will end up paying	-0.07	3.08	4330
	(0.07)		
Likelihood Gov't has info about my address	-0.13	3.24	4330
	(0.10)		
Likelihood Gov't has info about my tax status	-0.13	2.95	4330
	(0.13)		
Likelihood Gov't has info about my job	0.03	2.52	4330
	(0.09)		
Panel B: Gov't Efforts to Improve Tax Collection			
Agree efforts to collect taxes efficiently	0.01	3.58	4330
	(0.07)		
Agree efforts to ensure fair share paid	-0.18***	3.42	4330
	(0.07)		
Agree efforts to collect for useful purposes	0.08	3.04	4330
	(0.11)		
Panel C: Government Capacity and Competency			
% of taxes wastefully spent	-3.48	55.81	4330
	(4.64)		
Agree Gov't has capacity to improve roads	0.04	3.94	4330
	(0.11)		
Agree Gov't has capacity to improve water access	0.01	3.99	4430
	(0.11)		
Agree Gov't has capacity to improve waste management	0.03	3.98	4330
	(0.11)		
Overall Gov't competency rating	0.07	2.41	4330
	(0.07)		
Panel D: Satisfaction with Gov't Services			
Quality of tax collector services	-0.003	2.31	4330
	(0.05)		
Quality of tax authority services	-0.02	2.31	4330
	(0.05)		
Quality of overall Gov't services	-0.01	2.20	4330
	(0.05)		

This table presents technology impacts on beliefs and tax morale. Each row presents the technology treatment coefficient (in column 1) from estimating equation (2) on different outcomes (which are described to the left). Standard errors are clustered at the collection-unit. Column (2) presents the mean of the outcome variable in control areas, while column (3) shows the sample size. For a description of all the outcomes, see Data Appendix B.2. All regressions include household and collector controls (Section 4.2).

Table A6: Characteristics that Predict Tax Payment

	(1)	(2)	(3)	(4)	(5)
Panel A: Tax Payment Status					
Hard to observe index	0.056**				0.048**
	(0.022)				(0.023)
Income index		0.026**			
		(0.012)			
Taxpayer knowledge index			0.016		
			(0.020)		
Liquidity index				0.029**	
				(0.012)	
Assets index					0.031
					(0.021)
Panel B: Total Taxes Paid					
Hard to observe index	20.64***				18.96**
	(7.000)				(7.368)
Income index		14.73***			
		(4.774)			
Taxpayer knowledge index			1.980		
			(5.849)		
Liquidity index				5.360	
				(3.718)	
Assets index					6.006
					(6.487)
Outcome avg (Panel A)	1.255	1.255	1.255	1.255	1.255
Outcome avg (Panel B)	60.99	60.99	60.99	60.99	60.99
Control tax liability	X	X	X	X	X
Control block FEs	X	X	X	X	X
Observations	4353	4353	4353	4353	4353
Clusters	56	56	56	56	56

This table shows that the 'hard-to-observe' characteristics are robust predictors of tax payment outside of the experimental setting. In Panel A, the outcome is the tax payment status in the year prior to the experiment, which can take a value of 1 (=no payment), 2 (=partial payment), 3 (=full payment). In Panel B, the outcome is the total amount of taxes paid in GHC, prior to the experiment. Across columns (1) to (5), the outcome is regressed on the different indices that proxy for 'hard to observe' characteristics; in column (5), the index for assets is included as a control. For a description of the indices, please see Data Appendix B.4. All regressions include the level of property taxes due as a control. Moreover, all regressions include block fixed effects (approximately 7-8 properties per block). The standard errors are clustered at the collection-unit level.

Table A7: Impact of Technology on Collector Performance Measures

	# of unsuccessful visits per successful visit (1)	Total hours worked per week (2)	Average # of hours spent to deliver one bill (3)	Satisfaction & happiness on the job (4)
Technology	-1.222	-2.382	-0.798***	0.117
	(1.024)	(1.412)	(0.207)	(0.150)
	[0.23]	[0.11]	[0.00]	[0.42]
Collector controls	X	X	X	X
Survey round FE	X	X	X	X
Mean in CG	7.67	18.84	1.66	-0.07
Observations	141	141	111	139

This table presents impacts on technology on collector performance measures. All regressions pool the collector survey responses across the survey rounds, and include survey round fixed effects. All regressions also include the collector-level controls described in Section 4.2. Across columns, the outcome is: number of unsuccessful visits per property for every successful visit; hours worked per week; hours worked per bill delivered; and, satisfaction in job. Robust standard errors are reported in parenthesis; the randomization inference-based p-value is reported in brackets. For a detailed description of the outcomes, see Data Appendix B.5.

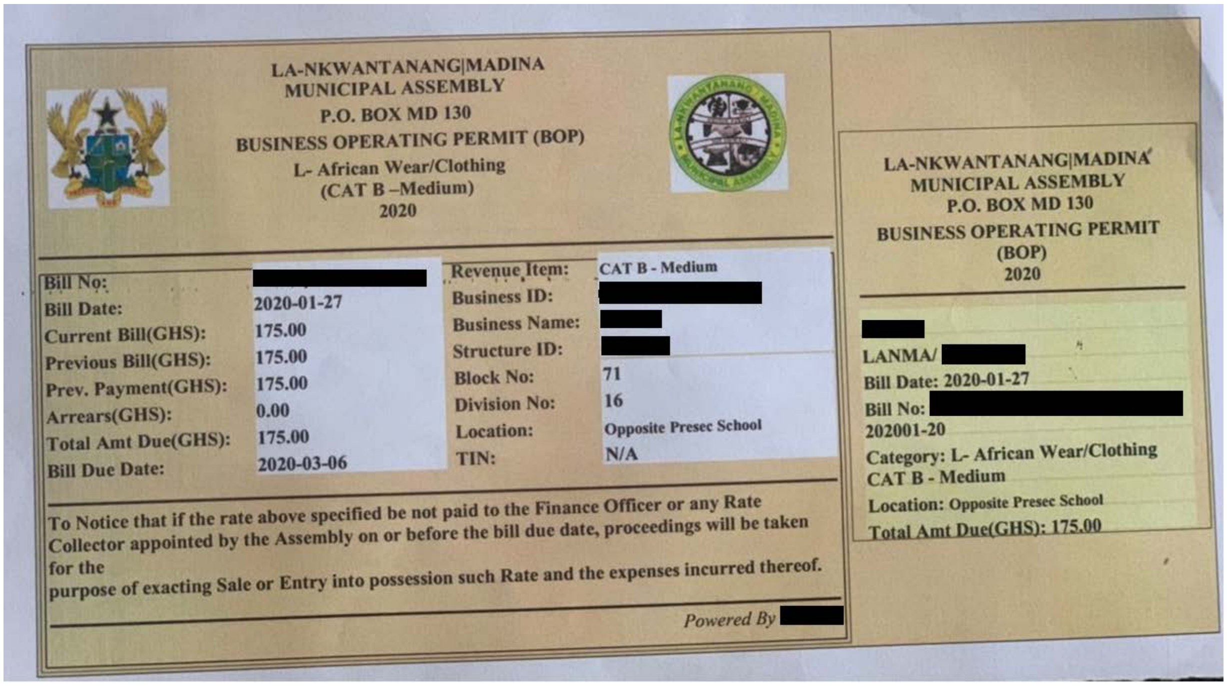
Table A8: Heterogeneity in Beliefs about Enforcement and Tax Morale

	Technology coefficient (β)	Heterogeneity coefficient ($\beta \times H$)
Outcome: Enforcement and Information Capacity Index		
Heterogeneity H: Liquidity index	-0.050	-0.016
	(0.056)	(0.053)
Heterogeneity H: Income index	-0.051	0.002
	(0.057)	(0.042)
Heterogeneity H: Taxpayer knowledge index	-0.050	-0.021
	(0.056)	(0.057)
F-test joint significance of interaction terms [F, p]	[0.09, 0.96]	
Outcome: Efforts to Improve Tax Collection Index		
Heterogeneity H: Liquidity index	-0.016	0.048
	(0.059)	(0.055)
Heterogeneity H: Income index	-0.010	0.059
	(0.060)	(0.039)
Heterogeneity H: Taxpayer knowledge index	-0.012	0.068
	(0.061)	(0.069)
F-test joint significance of interaction terms [F, p]	[1.27, 0.29]	
Outcome: Government Capacity and Competency Index		
Heterogeneity H: Liquidity index	0.048	0.039
	(0.070)	(0.063)
Heterogeneity H: Income index	0.063	0.012
	(0.073)	(0.040)
Heterogeneity H: Taxpayer knowledge index	0.064	-0.036
	(0.072)	(0.048)
F-test joint significance of interaction terms [F, p]	[0.32, 0.81]	
Outcome: Satisfaction with Gov't Services Index		
Heterogeneity H: Liquidity index	-0.018	-0.042
	(0.069)	(0.059)
Heterogeneity H: Income index	-0.009	0.011
	(0.069)	(0.032)
Heterogeneity H: Taxpayer knowledge index	-0.007	0.041
	(0.070)	(0.064)
F-test joint significance of interaction terms [F, p]	[0.45,0.72]	

This table presents heterogeneous technology impacts on beliefs and tax morale. Each row presents the technology treatment coefficient and the interaction coefficient, from estimating equation 2 augmented with the interaction between technology and the heterogeneity dimension H. Rows differ in the interaction (liquidity, income or taxpayer knowledge), and panels differ in the outcome. The F-test at the bottom of each panel tests the joint significance of the three interaction coefficients for a given outcome.

Figure A1: Illustrations of Tax Bill and Tablet

(a) Example of Business Property Tax Bill



LA-NKWANTANANG|MADINA
MUNICIPAL ASSEMBLY
P.O. BOX MD 130
BUSINESS OPERATING PERMIT (BOP)
L- African Wear/Clothing
(CAT B –Medium)
2020

Bill No:	[redacted]	Revenue Item:	CAT B - Medium
Bill Date:	2020-01-27	Business ID:	[redacted]
Current Bill(GHS):	175.00	Business Name:	[redacted]
Previous Bill(GHS):	175.00	Structure ID:	[redacted]
Prev. Payment(GHS):	175.00	Block No:	71
Arrears(GHS):	0.00	Division No:	16
Total Amt Due(GHS):	175.00	Location:	Opposite Presec School
Bill Due Date:	2020-03-06	TIN:	N/A

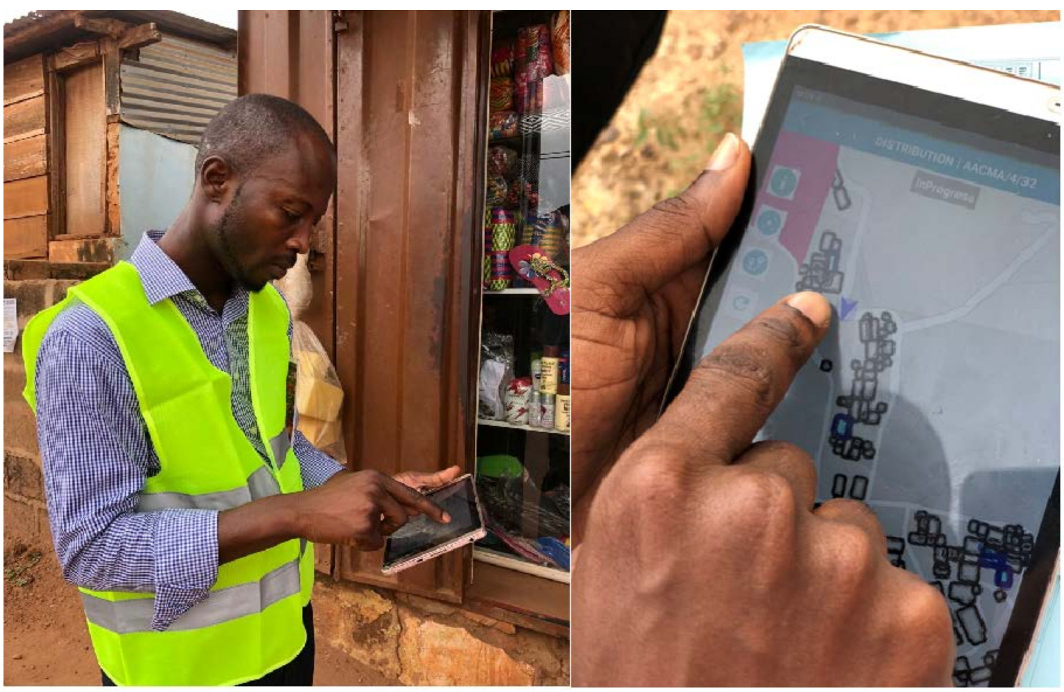
To Notice that if the rate above specified be not paid to the Finance Officer or any Rate Collector appointed by the Assembly on or before the bill due date, proceedings will be taken for the
purpose of exacting Sale or Entry into possession such Rate and the expenses incurred thereof.

Powered By [redacted]

LA-NKWANTANANG|MADINA
MUNICIPAL ASSEMBLY
P.O. BOX MD 130
BUSINESS OPERATING PERMIT
(BOP)
2020

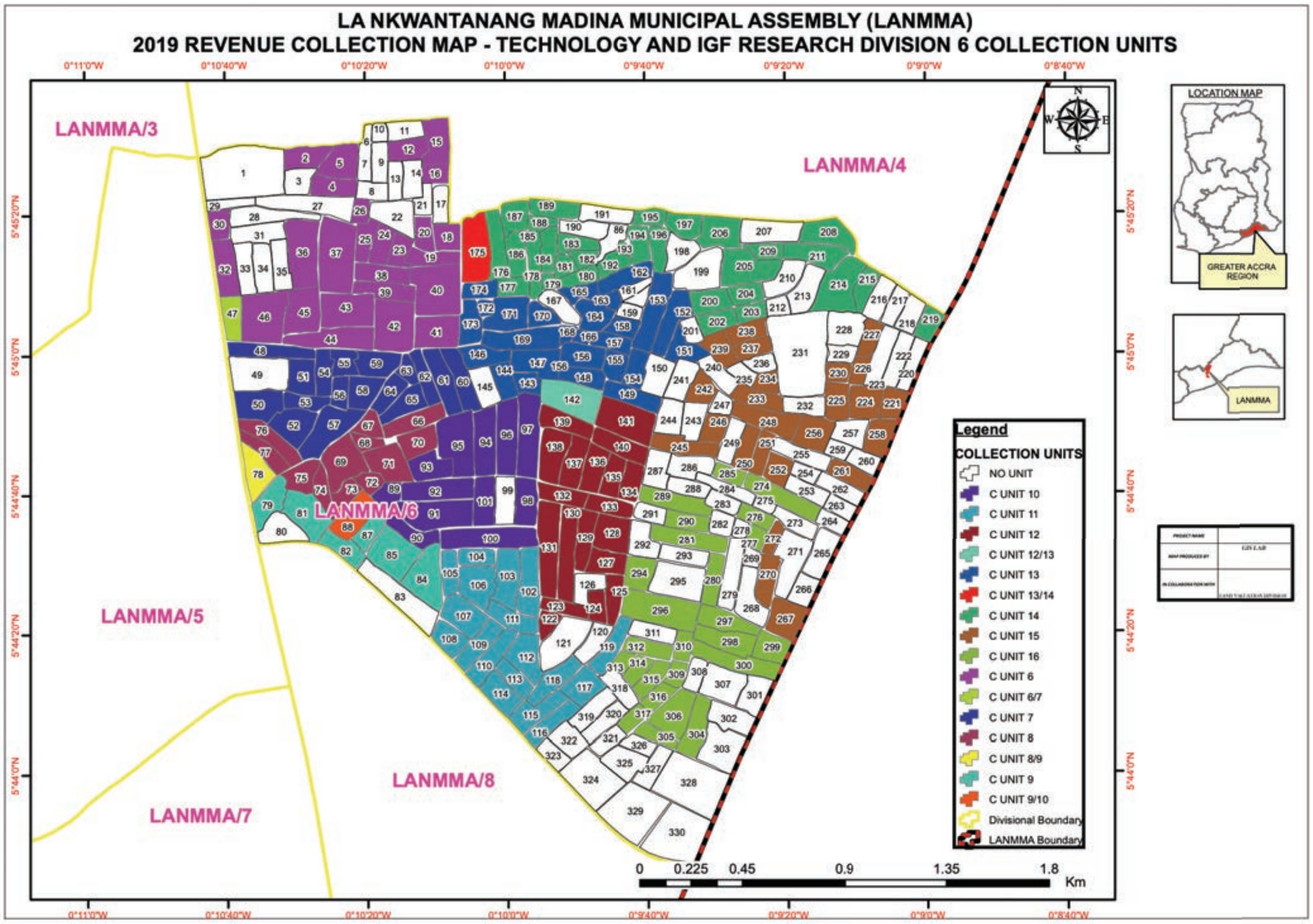
[redacted]
LANMA/ [redacted]
Bill Date: 2020-01-27
Bill No: [redacted]
202001-20
Category: L- African Wear/Clothing
CAT B - Medium
Location: Opposite Presec School
Total Amt Due(GHS): 175.00

(b) Navigational Assistance Provided by Tablet



Notes: These panels provide illustrations for a typical business property tax bill (Panel A), and the navigational assistance provided in the tablet that is used in treatment areas (Panel B).

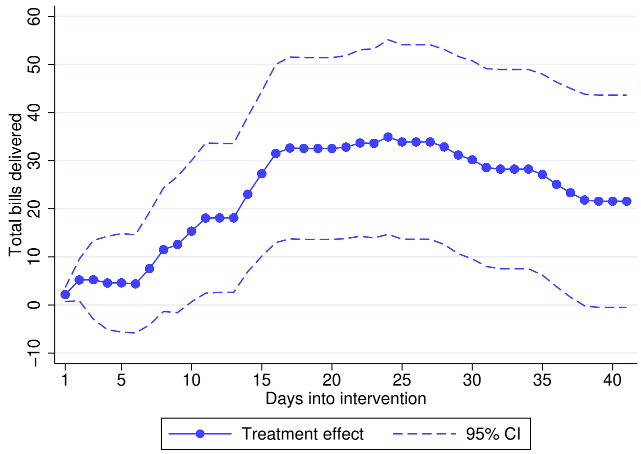
Figure A2: Illustration of Tax Collection Units



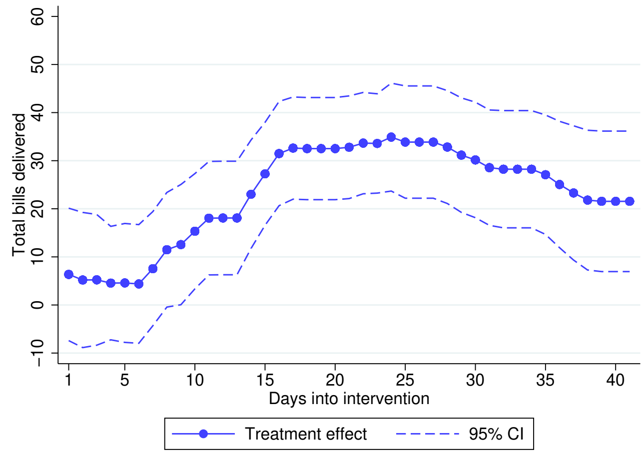
Notes: This graph provides an illustration of some of the collection units that exist in Madina. Due to confidentiality, these collection units are not necessarily included in the experimental sample.

Figure A3: Robustness for Impact on Bills Delivered

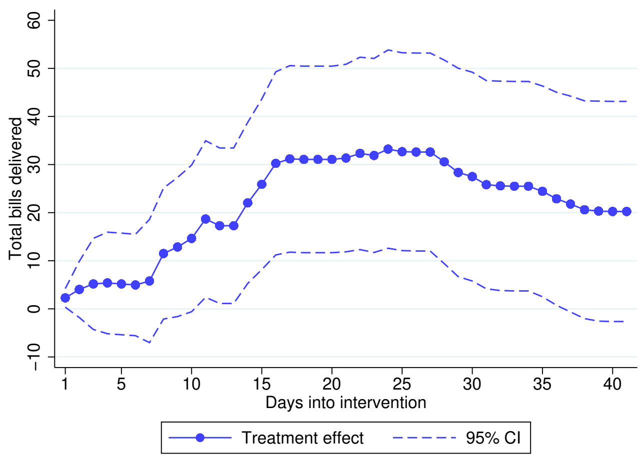
(a) Benchmark



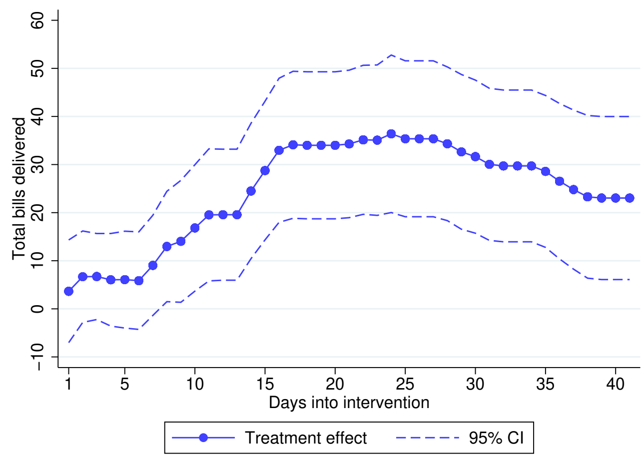
(b) Panel with Day and Collector FEs



(c) Winsorized Outcome



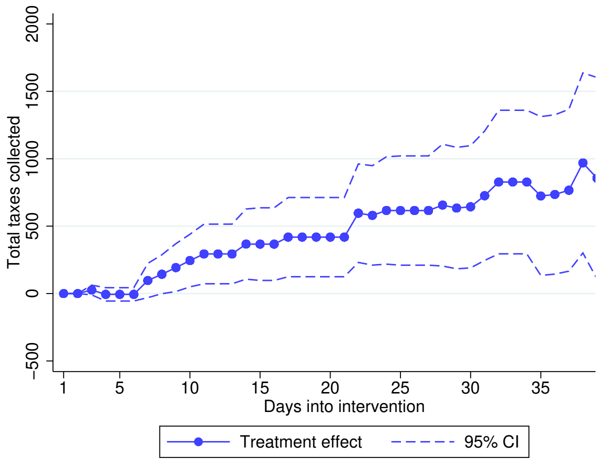
(d) Inclusion of Covariates



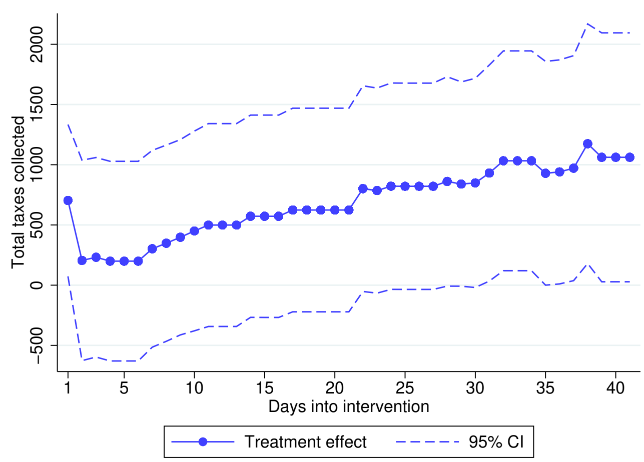
Notes: These panels show the robustness of the impact of technology on the number of property tax bills delivered. Panel A replicates the benchmark result from Figure 2, based on estimating equation (1). In Panel B, the benchmark estimation is augmented with collector fixed effects. In Panel C, the outcome is non-winsorized. In Panel D, controls are included: a dummy for whether the collector has previously worked in Madina; a dummy for whether the collector is assessed to be high performing; the total number of bills assigned to the collector; and, the average tax bill value per bill assigned. The analysis is based on the daily administrative data, described in Section 4.1.

Figure A4: Robustness for Impact on Taxes Collected

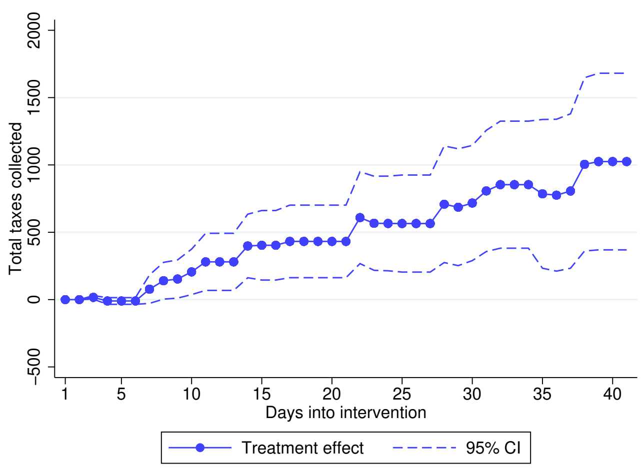
(a) Benchmark



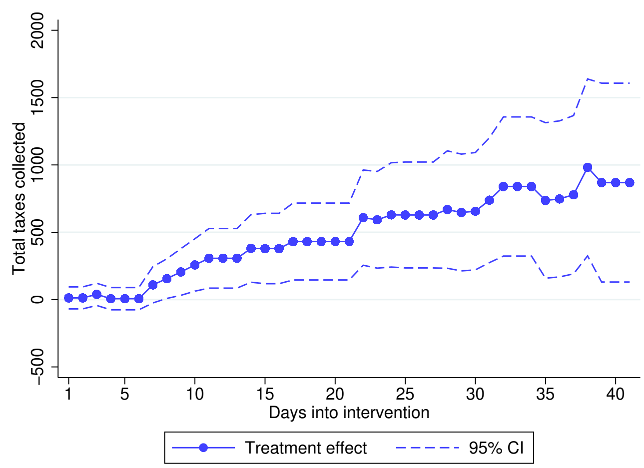
(b) Panel with Day and Collector FEs



(c) Winsorized Outcome



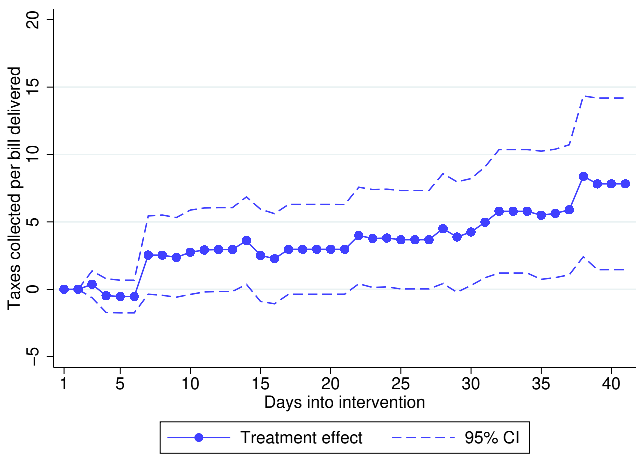
(d) Inclusion of Covariates



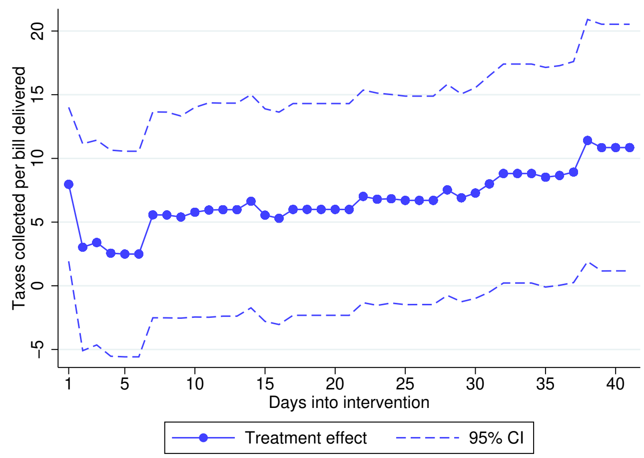
Notes: Notes: These panels show the robustness of the impact of technology on total taxes collected. Panel A replicates the benchmark result from Figure 3, based on estimating equation (1). In Panel B, the benchmark estimation is augmented with collector fixed effects. In Panel C, the outcome is non-winsorized. In Panel D, controls are included: a dummy for whether the collector has previously worked in Madina; a dummy for whether the collector is assessed to be high performing; the total number of bills assigned to the collector; and, the average tax bill value per bill assigned. The analysis is based on the daily administrative data, described in Section 4.1.

Figure A5: Robustness for Impact on Taxes Collected per Bill Delivered

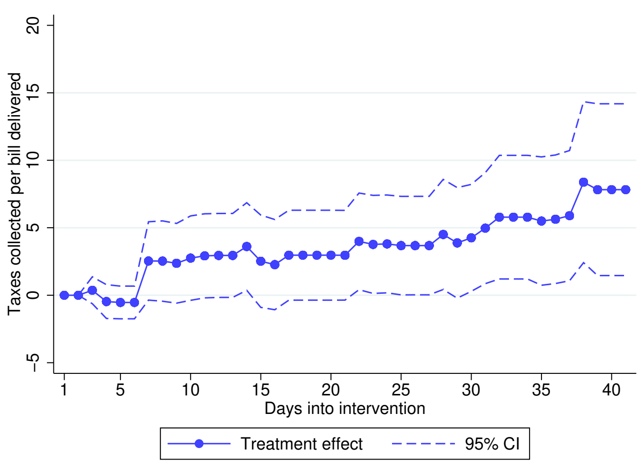
(a) Benchmark



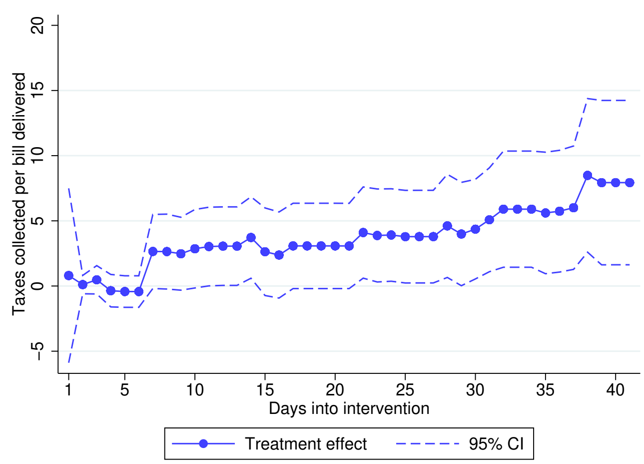
(b) Panel with Day and Collector FEs



(c) Winsorized Outcome



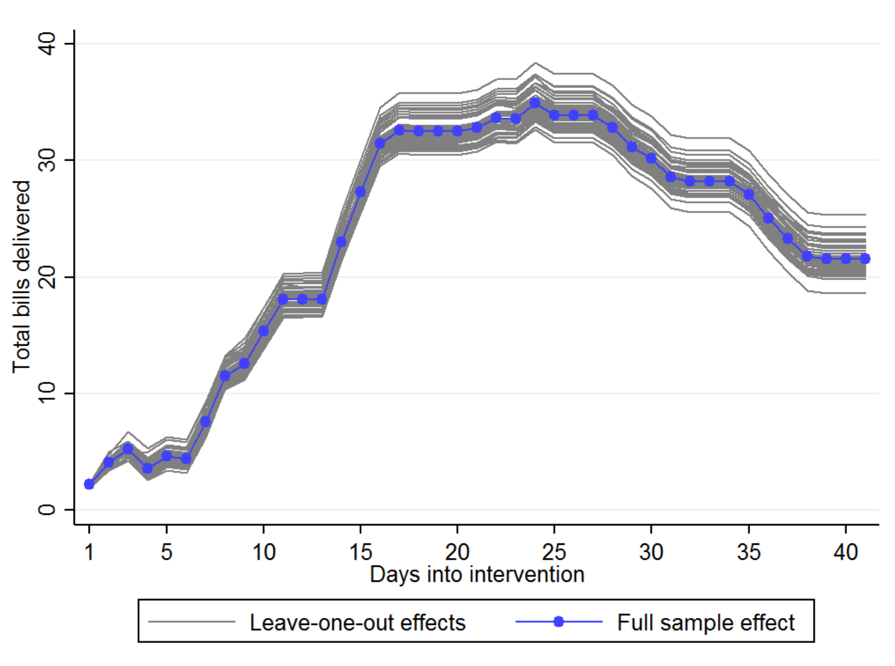
(d) Inclusion of Covariates



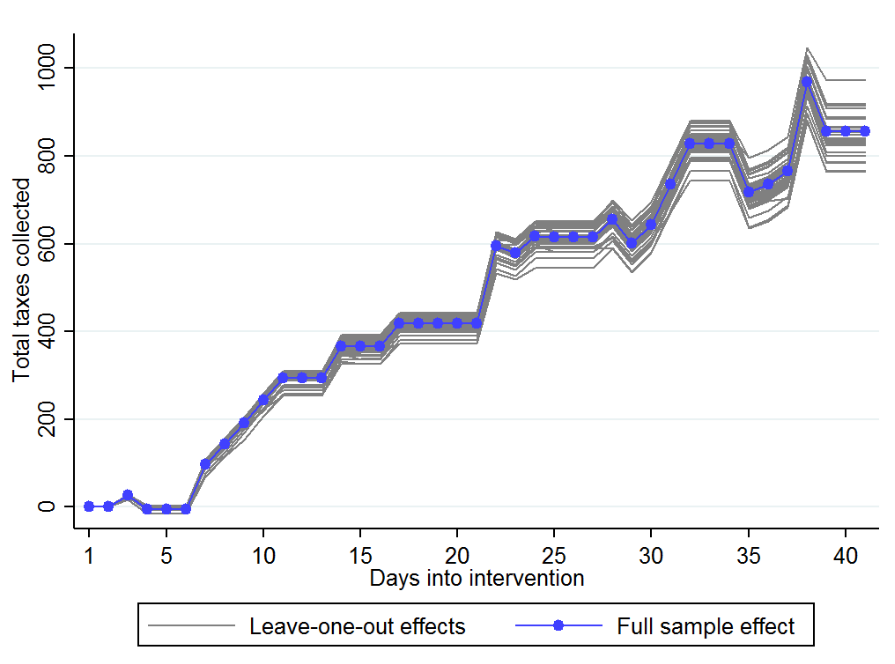
Notes: Notes: These panels show the robustness of the impact of technology on total taxes collected per bill delivered. Panel A replicates the benchmark result from Figure 4, based on estimating equation (1). In Panel B, the benchmark estimation is augmented with collector fixed effects. In Panel C, the outcome is non-winsorized. In Panel D, controls are included: a dummy for whether the collector has previously worked in Madina; a dummy for whether the collector is assessed to be high performing; the total number of bills assigned to the collector; and, the average tax bill value per bill assigned. The analysis is based on the daily administrative data, described in Section 4.1.

Figure A6: Robustness of Impacts to Leave-one-out Sample Restrictions

(a) Bills Delivered



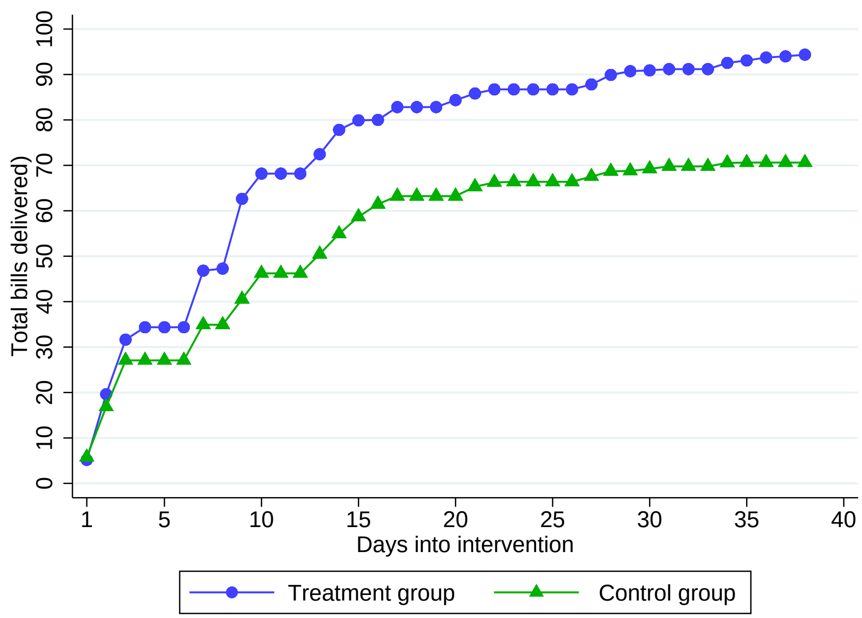
(b) Taxes Collected



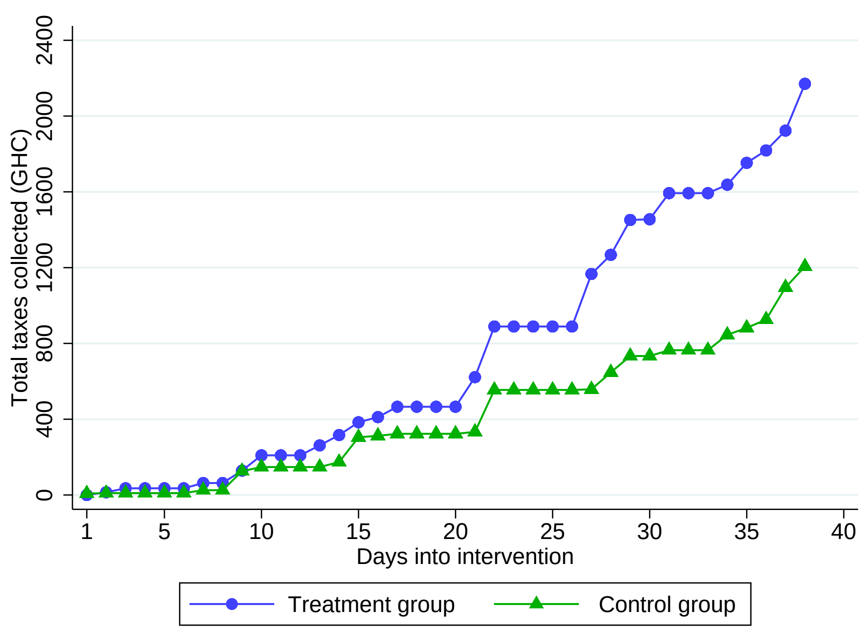
Notes: These panels show robustness of technology impacts on total bills delivered (Panel A) and total taxes collected (Panel B). In each panel, the blue dotted line represent the dynamic treatment effects estimated in the full sample (Panel B of Figure 2 and Figure 3, respectively). Each dark-gray line represents the dynamic treatment effects from estimating the same econometric model, but in individual sub-samples which remove one collector at a time.

Figure A7: Results from Pilot Experiment

(a) Average Number of Bills Delivered



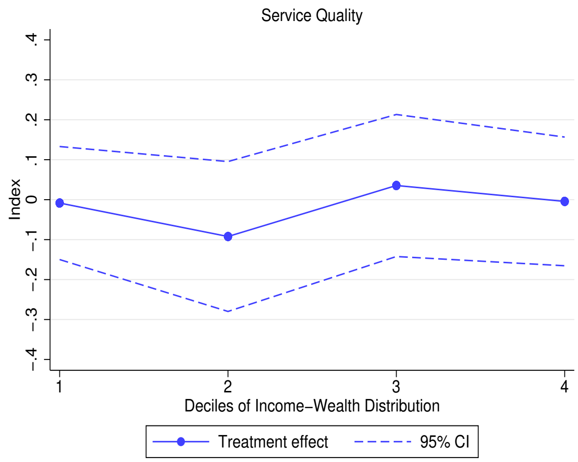
(b) Average Total Taxes Collected



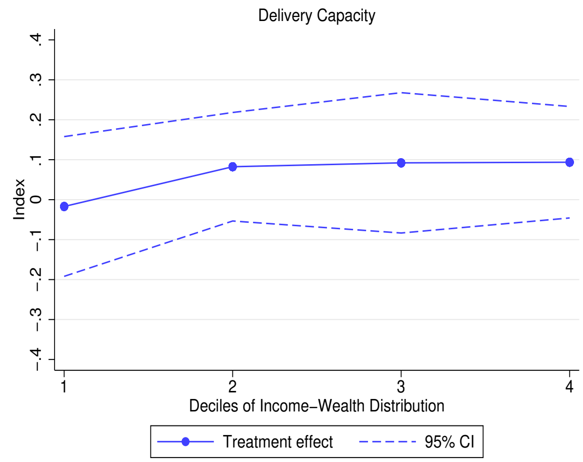
Notes: These panels show the impacts of technology on bills delivered and taxes collected based on the pilot experiment conducted in September-October 2019. The pilot was implemented in the same location as the main experiment, using the same technology, and following the same protocol for randomization and data-collection (see Section 3 for details). The pilot involved only 24 collectors and lasted 5 weeks, while the main experiment involves 56 collectors and lasts 6 weeks. Panel A (B) is constructed in the same way as Panel A of Figure 2 (Panel A of Figure 3). The treatment collectors had delivered 32% more bills at the end of the pilot experiment (compared to 27% at the end of the main experiment) and collected 74% more taxes (103%).

Figure A8: Distributional Effects on Beliefs about Government Capacity and Tax Morale

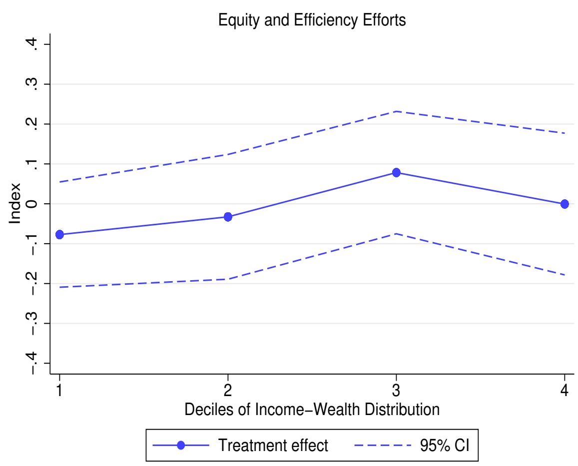
(a) Quality of Government Services



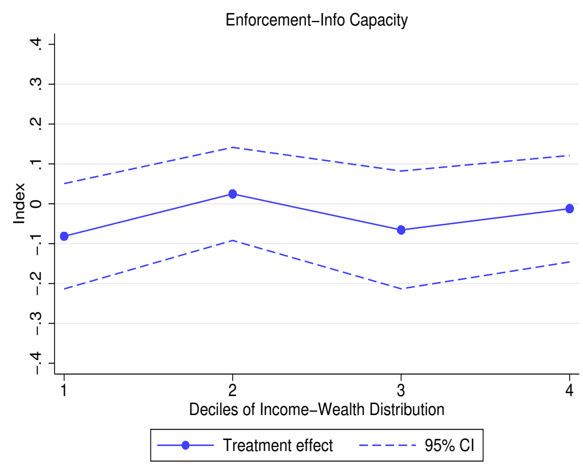
(b) Government Delivery Capacity



(c) Equity and Efficiency of Tax System



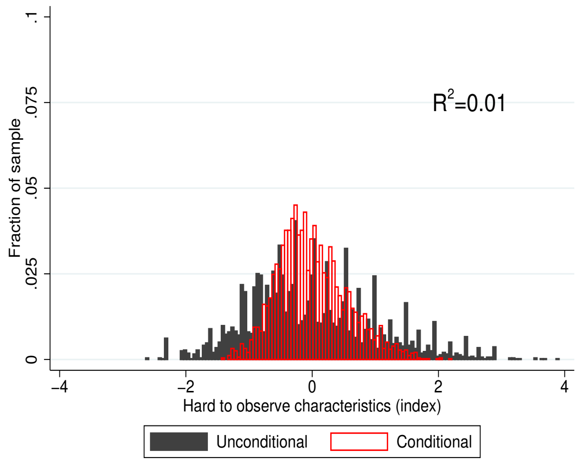
(d) Government Enforcement Capacity



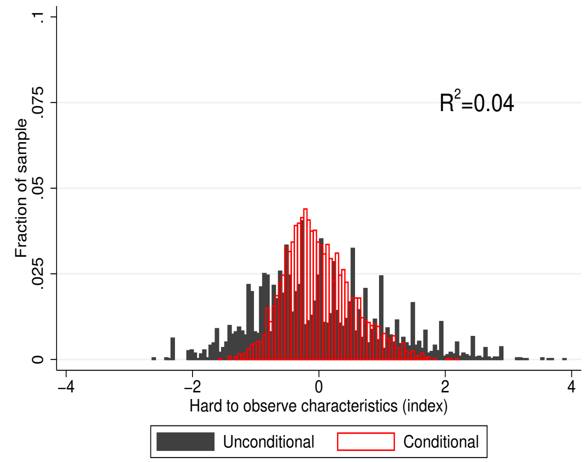
Notes: These panels investigate distributional impacts of technology on household beliefs about government capacity and tax morale. The four panels study four indices: satisfaction with the quality of government services (Panel A); capacity and integrity of local government (Panel B); government efforts to improve the efficiency and equity of the collection process (Panel C); the enforcement and information capacity of the local government (Panel D). Each panel displays the treatment effect coefficients on technology, separately by quartile of the income-wealth distribution, based on estimating equation (5). The income-wealth distribution is calculated as the unweighted average, by household, of the income index and the assets index. For a detailed description of the different indices, see Data Appendix B.2-B.4.

Figure A9: Correlation between Less-visible Index and More-visible Characteristics

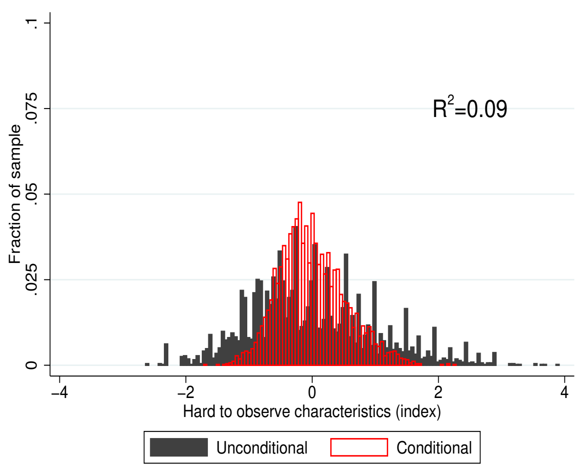
(a) Value of Tax Bill



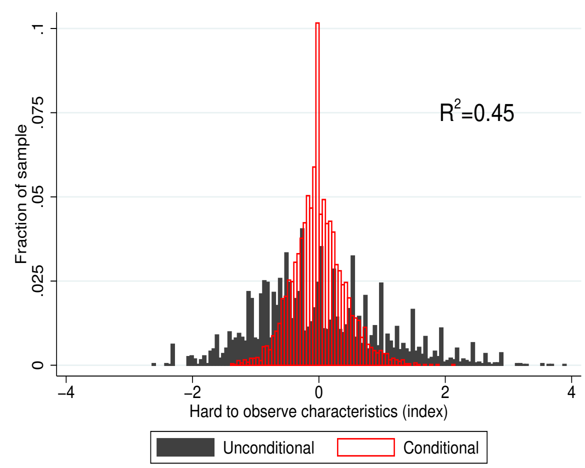
(b) Easily Observable Chars. (Index)



(c) Narrow Property Categories



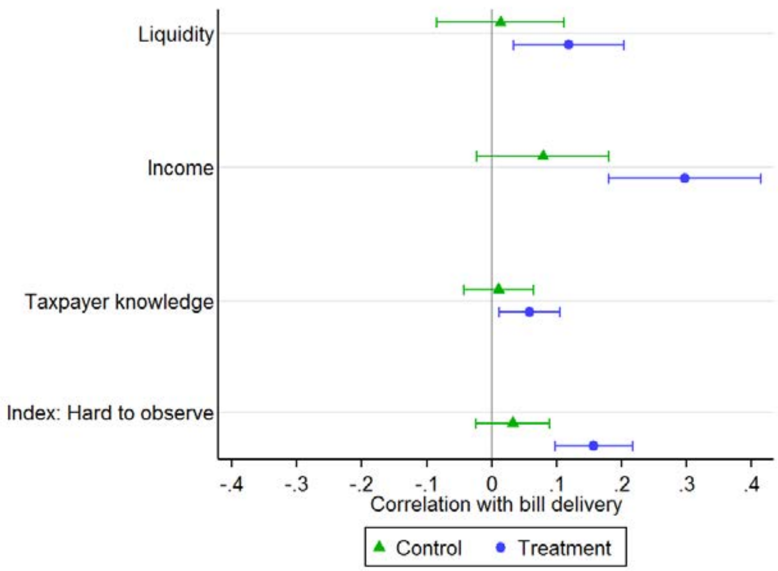
(d) Granular Location (Block Fixed Effects)



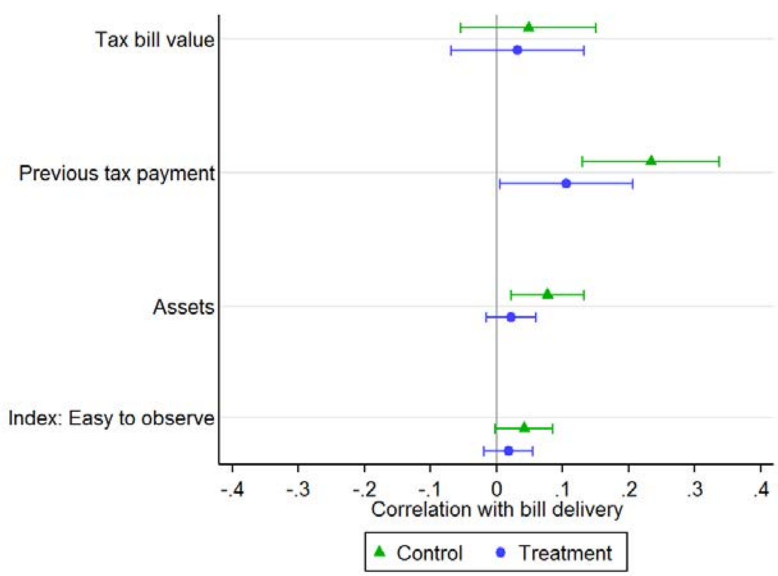
Notes: These panels show the distribution of the index for hard-to-observe characteristics, which measures the household's propensity to pay based on income, liquidity and taxpayer knowledge. In each panel, the grey-colored histogram shows the unconditional distribution of the index; the red-colored histogram shows the conditional distribution of the index, after controlling for more easily-observable characteristics. In the top-right corner is reported the R^2 of the regression of the unconditional index on the specific characteristics. Across panels, the included characteristic is: value of tax bill (Panel A); index for easily-observable characteristics (Panel B); categories of property quality; block fixed effects (Panel D). For detailed information on the indices, see see Data Appendix B.4. The block is a geographical cluster which contains 7 to 8 properties on average.

Figure A10: Characteristics of Households that Received a Bill by Treatment Status

(a) Less Visible Characteristics



(b) More Visible Characteristics

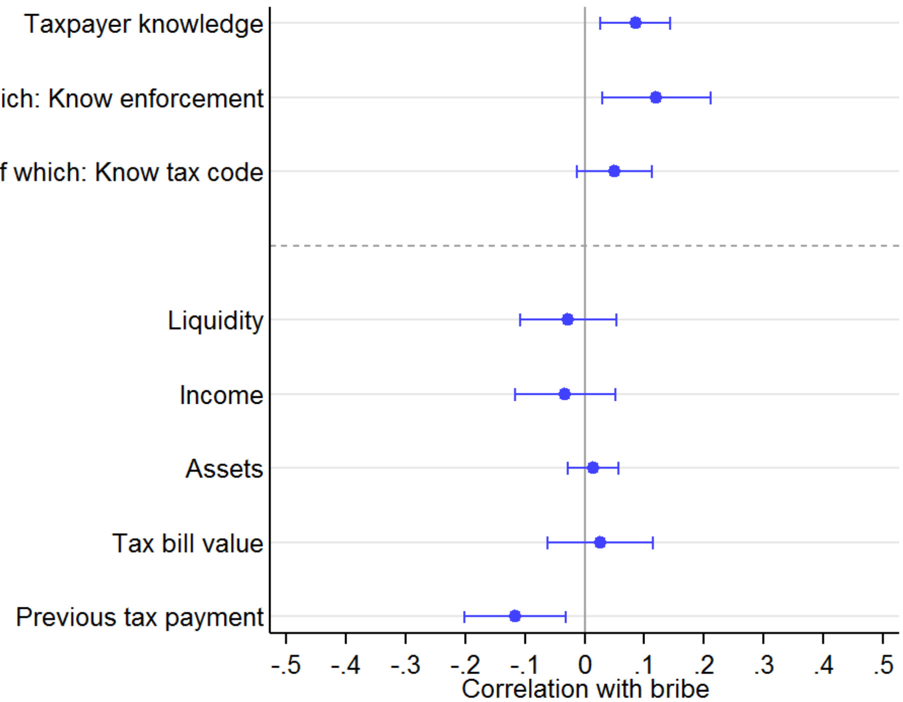


Notes: These panels show the selection on bill delivery payment for fixed household characteristics. The econometric model is the same as equation (3), except that the dummy for bill delivery is replaced with a dummy for any positive tax payment made. Formally, we estimate

$$y_{hc} = \theta \cdot \mathbf{1}(Taxpayment)_h + \beta \cdot [\mathbf{1}(Taxpayment)_h * \mathbf{1}(Tech)_c] + \Omega \cdot X_h + \mu_c + \epsilon_{hc}$$

For a detailed description of the household characteristics, see Data Appendix B.4.

Figure A11: Characteristics of Households Targeted for Bribes in the Treatment Group



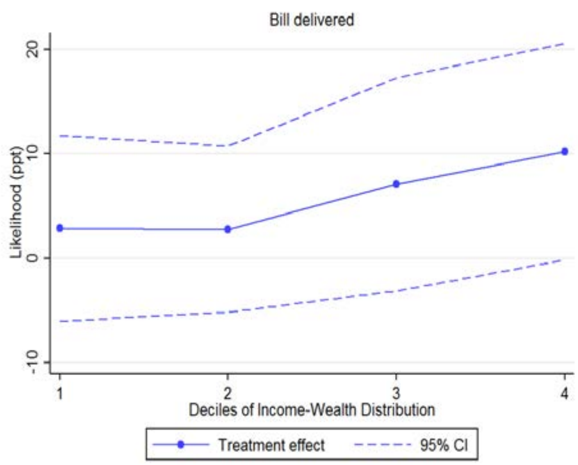
Notes: Notes: This panel shows the selection on bribe incidence for fixed household characteristics. The econometric model is the same as equation (3), except that the dummy for bill delivery is replaced with a dummy for any bribe incidence, and the analysis is limited to treatment areas. This dummy is the same as the outcome variable in Panel B of Figure 11. Formally, we estimate

$$y_{hc} = \theta \cdot \mathbf{1}(Bribe)_h + \Omega \cdot X_h + \mu_c + \epsilon_{hc}$$

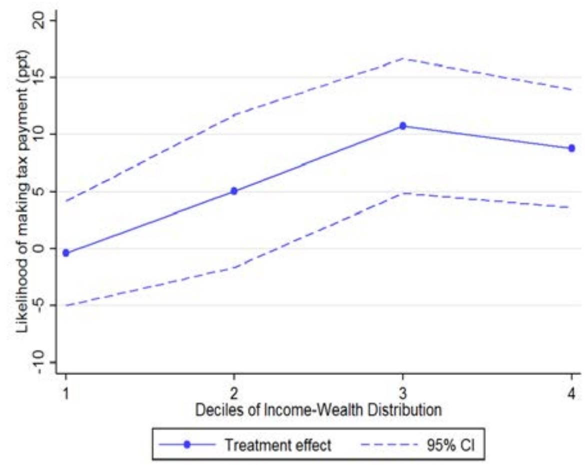
The fixed household characteristics are the same as those described in Figure A10. Moreover, in the top panel, the household characteristics measure knowledge about enforcement and knowledge about the tax code. For detailed description of the variables, see Data Appendix B.3-B.4.

Figure A12: Robustness of Distributional Impacts to Additional Tax Measures

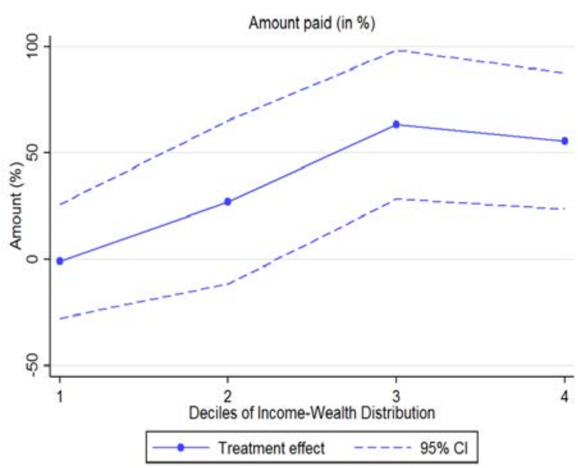
(a) **1**(Bill Delivered)



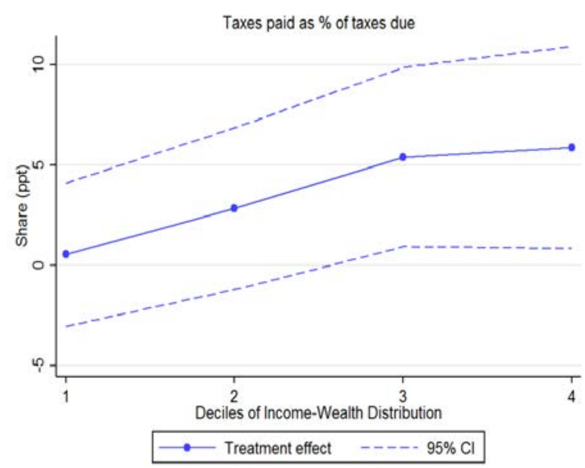
(b) **1**(Any Taxes Paid)



(c) Total Amount Paid (%)

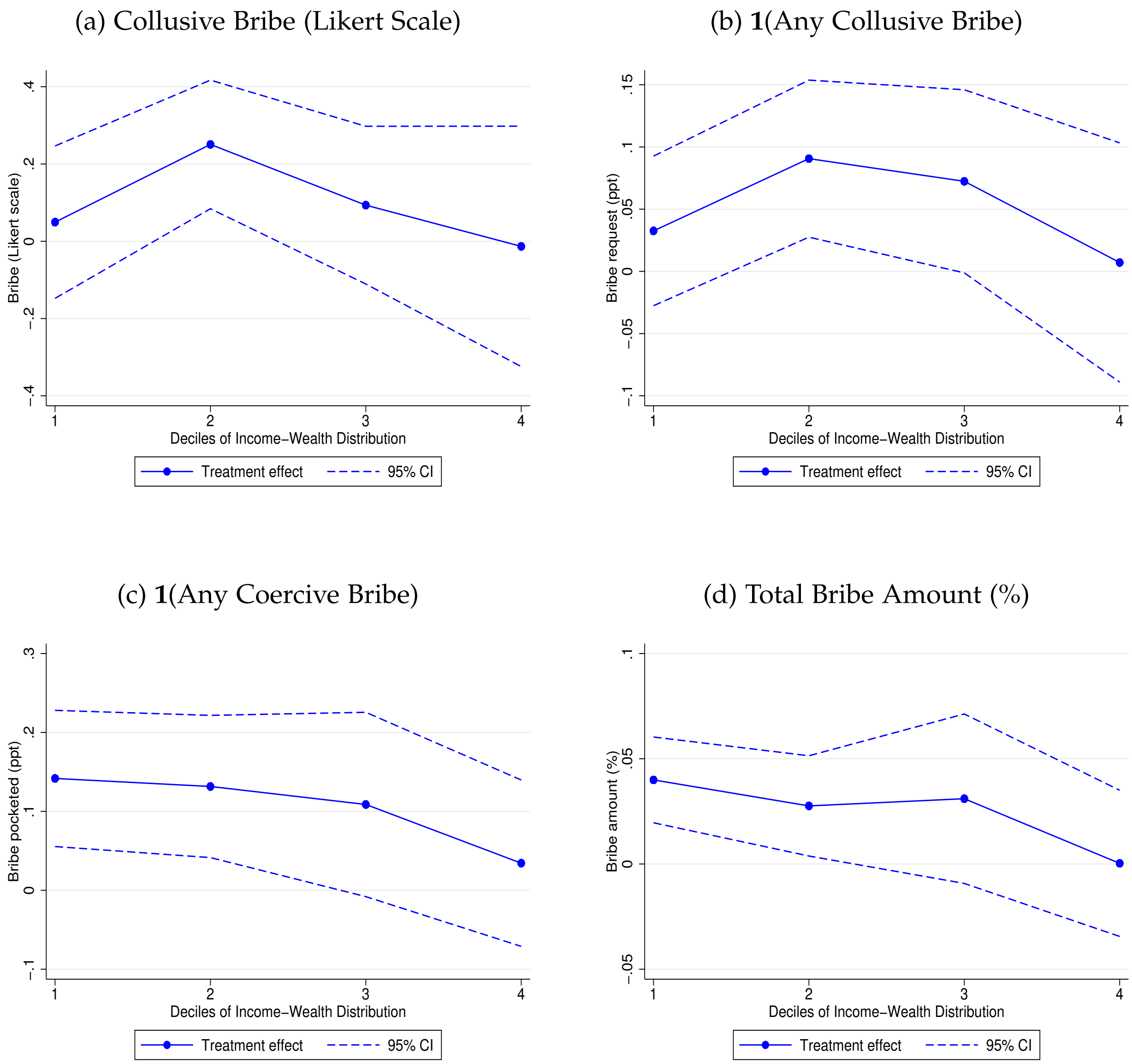


(d) Total Taxes Paid (% of Taxes Due)



Notes: These panels show the robustness of the distributional impact of technology. The econometric model is the same as in Figure 11, but the outcome varies across panels: a dummy for a bill delivered (Panel A); a dummy for any taxes paid (Panel B); amount of taxes paid, expressed as a % using the inverse hyperbolic sine (Panel C); and, the amount of taxes paid, expressed as a % of taxes due (Panel D).

Figure A13: Robustness of Distributional Impacts to Additional Bribe Measures



Notes: Notes: These panels show the robustness of the distributional impact of technology on payment capture by collectors. The econometric model is the same as in Figure 11, but the outcome varies across panels. For a detailed description of the outcomes, see Data Appendix B.3.

B Data Appendix

This section provides additional details on the variables considered in this paper.

B.1 Variables from Census of Local Governments

- *Share of bills delivered (%)* This variable is the answer to the question "Considering all the properties in your district, approximately what percent were sent a bill this year?" The answer ranges from 0% to 100%.
- *Taxes collected per bill delivered (GHC)* This variable divides the total taxes collected per capita (in Ghanaian Cedi) by the variable *share of bills delivered*.
- *Share of bills that are paid (%)* This variable is based on the answer to the question "Cumulatively, what share of bills are paid by the end of the year?". This answer is asked separately for business property taxes and for resident property taxes. We construct the district-level variable as the unweighted average over the responses for businesses and residents.
- *Share of properties with address (%)* This variable is the answer to the question "Approximately what percent of the properties in your assembly have an official address assigned to them?". The answer ranges from 0% to 100%.
- *Common to not locate property* This variable takes a value of 1 (0) if the respondent answers 'Yes' ('No') to the question "When delivering bills, it is common that you cannot locate the property/business for the bill to be delivered?"
- *Common to not locate owner* This variable takes a value of 1 (0) if the respondent answers 'Yes' ('No') to the question "When delivering bills, it is common that you locate the property/business but cannot locate the owner?"
- *Share of properties with valuation (%)* This variable is the answer to the question "Approximately what percent of the properties in the district are currently assessed by the Lands Valuation Board?". The answer ranges from 0% to 100%.
- *Share of tax payments made in cash (%)* This variable is the answer to the question "Approximately what percent of property rates are paid in cash?". The answer ranges from 0% to 100%.

- *Cost of collection (% of taxes collected)*. This variable is based on two questions asked to collectors in the census. The first question asks the collector what is their salary in a typical month. The second question asks what is the collector what is total revenue collected in a typical month. The variable is the ratio of salary to revenue collected, expressed as a a %.

- *Officials with post-secondary education*. This variable is a dummy variable equal to 1 (0) if the local official has completed any form of post-secondary education (has completed secondary education or less). In turn, we calculate the unweighted share of officials with post-secondary education in each district.

- *Officials' average years of work experience*. This variable is the answer to the question "For how many years and months have you worked in local government?". Note that this variable includes working in the local official's current district as well as other districts in the past. In turn, we calculate the unweighted average years of work experience in each district.

- *Legal capacity to enforce taxes*. This variable is a dummy variable equal which takes a value of 1 if the local assembly has gazetted the fee fixing resolution for the fiscal year 2017-2018, and zero otherwise.

- *Take tax defaulters to court*. This variable is a dummy variable equal to 1 (0) if the respondent answers 'Yes' ('No') to the question "Does the assembly normally take ratepayers/business owners to court for non-payment of property rates".

- *Main reason for no court: Legal*. This variable is a dummy variable equal to 1 if the respondent answers 'Legal constraints' in answer to the question "Why does your district not take more ratepayers/business property owners to court for non-payment?" and zero otherwise. The other possible answers are 'Not worth it'; 'Politically sensitive'; and, 'Yet to implement/prefer non-enforcement'.

- *Main reason for no court: Political*. This variable is a dummy variable equal to 1 if the respondent answers 'Politically sensitive' in answer to the question "Why does your district not take more ratepayers/business property owners to court for non-payment?" and zero otherwise. The other possible answers are 'Not worth it'; 'Legal constraints'; and, 'Yet to implement/prefer non-enforcement'.

- *Heard about local tax code*. This variable is a dummy variable equal to 1 if the respondent answers 'Yes' to the question "Have you heard about the fee fixing resolution?" and 0 if the respondent answers 'No'.

- *Agreement that tax compliance is conditional*. This variable is a dummy variable based on the answer to a question which has two statements: A ″Citizens should always pay their rates/fees″; B ″Citizens should only pay rates/fees if they believe it will lead to increases in local development″. The variable takes a value of 1 if the respondent ′strongly agrees′ or ′agrees′ with statement B and a value of 0 if the respondent ′strongly agrees′ or ′agrees′ with statement A.

- *Trust in officials relative to stranger*. This variable is based on two questions. The first question is ″Could you tell me how much you trust people you meet for the first time″, with possible answers ′Trust completely′, ′Trust somewhat′, ′Do not trust very much′, ′Do not trust at all′. The second question is ″Finally, how much do you trust local civil servants and locally elected officials?″, with the same range of possible answers. We assign numerical values from 1 to 4 for each answer, with higher values indicating higher trust. At the respondent level, the variable is constructed as the ratio of the trust in people met for the first time relative to trust in local civil servants and officials.

B.2 Variables from Household Survey Related to Tax Morale and Enforcement

- *Satisfaction with government services* This is an index variable, which is based on the average responses of households to three questions related to satisfaction with services. Possible responses are ′very satisfied′, ′somewhat satisfied′, ′neutral′, somewhat unsatisfied′, and ′very unsatisfied′. For each of the three questions, the answer is reverse coded such that higher values imply more satisfaction and all answers are standardized. The index variable is the unweighted average across the three standardized satisfaction questions outlined below

  1. ″In your personal dealings with tax collectors in Madina, how satisfied are you with the outcomes?″
  2. ″What has been your level of satisfaction with the overall quality of services offered by the local tax department of Madina″
  3. ″What has been your level of satisfaction with the overall quality of services offered by the local government of Madina?″

- *Integrity of government* This is an index variable, which is created as the unweighted average over the standardized responses to the different questions outlined below.

Questions are reverse coded where relevant such that higher answers always indicate more positive view on integrity and competency of the local government

1. "In your opinion, approximately what percent of the collections by the Madina Assembly will be put to good use for the benefit of the community?"
2. "If the Madina Assembly wants to improve all the roads, it will do this efficiently and without problems". There are five answers, ranging from 'strongly agree' to 'strongly disagree'.
3. "If the Madina Assembly wants to improve access to water for most citizens, it will be able to do so efficiently and without problems". There are five answers, ranging from 'strongly agree' to 'strongly disagree'.
4. "If the Madina Assembly needed to improve waste management, it would be able to do so efficiently and without problems". There are five possible answers, ranging from 'strongly agree' to 'strongly disagree'.
5. "Overall, how would you rate the Madina Assembly?". There are four possible answers, ranging from 'very competent' to 'not competent at all'.

- *Tax equity and efficiency efforts by government* This is an index variable, based on the respondent's strength of agreement with three statements. Possible answers to each question are 'agree strongly', 'agree somewhat', 'neither agree nor disagree', 'disagree somewhat', 'strongly disagree'. Answers are reverse coded such that higher values reflect stronger agreement, and standardized. The index is the average across the respondent's agreement with the statements below

  1. "Madina is making efforts to collect taxes in an efficient way"
  2. "Madina is making efforts to ensure everyone in their community pays their fair share of taxes"
  3. "Madina is making efforts to collect taxes that will be useful for local development of the community"

- *Enforcement and information capacity of the government* This is an index variable, which is created as the unweighted average over the standardized responses to the different questions outlined below. Questions are reverse coded where relevant such that higher answers always indicate stronger perceptions of enforcement and informational capacity

  1. "What share of households and businesses in the Madina Assembly do you think usually pay their taxes?" Answers range from 0% to 100%

2. ″Imagine a tax collector comes to your neighborhood, and someone refuses to pay. How likely do you think that the local government will pursue and enforce sanctions?″. There are four answers, ranging from ′very likely′ to ′very unlikely′.
3. ″Do you think the local government knows the precise address of your residence?″. There are four answers, ranging from ′very likely′ to ′very unlikely′.
4. ″Do you think the local government knows which of your neighbors did not pay property or business tax in 2020?″. There are four answers, ranging from ′very likely′ to ′very unlikely′.
5. ″Do you think the local government knows what you do for a living?″. There are four answers, ranging from ′very likely′ to ′very unlikely′.

B.3 Variables from Household Survey Related to Bribes

- *Any bribe (coercive or collusive)* This variable is based on two dummy variables. The first dummy variable takes a value of 1 if the household estimates that tax collectors will ask for any strictly positive unofficial payments when they are working in the field, and zero otherwise. This variable proxies for the likelihood of collusive bribes. The exact question is: ″Do you think it is likely that a local revenue collector will offer to take an unofficial payment from property owners/businesses in order to not make any return visits to their property?″ The possible answers were: ″very likely″; ″somewhat likely″; ″maybe″; ″not very likely″; ″very unlikely″. If a respondent answered ″very likely″, ″somewhat likely″ or ″maybe″, then the follow up question was: ″what is the amount that is typically asked for?″. We replace this answer with zero if the respondent′s first answer was ″not very likely″ or ″very unlikely″, and use this modified answer to construct the coercive bribe dummy. The second dummy variable takes a value of 1 if the household reports that the tax collector will pocket any positive amount out of a hypothetical 1000 Ghanaian Cedi collected from households (coercive bribe). The exact question is: ″Suppose a collector comes to a typical neighborhood in Madina and collects 1000 Ghanaian Cedi. How much of this money do you think the collector will submit to LANMA′s tax finance office account? And, how much will they put in their pockets?″. The variable used in the analysis takes a value of 1 if either the coercive dummy or the collusive dummy is equal to 1, and takes a value of 0 otherwise.
- *Any bribe (coercive)* This variable is a dummy variable which takes a value of 1 if the

household reports that the tax collector will pocket any positive amount out of a hypothetical 1000 Ghanaian Cedi collected from households. The exact question is: ″Suppose a collector comes to a typical neighborhood in Madina and collects 1000 Ghanaian Cedi. How much of this money do you think the collector will submit to LANMA's tax finance office account? And, how much will they put in their pockets?″.

- *Any bribe (collusive)* This variable is a dummy variable which takes a value of 1 if the household estimates that tax collectors will ask for any strictly positive unofficial payments when they are working in the field, and zero otherwise. The exact question is: ″Do you think it is likely that a local revenue collector will offer to take an unofficial payment from property owners/businesses in order to not make any return visits to their property?″ The possible answers were: ″very likely″; ″somewhat likely″; ″maybe″; ″not very likely″; ″very unlikely″. If a respondent answered ″very likely″, ″somewhat likely″ or ″maybe″, then the follow up question was: ″what is the amount that is typically asked for?″. We replace this answer with zero if the respondent's first answer was ″not very likely″ or ″very unlikely″, and use this modified answer to construct the coercive bribe dummy.

- *Collusive bribe (Likert scale)* This variable is the answer to the question ″Do you think it is likely that a local revenue collector will offer to take an unofficial payment from property owners/businesses in order not make any return visits to their property/business?″. The 5 possible answers range from ′very unlikely′ to ′very likely′. We assign numerical from 1 to 5 which are increasing in the likelihood.

- *Total bribe amount (in %)* This variable is constructed as the average % in coercive bribes amount and collusive bribe amount. The % coercive bribe amount is the %, of out a hypothetical 1000 Ghanaian Cedi collected by the collector, that the household estimates will be pocketed by the official. The exact question is: ″Suppose a collector comes to a typical neighborhood in Madina and collects 1000 Ghanaian Cedi. How much of this money do you think the collector will submit to LANMA's tax finance office account? And, how much will they put in their pockets?″ The collusive amount is the amount that the household estimates will be asked by the official as unofficial payment, expressed as a % of the household's actual property tax bill. The variable used in the analysis is the unweighted average of these two % measures. The exact question is: ″Do you think it is likely that a local revenue collector will offer to take an unofficial payment from property owners/businesses in order to not make any return visits to their property?″ The possible answers were:

”very likely”; ”somewhat likely”; ”maybe”; ”not very likely”; ”very unlikely”. If a respondent answered ”very likely”, ”somewhat likely” or ”maybe”, then the follow up question was: ”what is the amount that is typically asked for?”. We replace this answer with zero if the respondent’s first answer was ”not very likely” or ”very unlikely”, and use this modified answer.

- *Collusive bribe amount (% of tax due)* The collusive amount is the amount that the household estimates will be asked by the official as unofficial payment while conducting visits to the household, expressed as a % of the household’s actual property tax bill. The exact question is: ”Do you think it is likely that a local revenue collector will offer to take an unofficial payment from property owners/businesses in order to not make any return visits to their property?” The possible answers were: ”very likely”; ”somewhat likely”; ”maybe”; ”not very likely”; ”very unlikely”. If a respondent answered ”very likely”, ”somewhat likely” or ”maybe”, then the follow up question was: ”what is the amount that is typically asked for?”. We replace this answer with zero if the respondent’s first answer was ”not very likely” or ”very unlikely”, and express this modified answer relative to the value of the tax bill.

- *Coercive bribe amount (% of payment collected)* The coercive amount is the % that the household estimates will be pocketed by the tax collector out of a hypothetical 1000 Ghanaian Cedi that the official has collected as payments from households while working in the field. The exact question is: ”Suppose a collector comes to a typical neighborhood in Madina and collects 1000 Ghanaian Cedi. How much of this money do you think the collector will submit to LANMA’s tax finance office account? And, how much will they put in their pockets?”

- *Collusive bribe amount (in Ghanaian Cedi)* The collusive amount is the amount that the household estimates will be asked by the official as unofficial payment while conducting visits to the household. The exact question is: ”Do you think it is likely that a local revenue collector will offer to take an unofficial payment from property owners/businesses in order to not make any return visits to their property?” The possible answers were: ”very likely”; ”somewhat likely”; ”maybe”; ”not very likely”; ”very unlikely”. If a respondent answered ”very likely”, ”somewhat likely” or ”maybe”, then the follow up question was: ”what is the amount that is typically asked for?”. We replace this answer with zero if the respondent’s first answer was ”not very likely” or ”very unlikely”, and use this modified answer as the variable.

B.4 Variables from Household Survey Related to Learning and Targeting

- *Liquidity* This variable is created as the unweighted average over two household survey questions, which are outlined below. The survey questions are reverse coded such that higher values reflect lower liquidity constraints. Answers to both survey questions are standardized, and the liquidity index is in turn the unweighted average over the these two standardized survey variables. The two variables are

  1. "Think of a typical month. On how many days did you find yourself short of cash for basic expenditures for your house?". The answer can range from 0 to 30 days
  2. "In a typical month, imagine that one day you learn you need to pay an additional 300 Cedi fee in order to remain in your house. Could you find this money in the next 4 days?". The possible answers are 'Yes, with a little difficulty'; 'Yes, with great difficulty'; 'Very unlikely'; 'I could never pay this fee'

- *Income* This variable is based on the answer to the household question "What was the household's total earnings this past month?". The answer is given in Ghanaian Cedi. All survey answers are standardized.

- *Taxpayer knowledge* This variable is the unweighted average of six dummy variables which each take a value of 1 if the person answers 'Yes' to the individual questions outlined below, and take a value of 0 if the respondent answers 'No'. In turn, the unweighted average across the six variables is standardized, to facilitate comparison with the other hard-to-observe characteristics (*income* and *liquidity*).

  1. "Do you know of someone who received a letter from their MMDA summoning them to appear in court for non-payment of property rates"
  2. "Do you know of someone who was actually taken to court for non-payment of property rates?"
  3. Have you heard of any instance where a property owner had their property confiscated for non-payment of property rates?"
  4. "As best as you can remember, did you receive any text message earlier this year from your MMDA about paying the property rate?"
  5. "As far as you know, do the MMDA's have the legal authority to collect property rates?"

6. ″Have you heard of the fee-fixing resolution?″

- *Taxpayer knowledge – Enforcement* This variable is the unweighted average of three dummy variables which each take a value of 1 if the person answers ′Yes′ to the individual questions outlined below, and take a value of 0 if the respondent answers ′No′. In turn, the unweighted average across the three variables is standardized, to facilitate comparison with the other hard-to-observe characteristics (*income* and *liquidity*).

  1. ″Do you know of someone who received a letter from their MMDA summoning them to appear in court for non-payment of property rates″
  2. ″Do you know of someone who was actually taken to court for non-payment of property rates?″
  3. Have you heard of any instance where a property owner had their property confiscated for non-payment of property rates?″

- *Taxpayer knowledge – Tax code* This variable is the unweighted average of three dummy variables which each take a value of 1 if the person answers ′Yes′ to the individual questions outlined below, and take a value of 0 if the respondent answers ′No′. In turn, the unweighted average across the three variables is standardized, to facilitate comparison with the other hard-to-observe characteristics (*income* and *liquidity*).

  1. ″As best as you can remember, did you receive any text message earlier this year from your MMDA about paying the property rate?″
  2. ″As far as you know, do the MMDA′s have the legal authority to collect property rates?″
  3. ″Have you heard of the fee-fixing resolution?″

- *Index – Hard to observe* This variable is the unweighted average of the three index variables *Liquidity*, *Income* and *Taxpayer knowledge*

- *Tax bill value* This variable is based on the administrative data and measures the total amount of taxes that are owed. The total amount owed is the sum of the current year′s property taxes and outstanding arrears due to non-payment in full of the past year′s property taxes. The variable is standardized to facilitate comparison with other ′easy-to-observe′ characteristics (*previous tax payment* and *assets*).

- *Previous tax payment* This variable is based on the administrative data and measures the payment status from the previous fiscal year. It takes a value of 1/2/3 if the past year's property taxes were not paid at all/partially paid/fully paid. The variable is standardized to facilitate comparison with other 'easy-to-observe' characteristics (tax bill value and assets).

- *Assets* This variable is the sum over how many of the following assets the household currently possesses: motorbike; car or truck; television; electric generator; sewing machine; radio. In turn, the variable is standardized to facilitate comparison with other 'easy-to-observe' characteristics (*tax bill value* and *previous tax payment*).

- *Index – Easy to observe* This variable is the unweighted average of the three index variables *tax bill value*, *previous tax payment* and *assets*.

B.5 Variables from Collector Surveys

- *Challenge to navigate in the field* This variable is a dummy variable which takes a value of 1 if the respondent 'strongly agrees' or 'agrees' with the statement "Finding my way around my collection unit was a challenge for me this week"; the dummy variable takes a value of 0 if the respondent answers 'neither agree nor disagree', 'disagree' or 'strongly disagree.

- *Challenge to locate taxpayers* This variable is a dummy variable which takes a value of 1 if the respondent 'strongly agrees' or 'agrees' with the statement "Locating bill recipients was challenging for me this week"; the dummy variable takes a value of 0 if the respondent answers 'neither agree nor disagree', 'disagree' or 'strongly disagree.

- *Knowledge about households which are willing to pay* This variable takes a value of 1 if the respondents chooses statement A "I think I have a good understanding of which properties are more able and willing to pay and am able to focus my efforts on them" rather than statement B "I put a lot of effort to get my job done, but it remains unclear to me which exact properties are more likely or willing to pay their property rates". The variable takes a value of 0 if the respondent picks statement B. Respondents had to pick the statement which "you would say best characterizes your work in the field over the past weeks".

- *Focus on households that are able to pay* This variable takes a value of 1 if the respondent uses 'all the time' or 'often' the collection strategy "Go to areas on specific days where I know property owners are more likely to be able to pay"; the variable takes a value of 0 if the respondents uses this strategy 'only from time to time', 'not much' or 'never'.

- *Focus on households that are aware of tax payment duty* This variable takes a value of 1 if the respondent uses 'all the time' or 'often' the collection strategy "Go to areas where I know most taxpayers are aware of their duty to pay property rates"; the variable takes a value of 0 if the respondents uses this strategy 'only from time to time', 'not much' or 'never'.

- *Focus on households that are satisfied with public goods* This variable takes a value of 1 if the respondent uses 'all the time' or 'often' the collection strategy "Go to areas where I know owners are more satisfied with the delivery of public services and are more likely to pay"; the variable takes a value of 0 if the respondents uses this strategy 'only from time to time', 'not much' or 'never'.

- *Focus on collections with hard-to-observe household characteristics* This variable measures the frequency with which collectors make use of the three strategies that target hard-to-observe household characteristics: *focus on households that are aware of tax payment duty*, *focus on households that are able to pay*, and *focus on households that are satisfied with public goods*. The variable is the average reliance on these three strategies, and takes a value between 0 and 1.

- *Focus on collections with easy-to-observe household characteristics* This variable measures the frequency with which collectors make use of six strategies that target easy-to-observe household characteristics. For each strategy, outlined below, we measure reliance with a value of 1 if that collection strategy is used 'often' or 'all the time' and 0 if it is used 'only from time to time', 'not much' or 'never'. In turn, the variable is the average reliance across these six strategies, and takes a value between 0 and 1. The six strategies considered are

    1. "Go to areas where I know most taxpayers have paid property rates in the past year"
    2. "Go to areas where I know there are many properties with high property rates"
    3. "Go to areas where I know there are many property rate payers that have not yet paid this year's rates"
    4. "Go to areas which are close to the main road/center of activity"

5. ”Go to areas which are close to my home”
6. ”Go to areas which are closer to the Madina headquarters”

- *Difference in strategies: Hard versus easy to observe* This variable is the difference between the variable *’Focus on collections with hard-to-observe household characteristics’* and the variable *’Focus on collections with easy-to-observe household characteristics’*
- *Unable to locate properties and owners* This variable measures the collectors’ extent of agreement with two statements: ”Finding my way around my collection unit was challenging”; ”Locating bill recipients was challenging for me this week”. For each statement, the respondent can answer ’strongly disagree’, ’disagree’, ’neither agree nor disagree’, ’agree’, ’strongly agree’. We assign numerical values from 1 to 5, with larger values indicating stronger agreement. The answer to each statement is standardized, and the variable is the average over the two standardized answers.
- *Wrong information printed on bills* This variable measures the collectors’ extent of agreement with the statements: ”Some of the bills I tried to deliver this week had the wrong addresses”; ”Some of the bills I tried to deliver this week had the wrong amounts”. For each statement, the respondent can answer ’strongly disagree’, ’disagree’, ’neither agree nor disagree’, ’agree’, ’strongly agree’. We assign numerical values from 1 to 5, with larger values indicating stronger agreement. The answer to each statement is standardized, and the variable is the average over the two standardized answers.
- *Resistance from property to accept bill* This variable measures the collectors’ extent of agreement with three statements: ”Collection was challenging this week because bill recipients preferred not to pay in cash”; ”Collection was challenging this week because bill recipients preferred mobile payments, but I was not able to accept mobile payments”; ”Collection was challenging this week because bill recipients said that they did not trust me to collect their payment”. For each statement, the respondent can answer ’strongly disagree’, ’disagree’, ’neither agree nor disagree’, ’agree’, ’strongly agree’. We assign numerical values from 1 to 5, with larger values indicating stronger agreement. The answer to each statement is standardized, and the variable is the average over the three standardized answers.
- *Supervisors do not monitor field activities* This variable measures the extent to which collectors perceive that their supervisors are not monitoring their work. Specifically, we ask the collector’s extent of agreement with the statement: ”My supervisors spent a lot of time monitoring my work this week”. For each statement, the

respondent can answer ’strongly disagree’, ’disagree’, ’neither agree nor disagree’, ’agree’, ’strongly agree’. We assign numerical values from 1 to 5, with larger values indicating stronger *disagreement*. Values are standardized to be comparable with other outcomes.

- *Supervisors do not check mistakes made in the field* This variable measures the extent to which collectors perceive that their supervisors are not checking mistakes made by collectors in the field. Specifically, we ask the collector’s extent of agreement with the statement: ”My supervisors checked on me regularly this week to make sure I was not making mistakes”. For each statement, the respondent can answer ’strongly disagree’, ’disagree’, ’neither agree nor disagree’, ’agree’, ’strongly agree’. We assign numerical values from 1 to 5, with larger values indicating stronger *disagreement*. Values are standardized to be comparable with other outcomes.

- *Supervisors are unavailable for support* This variable measures the extent to which collectors perceive that their supervisors are not available to support the collectors in the field. Specifically, we ask the collector’s extent of agreement with the statement: ”My supervisors were available to help me this week when I needed them”. For each statement, the respondent can answer ’strongly disagree’, ’disagree’, ’neither agree nor disagree’, ’agree’, ’strongly agree’. We assign numerical values from 1 to 5, with larger values indicating stronger *disagreement*. Values are standardized to be comparable with other outcomes.

- # *Unsuccessful visits per successful visit* This variable is the answer to the question ”There are many challenges to getting things done in the field. Looking back at this past week, let us think about the unsuccessful visits you made to properties. A successful visit is a visit to a property where you were able to complete the task you had planned. For every successful visit, how many unsuccessful visits would you say that there were, for the typical property?”

- *Total hours worked per week* This variable is the product of the following two questions: ”How many days did you work this week?”; and, ”During the days where you did work this week, what would you say is approximately the number of hours you worked?”.

- *Average # hours spent to deliver one bill* This variable is the ratio of total bills delivered per week (self-reported by the collector) divided by the variable *total hours worked per week*.

- *Satisfaction and happiness on job* This variable measures the collectors' extent of agreement with three statements: "Overall, this was a productive week for me"; "Overall, I was content while working this week"; "Overall, I am satisfied with my job". For each statement, the respondent can answer 'strongly disagree', 'disagree', 'neither agree nor disagree', 'agree', 'strongly agree'. We assign numerical values from 1 to 5, with larger values indicating stronger agreement. The answer to each statement is standardized, and the variable is the average over the three standardized answers.

C Additional Analyses

C.1 Days Since Bill Delivery

Our results have shown that tax collectors face time-constraints when trying to deliver their bills, and technology allowed treatment collectors to more quickly find property owners and deliver bills. This frees up more time for treatment collectors to conduct repeat visits; if the likelihood that a property owner makes a positive payment is increasing in the number of repeat visits, this 'mechanical' relaxation of time-constraints would lead to a positive treatment effect on tax collection. In this sub-section, we investigate whether the full observed effect is driven by time-constraints, without any room for other mechanisms such as learning.[29] The idea is to control for the time since bill delivery in the equation which explains tax outcome (equation 2). If technology's impact is fully mediated through the mechanical time-advantage, then we should see no treatment effect remaining once we have controlled for time since bill delivery.

To implement this exercise requires data at the bill-level on both tax payment and the exact date of bill delivery. We attempted to collect delivery dates by asking tax collectors to maintain a diary during the tax campaign. We compiled the diaries at the end of the campaign which should, in principle, record the date of delivery for each bill that the collector was assigned to. In practice, the data quality of these diary entries is limited, for several reasons. First, conducting continuous quality-checks on diary entries during the tax campaign itself was challenging. Second, upon compiling the diaries at the end of the campaign, we learned that some collectors had been filling out entries at the end of each campaign week – introducing measurement error around the exact date of delivery for a particular bill. Third, while providing aggregate daily information on the number of bills delivered was part of the established process (Section 4.1), the requirement to maintain a diary was introduced during our experiment and was new to the collectors. We observe incomplete diary entries for some bills (e.g. where the bill was claimed to be delivered, but the information on delivery date is missing), and it is possible that collectors paid less attention to maintaining the diary than to submitting aggregate daily information to their supervisors. For these reasons, we view the results based on aggregate delivery date as more precise (Figure 2), and consider the results from this sub-section as secondary.

[29]Note that we have already established the existence of selection on bill delivery and tax payment in terms of hard-to-observe characteristics (Figures A10 and 8), which is consistent with targeting based on learning. Moreover, our model has quantitatively explored the importance of learning for the observed treatment effects. The results in this sub-section are complementary to those analyses.

With these caveats in mind, Figure A14 plots the density distribution of days since bill-delivery separately for the treatment and control groups. We focus on the sample of bills for which we also have household data, but results are similar based on the full experimental sample. We measure days since bill delivery as the number of days between the official end-date of the tax campaign and the date of delivery based on the diaries. Thus, a larger number indicates that the collector had more days available to conduct follow-up visits and collect payments before the end of the campaign. Consistent with the dynamics of bill-delivery based on aggregate data (Figure 2), this figure shows that treatment collectors delivered more bills in the early days of the tax campaign – and consequently have more days available to conduct follow-up visits. We can reject with confidence that the density distributions of the two groups are equal (Kolmogorov-Smirnov D-statistic= 0.095 with p-value= 0.001).

In the time-constraint mechanism, tax performance is an increasing function of days since delivery. Leveraging the fact that we observe tax outcomes and delivery dates at the bill level, in Figure A15 we plot tax performance as a function of days since delivery: the likelihood of making a tax payment in panel A, and the total amount paid in panel B. It is important to note that these are descriptive associations, since both the characteristics of the property that receives a bill and the date of bill delivery are endogenous. To visualize the associations, we create five days-since-delivery bins of equal size (quintiles), and calculate the average tax outcomes separately by quintile and treatment-control groups. Consistent with the time-constraint mechanism, the control group pattern shows that the likelihood of making a tax payment is increasing in the days since delivery. However, the profile in the treatment group is distinctly different – maintaining a positive slope, but being shifted upward everywhere relative to the control group profile. In the time-constraint mechanism, the profiles of the treatment and control groups would be identical – and the positive tax performance impact would come from the treatment group having more days since delivery (Figure A14). To formally test for statistical differences across profiles, we use the household sample to estimate

$$y_{hqc} = \beta_q \mathbf{1}(Tech)_c \cdot \pi_q + \pi_q + \epsilon_{hqc},$$

where y_{hqc} is the tax outcome of household h in quintile q and collection unit c, and π_q are fixed effects for the five quintiles of days-since-delivery. Standard errors are clustered at the collection unit. β_q is indexed with q because the treatment dummy is interacted with the quintile group fixed effects. In the top-left corners of each panel, we report the F-statistic which tests the joint significance of the five β_q coefficients. For both the likelihood of tax payment (F-statistic= 7.78, p-value= 0.007) and total tax payment

(F-statistic= 3.30, p-value= 0.011), we can reject that the two profiles are the same.

The difference in profiles suggests that the time-constraints may not be the only mechanism which drives our main results. To complete this investigation, we augment our main estimation equation for tax outcomes (equation 2) with the measure of days since delivery, $dayssince_h$:

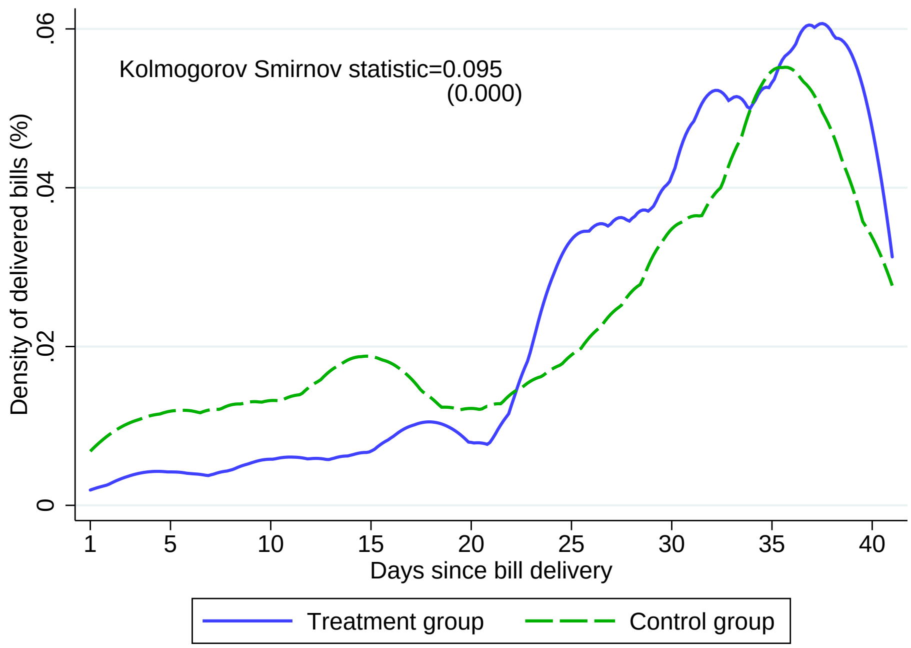
$$y_{hdc} = \beta \cdot \mathbf{1}(Tech)_c + \mu \cdot dayssince_h + \iota_d + \Omega \cdot X_{hc} + \epsilon_{hdc},$$

where the interpretation is that β reflects the impact of technology after controlling for time-constraints. We assume a linear relationship between the outcome and days-since-delivery; the results are robust to more flexible functional forms. As mentioned earlier, this interpretation is challenged by the issue that the variable $dayssince_h$ is endogenous – including to the treatment.[30] Notwithstanding this concern, the results are presented in Table A9; for each outcome, the odd-numbered (even-numbered) column corresponds to the specification without (with) the days-since-delivery control. In the first two columns, we observe that controlling for days since delivery reduces the treatment coefficient on total visits by almost 50% and the coefficient loses its statistical significance. The coefficient on days-since-delivery itself is positive and strongly significant. This suggests that technology's impact on total visits is largely mediated by its impact on date of bill delivery. In columns (3) and (4), we see that controlling for days-since-delivery does reduce the treatment impact on likelihood of tax payment, but only by 14% and the treatment coefficient remains statistically significant. In other words, time-constraints appear to account for only a (relatively small) part of the technology effect, with the remaining impact possibly driven by learning and targeting. For total tax payment (columns 5 and 6), the inclusion of days-since-delivery reduces the treatment effect by only 10.5% and the technology coefficient remains strongly significant. Finally, while these results were estimated in the full sample (households without a bill delivered were flagged with the dummy ι_h), we find similar results in the sample which conditions on a bill being delivered (columns 7 and 8).

The analysis in this sub-section remains limited due to identification and data-concerns, but the various pieces of evidence do suggest an important role for learning (or other mechanisms) above and beyond the 'mechanical' relaxation of time-constraints.

[30] We estimate the above equation in the full household sample, but $dayssince_h$ is only defined for the sub-set of households that were delivered (d) a bill. To maintain the full sample, we assign an arbitrary value (-400) to $dayssince_h$ to all households with no bill delivered, and include a fixed effect, ι_h, which flags these d values. Maintaining the full sample allows us to estimate with improved precision the coefficients Ω for the household and collection unit characteristics X_{hc}; results are qualitatively similar if we restrict the sample to households with bills delivered.

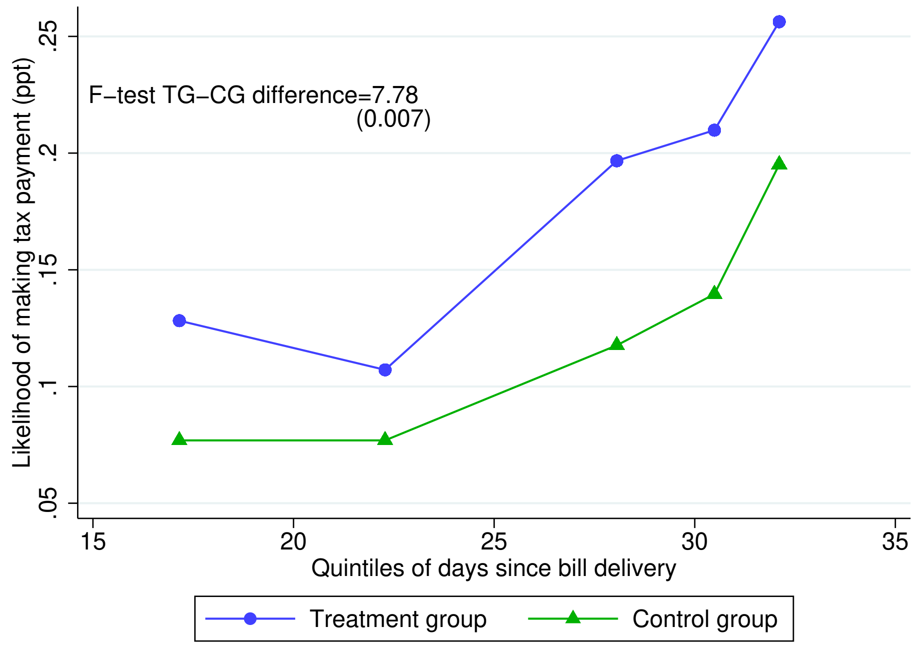
Figure A14: Distribution of Bill Delivery Dates



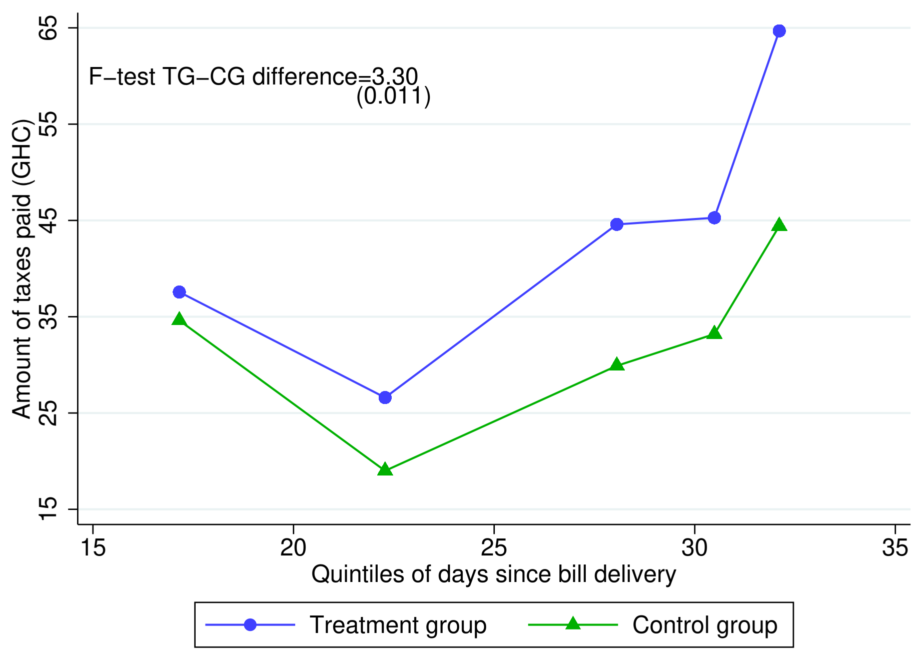
Notes: This figure shows the distribution of bills delivered by their delivery date, separately for the treatment group and the control group. The statistic reported in the top-left corner is the Kolmogorov-Smirnov D-statistic which tests that the treatment and control distributions are equal. The p-value for the D-statistic is reported in parentheses.

Figure A15: Tax Outcomes as a Function of Bill Delivery Dates

(a) Likelihood of Making Positive Tax Payment



(b) Taxes Collected



Notes: These panels show the associations between days since bill delivery and, respectively, likelihood of making a positive tax payment (panel A) and total taxes paid (panel B). The distribution of days since bill delivery (Figure A14) is separated into five quintiles (bins of equal size), and the average value of the tax outcome is calculated separately by quintile and group (treatment and control). The F-statistic reported in the top left-corner is the statistic which tests the hypothesis that the gaps between treatment and control are jointly zero in all five quintiles. This F-statistic is based on estimating equation C.1.

Table A9: Main Impacts While Controlling for Days Since Bill Delivery

	Total visits (in %)		Any positive tax payment		Total tax payment (in GHC)		Total payment per bill delivered	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Technology	0.094*	0.049	0.043**	0.037*	25.9**	23.2**	47.3**	48.3**
	(0.050)	(0.050)	(0.021)	(0.020)	(10.9)	(10.1)	(19.6)	(18.2)
Days since bill delivery		0.014***		0.005***		0.507		0.351
		(0.004)		(0.001)		(7.359)		(1.238)
Household controls	X	X	X	X	X	X	X	X
Collector controls	X	X	X	X	X	X	X	X
Strata FE	X	X	X	X	X	X	X	X
Mean in CG	0.67	0.67	0.16	0.16	41.0	41.0	80.9	80.9
Observations	4334	4334	4334	4334	4334	4334	2276	2276
Clusters	56	56	56	56	56	56	56	56

Notes:The regression model and outcomes in this table are the same as in Table 4, with the only addition that we include in even-numbered columns the variable which measures the number of days since the bill was delivered. This variable is based on the administrative data, and measures the number of days between the date when the bill was delivered and the end-date of the tax campaign. For a description of the regression model, please refer to Section 4.2.