Fiscal Exchange and Tax Compliance: Strengthening the Social Contract Under Low State Capacity*

Laura Montenbruck†

March 5, 2025

[Click here for the latest version]

Abstract

This article provides evidence that increased salience of public service provision can strengthen the social contract and increase tax compliance in a low-capacity setting. I conduct a field experiment randomizing information about public service provision across 5,494 property owners and tenants in Freetown, Sierra Leone. Receiving information increases property tax payments by 20% on average. The effect is driven by increases in tax compliance on both the extensive and intensive margin. Residents of low-value properties are 7–16 percentage points more likely to pay taxes when informed about public services that are both geographically accessible and respond to the citizens' most urgent needs, suggesting a benefit-based approach to taxation. Revenue effects are largely driven by residents of high-value properties, who depend less on the public provision of services, and for whom the treatment seems to act as a more general signal of government performance.

JEL codes: H20, O23, D73
Keywords: Social contract, Property tax, Public services, Tax compliance

*I am very grateful to Antonio Ciccone, Michèle Tertilt, Chris Udry, Silvia Vannutelli, Niccolò Meriggi, Dean Karlan, Molly Schnell, and Vanessa van den Boogaard, as well as participants at research seminars at Northwestern University and the University of Mannheim, the OXDEV workshop, and various conferences for valuable comments and feedback. I further thank my field coordinator Sheriff Kelo, my data collection supervisors, Ishmatu Juliet Sheriff, Mohamed Bangura, and Abdul B. Kamara, and my entire team of enumerators, without whom this project could not have been implemented. This research was funded through grants from the International Growth Centre (projects SLE-21175 and SLE-22083). Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project A04) is gratefully acknowledged. The project was pre-registered under AEA preregistration ID: AEARCTR-0009207.

†Stockholm University, laura.montenbruck@su.se

1 Introduction

The ability to tax is central to the process of state development (Besley and Persson, 2013). Tax revenue enables the provision of public services and allows countries to form capital from domestic sources (Kaldor, 1963). Despite its crucial role, many low-income countries fail to generate adequate levels of tax income.[1] The effectiveness of states—including their ability to tax—is associated with two key factors: coercive power (e.g., Hobbes, 1651), and the existence of a social contract (e.g., Locke, 1690; Rousseau, 1762). Low administrative capacity limits the ability to enforce tax policies in low-income countries (Besley and Persson, 2013). A salient contractual relationship between citizens and the state that acknowledges and fulfills reciprocal obligations, however, has the potential to encourage "quasi-voluntary" tax compliance (Levi, 1988).

In her historical account of the power to tax, Levi (1988) argues that citizens' quasi-voluntary tax compliance—the decision to pay taxes in a setting where non-compliance is technically punishable by law—depends crucially on their confidence that the state fulfills its part of the tax contract. Embedded in a growing literature on tax morale, recent empirical work has investigated whether highlighting the link between taxation and public service provision affects tax compliance (see, e.g., Luttmer and Singhal, 2014). Focusing mainly on high- and middle-income countries, these studies either emphasize the link between taxation and public service provision theoretically, or provide aggregate information about the use of tax revenue.[2] However, in low-capacity settings, where trust in the reciprocity of the government tends to be low (Prichard, 2017; Dom et al., 2022), and it is less recognized that taxes finance important public services (Mascagni, 2018), such general information may not be enough to credibly signal that the government fulfills the social fiscal contract.

In this paper, I test whether specific, targeted information about recently provided public services increases property tax compliance in a low-capacity setting. The study is situated in Freetown, Sierra Leone, where property tax is a major source of municipality level revenue.[3] The Freetown City Council (FCC) is making substantial investments to improve public service provision throughout the city and property owners report being willing to pay taxes in

[1] Whereas the tax-to-GDP ratio amounts to 34% in OECD countries, it is as low as 16% on average among 31 African countries for which tax data is available. With 12%, the ratio is even lower in Sierra Leone (OECD, 2022).

[2] "Your tax payment contributes to the funding of publicly financed services in education, health and other important sectors of society." (Bott et al., 2020) is an example of a theoretical message emphasizing fiscal exchange. "In the first 6 months of this year, CVP's collection contributed to placing 28 new streetlights, water connections in 29 streets and sewerage networks in 21 blocks." (Castro and Scartascini, 2015) is an example of aggregate information about the use of tax revenue.

[3] In 2021 and 2022, property tax accounted for 46% and 56% of municipality level revenue, respectively. These numbers only account for realized property tax payments.

exchange for the provision of public services (Grieco et al., 2019). However, despite being registered, only few property owners pay taxes.[4]

Against this background, I design an intervention that randomly provides residents of registered properties with specific and geographically targeted examples of public services delivered by the local government. To do so, I verify and geo-reference all locations at which public services have been provided in the three years prior to the intervention. I create a database including a total of 277 services aiming, among other things, at improving the quality of education, access to water, and sanitation. I link this information to administrative data available from the property tax registry, including the specific location of each property. This allows me to inform individuals about the closest location at which a certain type of service has recently been delivered.[5] In particular, treated individuals are informed about (a) a specific type of service, (b) an example location at which the service has been provided, and (c) the approximate time it takes them to get to said location from their house.[6]

To broaden our understanding of the mechanisms at play, I elicit and study the role of public service preferences. Laboratory experiments suggest that providing services in line with taxpayers' elicited preferences has the potential to increase tax compliance (Alm et al., 1993; Lamberton et al., 2018), supporting a theory of benefit-based tax compliance. However, evidence on the real-world link between public service preferences and tax payments is limited.[7] Thus, during a baseline interview, I ask individuals about both their most and least preferred type of service to be provided by the local government. More than half of all individuals indicate improved access to water as their top priority.[8] Notably, this priority coincides with the municipality's current focus on providing access to water.

I use elicited preferences to design two different treatment arms: In the first group, individuals receive information about the type of service they considered their most preferred service. I refer to this treatment as the *match* treatment. In the second group, I inform individuals about a service of their least preferred type. I refer to this treatment as the *mismatch* treatment, as it creates the largest mismatch possible between the type of service

[4] In 2021, the year prior to the intervention, property tax was paid for only 22% of all registered properties.

[5] To account for the fact that these locations are not selected randomly—they may be selected because of a previous lack of services, or reinforce existing differences in service provision across the city—and in the absence of detailed data on public services beyond those recently provided, I account for geographical differences in the base level of public services using ward fixed effects throughout my analysis.

[6] While the example location is the closest one at which the respective type of service has been provided, individuals are not explicitly informed about this fact.

[7] Survey experiments show that eliciting preferences can increase tax morale (Sjoberg et al., 2019; Casal et al., 2016). Khan et al. (2022) tests the effect of providing services aligned with individuals' preferences on property tax payments in Pakistan with no final results available yet. By contrast, I provide *information* about the provision of public services aligned or misaligned with preferences.

[8] As compared to this, slightly less than 50% of individuals claim disaster management and prevention or environmental management to be their lowest priority.

individuals receive information about and the type of service they consider their top priority. I compare these two treatment groups against a pure control group of individuals who receive no information at all.[9]

Based on the above-mentioned evidence, I expect that receiving information about services aligned with one's preferences has a positive effect on tax compliance. However, it is ex ante less obvious what effect information about one's least preferred service will have. If individuals have zero preference for the provision of their least preferred service, informing individuals about the latter will likely not affect tax compliance. If, however, individuals have a positive preference for the provision of their least preferred service, the effect can be expected to be positive and increasing in the closeness of preferences. The effect of the mismatch treatment may also interact with prior knowledge about the provided service, and as such the extent to which the treatment provides new information. For instance, individuals may choose not to acquire (costly) information about a service that they have a low preference for and therefore may not be aware of its provision. In this case, the mismatch message would on average provide more new information, thus shifting priors and potentially increasing tax compliance. Whether the mismatch treatment does or does not affect tax compliance, and if so, through what channel, will be investigated empirically.

I implement this intervention in a large-scale randomized controlled trial (RCT) with 5,494 residents of registered properties. Properties are sampled from the universe of registered properties, with eligibility being restricted to properties owned by a single-property owner for which at least one phone number is available (N=46,844).[10] I randomly allocate one third of these individuals to the control group, one third to the *match* and one third to the *mismatch* group. The treatment message is conveyed during a short phone call conducted by a qualified enumerator. Right after the call, treated individuals receive a short message (SMS) summarizing the main information provided during the call.

To investigate the effects of this intervention on property tax compliance, I use administrative tax payment data made available by the Freetown City Council (FCC). This data allows me to capture actual payments made rather than declarations of income or revenue, which determine the tax base, but not actual compliance. I complement this analysis with information from an endline survey with a subset of individuals. This data allows me to look at intermediary outcomes related to the reciprocity channel, as well as to investigate

[9]While this does not allow me to directly compare the effects of the fiscal exchange message against that of alternative messages such as reminders or deterrence messages, I look at these channels as potential drivers of the observed effects using suggestive evidence from an endline survey as well as results on the heterogeneity of effects.

[10]Single-property ownership is approximated using the names and phone numbers associated with each property. A more detailed description of the process can be found in Section 4. These properties represent approximately 45% of all registered properties.

potential alternative mechanisms through which the treatment could raise compliance, such as increases in perceived government capacity to track and punish evaders or awareness of one's tax duty.

I find that specific and geographically targeted information about recently provided public services significantly increases tax compliance by the payment deadline. Approximately six months after the intervention, treated individuals pay on average USD 1.5 (SLE 23.8) more taxes, a 20% increase as compared to the control group mean.[11] For the 3,660 treated individuals, the treatment increased timely payments by more than SLE 87,000, or approximately USD 5,600. Extending the intervention to all 46,844 eligible properties could increase tax revenue by more than USD 71,700. With a conservative per-individual cost of SLE 14.2 (including the cost of mapping all services as well as enumerator and phone costs for implementing the intervention), the intervention generates 1.7 times its cost. Almost half of the cost is driven by expenses for phone calls and SMS—which are particularly expensive in Sierra Leone. As the government often reaches more advantageous agreements with network providers, this is likely a lower bound of the return to providing public service information at scale.[12] These findings suggest that specific information about the provision of public services can be an effective tool to increase tax compliance.

To get a more nuanced understanding of the mechanisms driving these effects, I look at two different types of individuals: poor individuals, who reside in low-value properties, and rich individuals, who reside in high-value properties. Poor individuals rely heavily on public service provision and have lower access to public services.[13] Accordingly, these individuals react to the treatment message in different ways as compared to rich individuals. Residents of low-value properties increase tax payments in response to information aligning with their most preferred type of service, indicating a benefit-based approach to tax compliance as expected. The effects seem to be driven predominantly by increases on the extensive margin of tax compliance. Residents of high-value properties, on the other hand, increase tax payments in response to the mismatch treatment. As these individuals rely on average less on public service provision, tax compliance seems to be less closely related to personal benefit from the services they are informed about. Indicative evidence suggests that these individuals consider the treatment as a positive signal of government performance more generally. Whereas the mismatch treatment shifts individuals' priors about the government's performance in providing public services, the match treatment does not. This suggests that individuals may

[11] Nominal exchange rates for September 30, 2022 taken from https://www.exchangerates.org.uk/USD-SLL-spot-exchange-rates-history-2022.html.

[12] In previous interventions conducted by the government itself, network providers even allowed free usage of their network.

[13] Rich individuals, on the other hand, not only have higher access, but also often provide services privately and therefore depend less on the government's fulfillment of the social contract.

have been more informed about their most preferred service ex ante and explains why the match treatment does not affect tax compliance among residents of high-value properties.

The differential effects of the two treatment arms provide important insights into the role of personal benefit from the provision of public services for tax compliance. Beyond stated preferences, however, several other factors can determine the extent to which individuals benefit from the services they are informed about. Two aspects are of particular interest given the design of the intervention: *access* and *contextual importance* of specific types of services. Both factors implicitly form part of the treatment message: information about access can be derived from the measure of *distance* (approximated by commuting time) stated in the message; considered independently from preferences, different *types of services* are likely perceived as differentially beneficial among all individuals facing the same environment. In the context of Freetown, I focus on the role of access to water—a service that can unambiguously be considered essential—as compared to any other type of service. Not only is water a basic need and prerequisite for being able to benefit from other services. It is also a service that is poorly provided in the city of Freetown.[14]

I exploit the variation in distance to and type of service across individuals and within treatment groups to investigate heterogeneity in treatment effects with respect to the individualized content of the message. By design, the composition of the types of services individuals are informed about, as well as distance to that service, vary across treatment arms. However, both treatment arms as well as the control group are balanced in terms of distance to the closest service of each type and well-balanced with respect to the share of individuals indicating a specific service as their most or least preferred type.

Disentangling the effects of the treatment depending on the message's content reveals that the effects on tax compliance are driven by individuals who receive information about the provision of water close to where they live. In line with the idea that personal benefit of public service provision matters predominantly for those who depend on public provision of services, the specific content of the treatment message correlates with tax compliance only among residents of low-value properties. For this subgroup of individuals, the likelihood of making a payment increases by 7-16 percentage points, or 62–134% as compared to the respective control group, when informed about the provision of water at an accessible distance.[15] These results underline the importance of considering multiple layers of perceived benefit of public service provision, including not only stated preferences, but also access and the more general

[14]Indicative evidence suggests that around 39% of registered properties in Freetown have no access to water at all, and only around 45% are connected to the piped water system (Grieco et al., 2019). The situation is even more dire when considering the entire population of Freetown, with only 5% of individuals having access to piped water (Freetown City Council, 2019).

[15]As the intensive margin effects are driven almost exclusively by high-value properties, the content of the message does not correlate with these effects.

valuation of certain types of services.

To demonstrate that treated individuals indeed increase tax compliance based on a fiscal exchange (reciprocity) motive, I provide indicative evidence ruling out two alternative mechanisms: increases in perceived extractive capacity of the government, and increases in tax awareness. To show the former, I look at self-reported outcomes related to the perceived capacity of the government to collect taxes and punish evaders. I find no evidence that increased fear of being caught or punished when evading drives the effects on tax compliance. Importantly, this is also true for residents of high-value properties who received the mismatch treatment, suggesting that the effect of the mismatch treatment cannot be explained by fear of enforcement. The effects do not seem to be driven by increased awareness of one's tax duty either. Intuitively, this is because the treatment message was designed to focus on public service provision without mentioning tax obligations.[16]

Finally, I show that the effects on tax compliance persist over the entire tax cycle, suggesting that treated individuals do not only pay taxes earlier, but actually pay more taxes. By the end of the 2022 tax cycle, individuals living in low-value properties who received information about water at close distance are significantly more likely to have paid any tax. This translates into a 21% increase in the average amount of tax paid among individuals in the match group—who by design are about 6 times as likely to receive information about water as individuals in the mismatch group. Among residents of high-value properties, those who received the mismatch treatment pay on average 29% more taxes during the entire 2022 tax cycle than those in the control group.

This paper contributes to several strands of literature. First, I add to a growing literature on property taxation in low-income countries. As urbanization is expected to increase drastically, cities need to generate more revenue to provide adequate levels of public services.[17] The taxation of properties has been postulated to be the "best available instrument for funding local governments, because they are generally progressive, economically efficient, and potentially closely linked to local services" (Moore et al., 2018, p.151). However, countries face multiple challenges along the process, including the registration and valuation properties, determining optimal tax rates, and incentivizing compliance. Recent work has therefore investigated how countries can successfully register taxpayers (Weigel, 2020), assess the value of properties (Knebelmann and Pouliquen, 2023), determine the optimal tax rate (Bergeron

[16]This conclusion is also supported by several other pieces of evidence. First, I show that the treatment effect is driven by individuals who previously contributed to the provision of community services, i.e., individuals who care disproportionally about the provision of public services. Second, the fact that the message's specific content matters contradicts the hypothesis that being treated increases tax compliance via increased tax awareness independently of the fiscal exchange mechanism.

[17]In 2018, approximately 42.5% of the African population lived in urban areas. This number is expected to increase to 58.9% by 2050 (United Nations, 2019).

et al., 2023; Brockmeyer et al., 2023), or increase property tax compliance (Khan et al., 2019; Krause, 2020; Balán et al., 2022; Collin et al., 2022; Fernandez et al., 2023). I contribute to this literature by providing evidence on the effectiveness of specific, targeted information about public service provision on tax compliance in a low-enforcement setting.

Second, I add to the literature on the effects of messages emphasizing the link between taxation and public service provision, thereby trying to strengthen the social fiscal contract and nudge tax compliance.[18] Previous studies from high- and middle-income countries have found that these types of messages are generally insufficient to increase compliance.[19] Evidence from low-income countries is more scarce, but largely supports these findings. In particular, studies from Rwanda and Uganda find no effects of such general messages on average (Mascagni and Nell, 2022; Cohen, 2023).[20] Importantly, previously studied messages are general in nature and merely inform individuals that tax revenue is necessary to finance public services or provide aggregate, hard-to-verify information about the use of tax revenue. By contrast, I provide individuals with targeted information about services delivered by their local government. This distinction is substantial as citizens often lack awareness of the public services provided to them, or have biased beliefs about how the state spends its revenue (Khan et al., 2022; Giaccobasso et al., 2023) and confidence in the reciprocity in the government is low (Prichard, 2017; Moore et al., 2018; Dom et al., 2022). This can explain why more generic information may not be enough to motivate them to pay taxes. Crucially, by providing individuals with specific examples of services delivered by their government, my treatment makes the government's effort in fulfilling its part of the social fiscal contract more salient. My work shows that providing specific information about public services delivered by the local government increases property tax compliance. This finding adds important context to the more pessimistic previous results on the effects of messages highlighting the link between tax compliance and public service provision in a more abstract way. These insights are relevant for policymakers who weigh the costs and benefits of different policies aimed at raising tax compliance.

[18]The role of reciprocity has also been studied in theoretical work. For instance, Besley (2020) proposes a model linking tax compliance to reciprocal motivations.

[19]E.g., Blumenthal et al. (2001); Castro and Scartascini (2015); De Neve et al. (2021) and Bergolo et al. (2023) find no effects of messages emphasizing fiscal exchange. Hallsworth et al. (2017) and Bott et al. (2020) find positive effects on the timing of payments and income declared in the UK and Norway, the latter pointing out favorable conditions of high trust in the government and high initial tax compliance as potential drivers of the effects. Hernandez et al. (2017) even finds negative effects on tax compliance in Poland. Antinyan and Asatryan (2020) conducts a meta-analysis of more than 40 interventions and concludes that non-deterrence tax messages, including fiscal exchange messages, are on average ineffective.

[20]One exception is Collin et al. (2022), which finds positive and economically significant effects of reciprocity messages in Tanzania. This may indicate a more favorable environment for fiscal exchange messages, or could be related to the fact that around 80% of treated had not received a tax bill prior to the data collection.

Finally, I add to our understanding of the drivers of perceived fiscal exchange by disentangling the different components of the treatment message. Evidence from lab (Lamberton et al., 2018) and survey experiments (Sjoberg et al., 2019; Abbiati et al., 2020) suggests that individuals are more willing to pay taxes when asked about their preferences on how to use tax revenue. Similarly, lab experiments show that allowing individuals to influence the allocation of revenue to services they prefer over others increases tax contributions (Alm et al., 1993; Casal et al., 2016). Yet, evidence from the field confirming these findings is scarce (Khan et al., 2022; Giaccobasso et al., 2023). I show that individuals who depend substantially on public provision of services indeed react to information on services aligned with their elicited preferences. However, these effects are largely driven by information on services which align with the more general needs of the population (water) as well as by accessibility of these services. While the former suggests that the elicitation of individual preferences may be of secondary order as long as the government can proxy citizen needs on an aggregate level, the latter shows that characteristics of public service provision which have so far not been incorporated in lab experiments are crucial indicators of the willingness to pay taxes in real life settings.

2 Institutional Background

Property tax reform In 2019, the city of Freetown introduced a new property tax system. Prior to this reform, the valuation of properties was done manually and was purely based on the floor area of a property. Whereas this method was simple, it failed to account for important property characteristics such as the location and quality of a building, and as such was both inaccurate and regressive (Jibao, 2017). The new 'points-based method' combines information on the surface area with additional observable characteristics of the property and surrounding area, assigning a standard number of points for the former and adding or deducting points for positive or negative property features (Fish, 2018). Using a model informed by the characteristics and rental value of a subset of properties, this information is then translated into an estimated property value (Grieco et al., 2019). Importantly, the assessed value of a property also depends on access to public services, such as water, drainages, and streets. Thus, individuals with better access to public services have on average higher tax liabilities.

The development of the new property assessment scheme went hand in hand with the registration of all qualifying buildings into the property registry.[21] Buildings were identified

[21]This does not include non-permanent and uncompleted properties and illegally constructed properties. Furthermore, public buildings such as clinics, schools, universities, and places used for religious worship are

using satellite imagery and were then visited by enumerators who geo-referenced them and collected the data on observable characteristics necessary to estimate the property value. During this process, more than 50,000 new properties were registered, increasing the tax base to roughly 110,000 taxable properties throughout the city (Kamara et al., 2020).

Table 1: Change in tax liabilities by property value quintiles (in USD)

Average tax payable	Existing system	New system	Average change
1st Quintile	\$14.33	\$4.31	-70%
2nd Quintile	\$15.85	\$9.48	-40%
3rd Quintile	\$16.10	\$17.40	+8%
4th Quintile	\$23.38	\$36.94	+58%
5th Quintile	\$41.64	\$142.25	+242%

Source: Table adapted from Kamara et al. (2020)

The introduction of the new property tax system had considerable effects on the tax liabilities of property owners. Whereas the tax rate for properties in the first two quintiles of average tax payable decreased by more than half, tax liability more than tripled for the top 20% of properties (see Table 1). However, despite increased transparency and accuracy of the property assessment, as well as the introduction of penalties for late payers, tax compliance remains low. In 2021, i.e., after the reform, but prior to the intervention examined in this paper, only 22% of property owners paid property taxes. At the same time, property tax is the most important source of income for the municipality. 46% of local government revenue consisted of property taxes in 2021, increasing to 56% in 2022 (see Appendix Figure A1). Calculating potential tax revenue—that is, tax revenue if all annual tax liabilities were paid—property tax would make up around 80% of total revenue.[22] This underlines the importance of property tax and points to the need of testing additional means to increase tax compliance, such as messages related to tax morale.

Payment logistics Property tax payments for a specific tax cycle are made in the same tax year, i.e., property tax payments for the 2022 tax cycle are made in 2022. Between January and March, rate demand notes (RDNs) are issued and delivered to the properties via mail (printed, not online). RDNs specify the current tax liability including arrears, penalties, and credit. They also contain information about the value assessment of the property. Tax payments can be made either directly at the Freetown City Council or at a number of bank

exempted from the property tax.

[22]This includes all tax liabilities from domestic, commercial, and institutional properties and holds all other revenue equal.

branches throughout the city. Payments can be made in installments of any size, which will be credited to the outstanding tax liability.

Individuals who pay before March 31 receive a 5% credit on the amount paid, which is deducted from the tax due in the following year. The deadline for tax payments is September 30. Individuals who exceed this deadline have to pay a penalty of 5%, which is added to the tax payment due in the following year. Despite the theoretical penalty mechanism, enforcement of payments is extremely low. As of now, liabilities and penalties are mostly accumulated over time without actual consequences for the debtor. This is somewhat different for properties with very high value, for which, anecdotally, additional efforts to collect taxes, especially after the payment deadline, are higher. By law, properties can be seized if individuals are non-compliant despite official warnings and penalties. Again, this is not something that regularly happens in practice.

3 Data and Descriptives

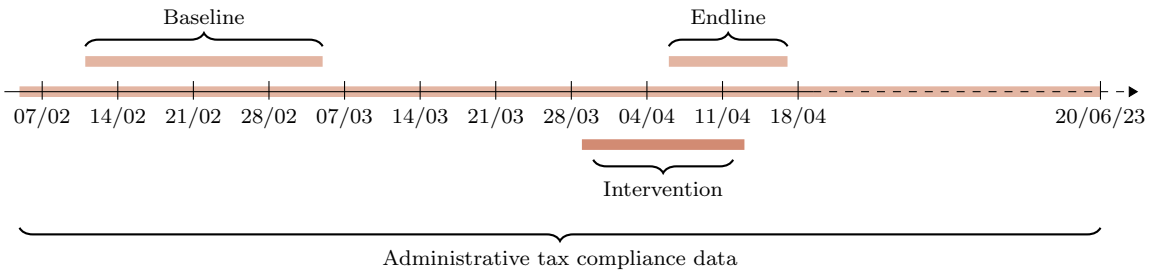
I use both administrative records and survey data. I describe this data briefly here, and in more detail in Appendix A.1. Through an agreement with the Mayor of Freetown and the Freetown City Council (FCC), I had access to the database of registered properties in Freetown. This database was used to sample study participants and track property tax payments over time. To inform the content of the treatment message, I generated a dataset of recently provided public services including geo-referenced locations. Additionally, baseline data was collected from all study participants to elicit preferences for local government-provided services and to ascertain beliefs related to taxation and governance. For a subset of individuals, I collected endline data to explore potential alternative drivers of tax compliance.

The study was carried out between February and April 2022 and was conducted purely over the phone. Figure 1 illustrates the timeline of the data collections as well as the intervention. The baseline survey was conducted between February 11 and March 4, information was provided to treated individuals between March 29 and April 12, and endline interviews started on April 6 and ended on April 16.[23]

Administrative data The database of registered properties contains all properties in Freetown which are legally obliged to pay property tax, including domestic, commercial, and

[23]Whereas the timing of the intervention and the endline data collection overlap overall, there are on average 9 days between the intervention and the endline survey for treated individuals, and 96% of endline interviews were conducted between 8 and 12 days after the treatment took place. See Appendix Figure A2 for the exact distribution of days between intervention and endline survey in the treatment group. Note that the endline data collection was organized in a way that aimed at collecting equal shares of control group and treatment group data every day to ensure comparability.

Figure 1: Timeline of data collection (2022)



institutional buildings. In January 2022, the registry counted 103,407 properties. Besides information about the value of properties and the associated property tax rate, the database contains information about the exact location of the property (address and geo-referenced location) as well as contact information (phone numbers) of the owner or another individual associated with the property, e.g., a tenant, a family member of the owner, or a caretaker.[24] The registry records every tax payment made over the course of the year together with the date on which the payment was made. I have access to this data for the 2021 and 2022 tax cycles as well as part of the 2023 tax cycle, including payments until June 20, 2023.

I restrict the set of properties to properties used for domestic purposes, which are owned by a single-property owner, and for which at least one phone number is available. Multiple-properties ownership is used as a proxy for wealthier households. Since these households generally rely less on *public* service provision and instead tend to provide services *privately*, they are expected to be affected less by information about publicly provided services. To concentrate efforts on the population which I anticipate to respond most to the treatment, these properties are excluded. This results in a set of 46,844 eligible properties, out of which 15,000 were randomly drawn to build the sample frame.[25] The sample frame exceeded the planned amount of 7,000 interviews to account for potential non-response.

Geo-mapping of public services To provide accurate and targeted information about recent public service provision, I used a publicly available information brochure issued by the Freetown City Council, which lists public services provided between 2019 and 2021 for each

[24]The desktop version of the property registry also contains detailed information about the characteristics of each property used to assess the property value. Unfortunately, I do not have access to this data in an aggregated manner.

[25]Note that in few cases, several properties use the same phone number without necessarily being owned by the same person. To avoid sampling the same phone number twice, properties were redrawn until no phone number appeared for more than one property. Appendix Table A1 shows that sampled and non-sampled properties are well-balanced with respect to the few characteristics available for all properties: there are no differences in property value and only minor differences in terms of property locations. Finally, selected properties are slightly more likely to be used for both domestic, and commercial or institutional purposes.

ward of Freetown.[26] Using this brochure, enumerators were sent to locate the mentioned services to (a) confirm their existence and (b) save their exact geo-referenced location.[27] In total, 277 services were geo-referenced. This list includes all services provided and listed by the FCC for which existence could be confirmed by the enumerators.

Based on the services provided by the local government, I identified eight key areas of municipality level public service action: (1) access to water, (2) disaster management and prevention, (3) environmental management, (4) sanitation, (5) the improvement of market infrastructure, (6) health care, (7) education, and (8) road maintenance. Table 2 gives an overview of these categories as well as a few examples of the specific services provided by the FCC for each category. Column (3) specifies how many services were provided between 2019 and 2021 in each category.

Table 2: Public services provided by the FCC

Category	**Examples**	**#**
Access to water	Installation of boreholes, water pumps, or milla tanks	137
Disaster management and prevention	Clearing or construction of drainages, or reconstruction support after floods and fires	53
Environmental management	Tree planting, or the construction of pedestrian crossings and street steps	32
Sanitation	Construction of public toilets, or the cleaning of illegal dumpsites	23
Market infrastructure	Repairs or improvements of market infrastructure such as roofs and floors	14
Health care	Infrastructural improvements of peripheral health units (PHUs) and hospitals	7
Education	Provision of furniture to schools, or the construction of early learning centers	6
Road maintenance	Road repairs, or the pavement of roads	5

Notes: This table provides an overview of services with a fixed location delivered by the Freetown City Council between 2019 and 2021. Column (1) illustrates the different categories of services provided. Column (2) gives specific examples of what has been provided. Column (3) shows how many services have been provided in each category. The table excludes services with no fixed location and services for which existence could not be verified.

Figure 2 illustrates the geographical distribution of services provided by the Freetown City Council. While this map does not reflect overall public service provision in Freetown, it can be seen as a proxy for the FCC's current policy priorities in the area of service provision.

[26]The Transform Freetown Third Year Report can be found here. Information about all service provided between 2019 and 2021 can be found on pages 84 to 105.

[27]I excluded services with no fixed location, such as "the set up of youth tricycle groups to collect garbage" from the geo-referencing.

Approximately 50% of all recorded projects fall in the category of water provision. As shown by the red squares, water projects are widely scattered throughout Freetown. They include the provision of water tanks or street taps, as well as the building of water wells. With roughly 19%, projects related to disaster management and prevention rank second. These types of projects, shown in dark blue, are more densely provided in the central parts of the city, and less so in the Western and Eastern area. Roughly 12% of all projects are related to environmental management. These services, marked in turquoise, are predominantly located at the Eastern part of the city center and in the Eastern outskirts of the city and include, among others, the clearing and building of drainage systems. With slightly above 8% and around 5%, sanitation and market infrastructure projects rank forth and fifth in the agenda of FCC service provision. Finally, health care (2.5%), education (2%) and road maintenance (2%) are least represented in the provision of services by the FCC. The latter is at least partly due to the fact that the construction and maintenance of roads officially falls under the mandate of the federal rather than the local government.[28] Whereas this is also true for the provision of access to water—especially in terms of connecting neighborhoods and individual houses to the piping system—the FCC has made water a priority in the absence of sufficient action by the national government.

Whereas many of the located services seemed to be well-functioning, others were rather short-lived. For instance, many of the drainages that had been cleared by the FCC at some point were again flooded with trash at the time of the geo-coding procedure. Overall, enumerators marked approximately 4% of all registered services as not currently functioning.[29]

Survey data During the baseline data collection, enumerators successfully conducted phone interviews with a total of 5,827 individuals.[30] Interviews were held with respondents who confirmed to live in or own the respective properties.[31] Given that effects on property tax payments were expected predominantly through an interaction with property owners, the latter were oversampled.[32]

Participants were asked about their attitudes towards taxation and public service provi-

[28]Note that the few road *construction* projects commissioned by the FCC were dropped from the list of services and only projects related to road *maintenance* were kept. This decision was taken to align the information provided to the treatment group (see Section 4.1) more carefully with the FCC's mandate.

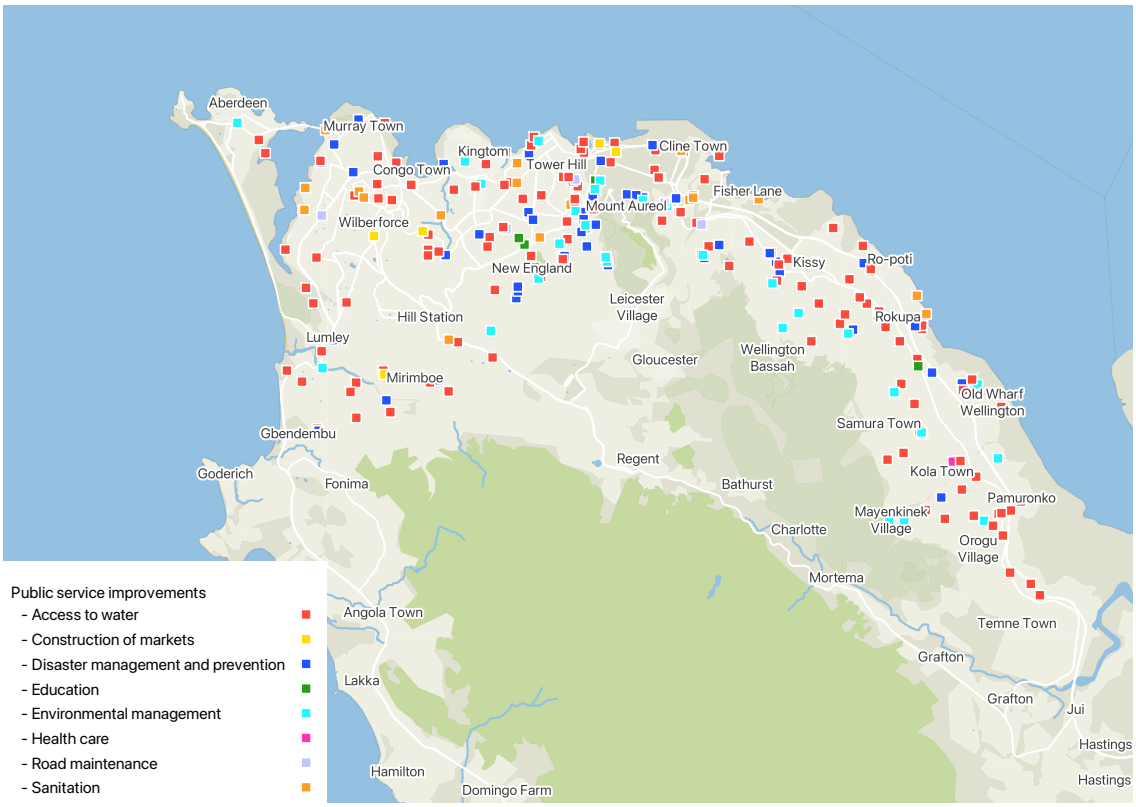
[29]I did not exclude these services from the list of services I provided information about. However, only 2.5% of all treated individuals ended up receiving information about a service considered to be non-functional.

[30]Due to uncertainties concerning the funding at this stage, it was not feasible to continue the data collection until the envisioned 7,000 individuals were reached.

[31]In total, 0.5% of respondents stated to own, but not to live in the respective property.

[32]This was done by splitting the data collection process into two batches and interviewing only owners in the second batch, while interviewing respondents without further selection in the first batch. The exact procedure is described in Appendix A.1.

Figure 2: Locations of services provided by the local government



Notes: The figure illustrates the locations at which public services have been provided by the Freetown City Council between 2019 and 2021, color-coded by the type of service provided at each location.

sion by the local government, and, importantly, about their preferences in terms of public service provision.[33] In particular, enumerators asked individuals to choose one out of eight areas of public service provision, which they would most want the FCC to provide services in and one area in which they would least want the FCC to provide services in.[34] The eight categories correspond to the previously identified types of services that had recently been provided by the FCC.

To decrease the likelihood of providing information about services that are irrelevant to the respondents due to their distance, individuals above the 97th percentile of distance to their most or their least preferred type of service were excluded from the study prior to

[33] A subset of individuals was also asked about their attitudes towards the local government more generally as well as their political participation. More information about the selection of these individuals can be found in Appendix A.1.

[34] The exact framing of the survey questions was as follows: 1. In the following, I will read a list of areas of public service provision to you. Imagine the FCC was to provide services in only one of these areas. Which one would you personally prefer? 2. Now think about the same list of public service areas and imagine again that the FCC was to provide services in only one of these areas. Which one would you personally be least interested in?

randomization. Similarly, individuals with missing data on their most or least preferred public service were excluded. This leaves me with a final sample of 5,494 individuals. The sample is representative for the set of eligible properties in terms of property value, but contains a set of properties that is slightly more likely to be used for multiple purposes and different with respect to the geographical distribution of all eligible properties, suggesting selection into being surveyed based on observable characteristics. Balance results can be found in Appendix Table A2. Imbalances are accounted for in the analysis by controlling for both multiple use and the ward in which the property is located.

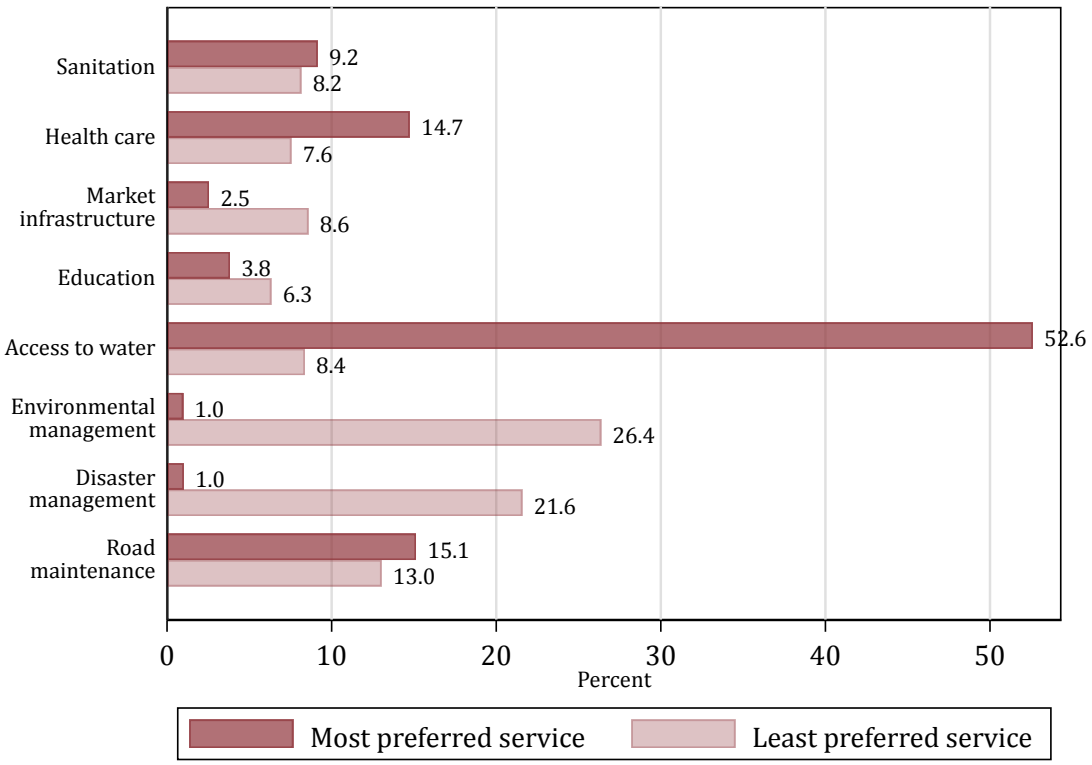
Approximately 8 to 12 days after the intervention, an endline survey including questions about attitudes related to taxation, public service provision and the local government more generally was conducted. I use this data to get a better understanding of the potential mechanisms driving the effects on tax compliance. This survey was administered to a selected subset of the individuals interviewed at baseline among which property owners were not oversampled. As such, the endline sample is not fully representative of the overall study sample, in particular in terms of the share of owners.[35] To account for this, I conduct several adjustments which I explain in detail in Section 7.3.

Public service preferences Figure 3 illustrates the public service preferences reported by the study participants at baseline. Shown in dark red is the percent of individuals who are most interested in the provision of a particular type of public service. The light red bars illustrate the types of services individuals are least interested in seeing provided. More than 50% of individuals declare water to be their top priority. This overlaps almost perfectly with the municipality's priority for providing water as presented in Table 2. Water is followed by a priority for roads and health care with 15.1% and 14.7%. Only 1.0% of individuals have a strong preference for environmental or disaster management, the FCC's second most important categories in terms of services provided. In fact, these two types of services rank first when it comes to least preferred services, followed by roads with 13.0%.

Elicited preferences can capture several layers of preference formation. First, individuals may select a specific type of service as their most preferred type because, all else equal, they value this type of service more than any other type of service. At the same time, when asked for the types of services *to be provided*, individuals' responses are likely informed not only by general preferences, but also by the baseline level of public service provision, or the

[35]The share of owners in the endline sample is 64.4% as compared to 91.3% among those who did not participate in the endline survey. In addition, endline participants are on average somewhat younger, live further away from the closest location at which a service of their most preferred type was provided by the FCC, and are less likely to have a preference for the provision of water. Finally, they live in properties with somewhat lower value. See Appendix Table A3 for more details. The sampling procedure is described in more detail in Appendix A.1.

Figure 3: Public service preferences at baseline



Notes: The figure illustrates the percentage of individuals indicating a specific type of service as their most (dark red) or least (light red) preferred service, respectively. The data used stems from the baseline survey and includes responses from all individuals included in the final sample.

perception thereof. If individuals care a lot about environmental services, but at the same time don't have access to water, they are likely to prioritize the provision of water over the provision of environmental services out of necessity. Thus, the measures of preferences used in this paper likely captures a mixture of general preferences for and perceived need of a particular service.

As I do not directly observe the overall level of services provided in the city of Freetown, I look at several alternative variables to get a better understanding of how stated preferences are formed. I use these variables to predict the likelihood of choosing a specific type of service as one's most or least preferred service in a seemingly unrelated regression framework. Preferences for a specific type of service are uncorrelated with distance to the closest service of that type, as well as with the amount of services recently provided within a 3 kilometer radius (see Appendix Figure A4).[36] However, the lower perceived availability of or satisfaction with the provision of a specific type of service, the more likely individuals are to indicate this

[36]Results are similar when restricting to a 1 kilometer radius or extending to a 5 kilometer radius.

service as their most preferred one (see Appendix Figure A5). This suggests that preferences are indeed formed to a certain extent on the basis of perceived needs. Finally, preferences also vary by property value and property ownership.[37] To account for the fact that preferences correlate with specific characteristics of properties and their residents, I control for the above-mentioned variables in all regressions aiming at disentangling the treatment mechanisms.

4 Experimental Design and Randomization

4.1 Experimental Design

I design and study two interventions disseminating information about public services provided by the Freetown City Council (FCC), a preference *match* and a preference *mismatch* intervention. To do so, I leverage the previously described geo-location of public services as well as the collected information about individual preferences. The information interventions were carried out as a combination of a call and a follow up short message (SMS) sent on the respondent's phone immediately after the call, and aimed at increasing awareness about municipality level public service provision. During the phone call, individuals were informed about one of the eight types of services identified above, including an example location at which this type of service had been provided, as well as the approximate commuting time from the respondent's home.[38]

The type of service mentioned during the call varied by individual depending on their preferences for public service provision elicited at baseline. In the *match* intervention (T_M), individuals received information about their most preferred service, whereas individuals randomized to receive the *mismatch* intervention (T_L) received information about their least preferred service. To target the information treatment geographically, individuals were informed about the closest location at which a specific type of service had recently been provided.

Individuals in the control group (Control or C) did not receive any information call or SMS. This pure control group does not allow me to directly assess whether the effects of my treatment are actually driven by the fiscal exchange content of the message, or whether individuals react to the treatment based on an alternative channel.[39] Thus, to shed light on

[37] In particular, higher property value is associated with a higher likelihood of having a preference for the provision of water. Property owners are significantly more likely to have a preference for roads than individuals who do not own the property they live in.

[38] To provide information on the commuting time, I made use of two datasets. First, I used the geo-locations of properties available in the tax registry. Then, I used the geo-mapped locations at which public services had been provided by the FCC and generated approximate travel times based on the distance between the closest service of each type and the respondent's home.

[39] Note that due to budgetary constraints, it was unfortunately not possible to include an additional control group receiving a non-fiscal exchange message in the study.

the role of potential alternative mechanisms, I present suggestive evidence from heterogeneity analyses as well as analyses of the endline data in Section 7.4.

Appendix Figure A6 depicts the script used for both the *match* and the *mismatch* intervention. The two scripts differ in only one sentence, as can be seen when comparing Figure A6a against Figure A6b. Whereas the *match* group was reminded of their baseline preference in order to make the match more salient, the *mismatch* group was simply informed that some of the FCC's budget went into providing a certain type of service.[40] Immediately after the phone call, enumerators sent an automated SMS to the respondents. This SMS summarized the most important information given during the phone call, i.e., the type and location of a service provided by the FCC, as well as the individual's approximate commuting time.[41]

While the phone call script differs only insubstantially across treatment arms, the content of the treatment message varies depending on individual preferences and the location of the property in which individuals live. Given that the distribution of most and least preferred services differs (see Figure 3), so does the distribution of the types of services individuals are informed about in the match versus mismatch treatment (see Appendix Figure A8).[42] Overall, roughly 30% of individuals were informed about the provision of access to water, the vast majority of which are individuals in the match group. Information about water makes up more than half (53.9%) of the match messages. Road maintenance and health care rank second and third for the match group. Whereas environmental management and disaster management rank last with 0.8% each in the match group, these two categories rank first and second in the mismatch group.

These differences illustrate well that there is non-random variation in the types of services individuals are informed about across treatment groups. Similarly, conditional on being informed about the same type of service, distance to services differs across treatment arms for some services.[43] The implications of this variation for the interpretation of treatment effects are discussed in detail in Section 6.

[40] To avoid negative associations with the phone call, individuals in the mismatch group were not specifically told that they received information about what they previously mentioned to be their least preferred good.

[41] An example of such message is shown in Appendix Figure A7.

[42] Figure A9 in the Appendix shows the share of individuals receiving information about a certain type of service independent of the treatment group assignment.

[43] In particular, individuals in the match group who received information about sanitation are significantly closer to the nearest location at which sanitation services have been provided than individuals in the mismatch group who received information about sanitation. Similarly, individuals in the mismatch group informed about services related to market infrastructure and road maintenance are significantly closer to these services than individuals in the match group informed about the same types of services. There are no significant differences in distance to the respective service between the match and mismatch group for the remaining types of services.

4.2 Randomization

The treatment was randomly assigned on the individual level. Among all 5,494 study participants, one third (1,834) of individuals was randomized into the *control group* (C), one third (1,829) into the *match treatment* (T_M), and one third (1,831) into the *mismatch treatment* (T_L), respectively. The randomization was done using stratification by gender, education, ownership status, and geography, as well as the survey batch and a dummy variable indicating a baseline preference for road maintenance.[44] The latter were included to guarantee balance across treatment groups with respect to service preferences that do not reflect the local government's mandate or priorities. The randomization was successful in generating a well-balanced sample with respect to the main baseline characteristics, with minor imbalances in terms of respondent age (see Appendix Table A4).[45] Importantly, individuals are also well-balanced with respect to baseline preferences and perfectly balanced in terms of distance to the closest location at which services of each type were provided by the FCC (see Appendix Table A6.)

5 Fiscal Exchange: Theory and Hypotheses

Fiscal exchange refers to the concept of a reciprocal relationship between the government and the taxpayer, where individuals are willing "to pay taxes in exchange for benefits that the state provides to them or to others even though their pecuniary payoff would be higher if they didn't pay taxes" (Luttmer and Singhal, 2014, p. 150). The stronger this relationship, the more likely they are thought to pay taxes. But what defines the (perceived) strength of the fiscal social contract and how can public service messages affect it?

A crucial dimension of fiscal exchange is the individual's own perceived benefit of public service provision: specific services can be relevant for some individuals or households, but not for others. Several papers have pointed to the importance of this channel. Evidence from the lab suggests that individuals are less likely to evade taxes when tax revenue is allocated to programs they support (Alm et al., 1993). Similarly, in a field experiment, Giaccobasso et al. (2023) shows that information about the share of property tax revenue used to finance public schools affects the probability of filing a tax appeal differentially for households with

[44] Survey batches refer to whether individuals were in the random respondent or owner batch. Whereas the former were to be interviewed again at endline to measure self-reported outcomes, the latter were included to enlarge the sample used to measure actual tax compliance. More details on the two batches can be found in Appendix A.1.

[45] Table A5 in the Appendix further shows that balance is also given for all characteristics but age and water being one's least preferred service for the subsample of individuals who received an endline survey, which is used to look at self-reported outcomes.

children, who directly benefit from schooling expenditures, and households without children, who do not benefit directly.

These examples illustrate that the extent to which individuals directly (including their household members) benefit from public service provision can affect tax-related outcomes. It also crucially demonstrates that personal benefit can have various dimensions. Individual level preferences as well as the accessibility of services determine how beneficial public service provision is perceived to be. Furthermore, depending on the specific context, certain types of services may be perceived to be more essential than others and can thus play a role independent of stated individual-level preferences for public service provision.

The design of my experiment allows me to look at the role of individual preferences, accessibility, and contextual importance of a specific service separately by investigating how individuals react to the treatment depending on the exact content of the message. Through the match and mismatch treatment, information about public service provision is either aligned or misaligned with personal preferences. Information on accessibility with respect to the location of the service is given through the approximate commuting time between the location of the service and the individual's house. Finally, the type of service individuals are informed about can be considered independently of individual-level preferences as reflective of the more general needs of the population.

In any setting, and in particular in the context of Freetown, access to water can be considered a basic need without which many other public services have less value. For instance, without access to water, health facilities cannot fully operate. In Freetown, water access is still far from universal. Evidence from a pilot study in a subset of wards suggests that around 39% of registered properties have no access to water at all, and only around 45% are connected to the piped water system (Grieco et al., 2019). While this data is not representative for the entire city, it is important to mention that these numbers do not take informal settlements into consideration, where the situation is even more dire.

In response to the under-provision of access to water through the federal government—which is officially responsible for the provision of water, and in particular the piped water system—the Freetown City Council has increased its efforts to provide water to the citizens of Freetown. This is clearly visible when considering the share of services provided with respect to access to water: between 2019 and 2021, 49.5% of all services provided were related to water (see Table 2). At the same time, while many individuals state other preferences, the majority of individuals interviewed for this project considers increased access to water to be the top priority for municipality level public service provision (see Figure 3). Taken together, these facts underline the crucial importance of increasing access to water in Freetown.

In line with this, the main part of this paper focuses on disentangling the role of three

potential indicators of perceived personal benefit. To do so, I look at variation in the information provided to individuals during the treatment call along the following dimensions: (1) the alignment with personal preferences (*match* vs. *mismatch* group), (2) distance of the service, and (3) whether individuals were informed about services providing access to water or any other type of service.

Given that more personal benefit is considered to positively affect individuals' tax morale, I expect individuals to increase tax compliance if they receive information about (a) their most preferred service, (b) a service considered to be accessible with respect to its location (i.e., a service considered to be close), or (c) a service that is unambiguously considered to be important independent of own stated preferences (i.e., water). Tax payments are expected to increase in the share of information aligned with personal benefit. At the same time, it is ex ante unclear whether information that does not suggest personal benefit has no or even negative effects on compliance. This question will be answered empirically in the Section 7.

6 Measurement and Estimation

This article examines the effects of information about public service provision on tax compliance using administrative data. A number of studies look at declarations of income or revenue, which determine the tax base, but not necessarily true compliance (e.g., Bott et al., 2020; Mascagni and Nell, 2022). By contrast, administrative data allows me to capture actual tax payments made. I predominantly look at tax payments made for the 2022 tax cycle between April and September, 2022, i.e., after the intervention and up until the official tax payment deadline for the 2022 tax cycle. To further examine the longevity of the treatment effects, I additionally consider payments made within the entire 2022 tax cycle, i.e., until December 2022. I define three main outcomes: (1) a variable measuring the sum of all payments made by a certain point in time, (2) a dummy variable indicating whether any payment was made, and (3) a variable measuring the sum of all payments made conditional on any payment.

I use OLS to estimate the following equation for the average treatment effect of receiving information about public services provided by the FCC :

$$y_{ik} = \beta_1 T_{ik} + \alpha_k + \mathbf{X}_{ik}\mathbf{\Gamma} + \epsilon_{ik}, \tag{1}$$

where i indexes individuals and k the randomization strata. T_{ik} is an indicator for individuals that received any information treatment. β_1 estimates the average causal effect of assignment to any treatment on the respective outcome of interest, y_{ik}. α_k are strata fixed effects and $\mathbf{X}_{ik}$ are individual-level covariates. In the specifications presented in the main

part of this paper, these include the baseline value of the outcome, y_{ik0}, as well as baseline enumerator fixed effects. In addition, I control for all variables which are either unbalanced across treatment groups or unbalanced with respect to the set of properties the sample is expected to represent. I therefore include respondent age, ward fixed effects, and a variable indicating whether a property is used for commercial or institutional purposes on top of domestic ones. When looking at self-reported endline outcomes, I further include enumerator fixed effects for the endline interview and a dummy variable indicating the timing of the endline data collection (first vs second week). All regressions are estimated using robust standard errors. Finally, I use inverse probability weights to adjust for the fact that the probability of assignment to a treatment group or the control group varies slightly across strata.

To look at the effects of the two treatment arms separately, I estimate

$$y_{ik} = \beta_2 T_{Mik} + \beta_3 T_{Lik} + \alpha_k + \mathbf{X}_{ik}\mathbf{\Gamma} + \epsilon_{ik}, \tag{2}$$

where T_{Mik} and T_{Lik} refer to the *match* and the *mismatch* treatment, respectively, and β_2 and β_3 estimate the effects of being assigned to either of the two treatment groups on the respective outcome of interest, y_{ik}.

Whereas the treatment arms were designed to causally investigate the role of public service preferences, non-random variation in the type of service that individuals were informed about, as well as the distance to that service complicates the interpretation of the results. Thanks to the randomized assignment, individuals are balanced in terms of preferences across treatment groups. At the same time, the share of individuals who mention a specific type of service as their most preferred service differs substantially from the share of individuals mentioning the same service as their least preferred service. Thus, the distribution of services individuals were informed about in the two treatment groups differs, and with it distance. Potential differences in effect sizes across the treatment arms could therefore be driven by differences in the reaction to information on specific types of services or services within a certain range from the individual's property.

To account for this, I do two things. First, I introduce additional control variables which may affect the formation of preferences (see Section 3). Second, I exploit the variation in the treatment message to look at heterogeneity depending on the type of service and distance to the service individuals are informed about. This helps me to separate the role of service preferences from that of distance and service type, and thus allows me to disentangle several potential drivers of perceived personal benefit.

7 Results

7.1 Average Treatment Effects on Tax Compliance

This section presents the effects of providing specific, targeted information about local level public service provision on tax compliance. I begin by illustrating the effect of receiving any type of information about local level public service provision on three main tax compliance outcomes: (1) the sum of all payments made, (2) a dummy indicating whether any payment was made, and (3) the sum of all payments made conditional on payment. All three outcomes are measured at the official tax payment deadline end of September 2022, approximately six months after the intervention.

Table 3: Treatment effects on tax compliance by the payment deadline (September 2022)

	(1) Amount paid (SLE)	(2) Paid any tax (d)	(3) Amount paid (SLE, cond.)
Treated	23.784**	0.012	74.983
	(11.322)	(0.008)	(55.252)
Obs.	5,494	5,494	1,159
Control mean	118.829	0.201	590.605
SD	418.244	0.401	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Column (1) of Table 3 shows that the treatment significantly increases timely tax payments by SLE 23.3 (USD 1.5), an increase of 20% as compared to the control group mean.[46] This suggests that targeted information about public services provided by the institution responsible for collecting taxes can indeed have positive effects on tax compliance in a time frame crucial for the tax authorities. With 3,660 treated individuals, the treatment increased tax payments before the deadline by more than SLE 87,000, or approximately USD 5,600.[47] Extending the intervention to all 46,844 eligible properties could increase tax revenue by more than USD 71,700. With a conservative per-individual cost of SLE 14.2 (including the cost of mapping all services as well as enumerator and phone costs for implementing the intervention), the intervention raises 1.7x its cost. Given that almost 50% of the cost goes back to phone costs and the government often reaches more advantageous agreements with

[46]The effect is robust to winsorizing payments at the 95th percentile of the distribution of the share of tax paid, though somewhat smaller in size (see Appendix Table B7).

[47]Nominal exchange rates for September 30, 2022 taken from https://www.exchangerates.org.uk/USD-SLL-spot-exchange-rates-history-2022.html.

network providers, this figure can be interpreted as a lower bound for the cost-effectiveness of providing public service information at scale.[48]

The effect on overall revenue could be driven by several reactions to the treatment: an increase in the likelihood of any tax being paid (extensive margin), an increase in the amount paid by those who do pay (intensive margin), or a combination of the two. While neither the extensive (Column 2) nor the intensive margin effects (Column 3) are significantly different from zero on their own, both coefficients are positive and sizable. This suggests that the treatment may incentivize some individuals to switch from not paying to paying taxes, while it encourages others to pay more than they otherwise would have paid.[49]

To get a more nuanced understanding of how individuals react to the treatment, I look at two different types of individuals: poor individuals, who reside in low-value properties, and rich individuals, who reside in high-value properties. These individuals vary along several dimensions. By definition, individuals residing in high-value properties face higher tax liabilities in absolute terms. They are also more likely to pay taxes in the absence of treatment.[50] In addition, poor and rich individuals vary in their access to and reliance on public services. As the assessed value of a property depends, among other things, on access to public services, residents of low-value properties have on average lower access to public services. At the same time, the are on average poorer, and as such more budget-constraint. Thus, they depend quite heavily on the provision of public services. Residents of high-value properties, on the other hand, often provide services privately and as such depend less on the government's fulfillment of the social contract.

I look at the variation in treatment effects empirically by splitting my sample into properties with below (low-value) and properties with above (high-value) median value. I present the effects on tax compliance separately for both subsamples. Columns (1) and (2) of Table 4 show that tax payments increase significantly for both below and above median value properties, the effect being significantly larger among the latter. This is not surprising given that these properties have on average tax liabilities that are more than five times as high as those of below median value properties.[51] At the same time, it is not obvious that individuals living in high-value properties, who have on average better access to public services and are more likely to provide services privately, would react to a fiscal exchange message at all.

[48] In previous interventions conducted by the government itself, network providers even allowed free usage of their network.

[49] Note that individuals in the control group who did pay taxes by the end of September paid on average 90.5% of their outstanding tax liability.

[50] This could be driven by the positive relationship between property value and the probability of experiencing enforcement. It could also be explained by individuals residing in low-value properties being more cash-constraint than those living in high-value properties.

[51] This is true both for annual and outstanding (incl. arrears and penalties) tax liabilities.

Table 4: Treatment effects on tax compliance by property value (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) < median	(2) > median	(3) < median	(4) > median	(5) < median	(6) > median
Treated	3.861**	47.718*	0.012	0.009	8.554	51.419
	(1.785)	(24.763)	(0.010)	(0.013)	(7.904)	(96.423)
p-val T S1 = S2	0.054		0.486		0.610	
Obs.	2,747	2,747	2,747	2,747	446	713
Control mean	23.681	209.321	0.153	0.247	154.533	848.112
SD	60.846	566.688	0.360	0.431	62.768	872.514

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. p-values below the treatment effect show whether there are significant differences between the treatment effects in sample 1 (below median property value) and sample 2 (above median property value).

Understanding the dynamics behind these effects thus merits further investigation.

Columns (3) – (6) illustrate the extensive and intensive margin effects separately. Again, while not significant on their own, both effects seem to play a role. The point estimates suggest that the treatment mainly incentivizes residents of lower-value properties to pay taxes. As compared to the control group mean of 15.3%, they are up to 8% more likely to pay taxes. Among above median value properties, who are around 61% more likely to pay taxes in the absence of any treatment, the effect would be less than half that size. As compared to this, the coefficients on the amount paid conditional on any tax paid are similar in relative size, but much larger for high-value properties in absolute terms.

Based on these results, I want to gain a better understanding of the potential mechanisms driving different reactions to the treatment. I start by looking at the role of service preferences measured through the two different treatment arms. To compare the effects of the two treatment arms for individuals with identical preferences, I add service type fixed effects indicating individuals' most and least preferred service. To further account for the fact that preferences are formed endogenously, I control for several variables that are likely to affect the formation of preferences. In particular, I include distance to the closest location at which the most and least preferred services are provided, perceived availability of the most and least preferred service at baseline, property value, and property ownership.

Table 5 presents the effects on tax compliance by treatment arm. Differently than expected, the average treatment effect seems to be driven by individuals who receive information about their least preferred service.[52] The effects on the amount paid, both unconditional and

[52]Appendix Tables B1 to B3 show how the effects change with the gradual inclusion of these control

conditional on payment, are significant and differ from the effects of the match treatment. This has two important implications. First, in this setting, there is no backlash effect of receiving potentially negative information about the provision of a service previously indicated to be one's least preferred service. Second, there has to be another dimension explaining why, on average, individuals react to the mismatch, but not the match treatment.

Table 5: Treatment effects on tax compliance by treatment arm (September 2022)

	(1) Amount paid (SLE)	(2) Paid any tax (d)	(3) Amount paid (SLE, cond.)
Match	7.894	0.012	13.145
	(9.857)	(0.009)	(43.045)
Mismatch	35.413**	0.015	143.640**
	(14.277)	(0.009)	(56.784)
p-val. Match = Mismatch	0.057	0.778	0.011
Obs.	5,469	5,469	1,151
Control mean	118.829	0.201	590.605
SD	418.244	0.401	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

To get a more differentiated view, I once more look at the effects for below and above median value properties separately. Table 6 offers a new perspective, showing that the large and significantly different effects of the mismatch treatment are driven by individuals residing in properties with above median value. Residents of below median value properties, on the other hand, seem in fact to react (more) to the match treatment. These results indicate that the anticipated fiscal exchange channel may indeed be more relevant for individuals who are more dependent on the provision of services by the government. At the same time, the coefficient on the mismatch treatment is still positive and insignificantly different from the coefficient on the match treatment, suggesting that the value of targeting information to individual preferences may be lower than anticipated based on previous research (Alm et al., 1993; Lamberton et al., 2018; Giaccobasso et al., 2023). This finding is likely related

variables. The effects are similar across specifications, with the effects of the mismatch treatment becoming significantly different from those of the match treatment only once controlling for type of service, distance, and perceived availability at baseline. The results are also robust to replacing linear distance to services controls with indicators for type-specific above median distance, distance quartiles, or linear distance to the closest service of each type (see Appendix Tables B4 to B6).

Table 6: Treatment effects on tax compliance by property value and treatment arm (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) < median	(2) > median	(3) < median	(4) > median	(5) < median	(6) > median
Match	5.156**	10.153	0.018	0.004	12.238	-51.891
	(2.263)	(20.812)	(0.012)	(0.015)	(9.271)	(85.179)
Mismatch	3.379	70.854**	0.012	0.013	0.625	164.088*
	(2.139)	(29.996)	(0.012)	(0.015)	(11.049)	(89.828)
p-val. Match = Mismatch	0.465	0.038	0.661	0.578	0.334	0.015
p-val. Match S1 = S2	0.290		0.666		0.808	
p-val. Mismatch S1 = S2	0.046		0.456		0.299	
Obs.	2,736	2,733	2,736	2,733	444	707
Control mean	23.681	209.321	0.153	0.247	154.533	848.112
SD	60.846	566.688	0.360	0.431	62.768	872.514

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

to context-specific factors such as the overall level of public service provision and the related absolute valuation of services provided, independent of their ranking. Importantly, in low-capacity settings, where public service provision is generally insufficient, specific information about any type of service, independent of preferences, may be considered valuable and incentivize a positive reaction towards the state.

These results improve our understanding of the role of different types of information taking both preferences and the allocation of public services within a setting as given. However, individuals receive—by definition of the preference matching—information about different types of services which are located differently close to their homes. These factors can be indicative of how much individuals benefit from a specific service independently, and they are likely to interact with the effect of matching preferences. To investigate whether that is indeed the case, the following section looks at the role of distance and type of service, and their interaction with public service preferences, in more detail.

7.2 The Components of the Treatment Message

To disentangle the different dimensions along which perceived personal benefit of public service provision can affect tax compliance, I look at three components of the treatment message: alignment with individual preferences, distance, and type of service. As before, I look at the effects of the *match* and the *mismatch* treatment to investigate the role of service preferences. To examine the role of information about the provision of access to water, I introduce two dummy variables indicating whether water is an individual's most or least preferred service, respectively.

Finally, to assess the role of differences in access to services, I introduce an indicator for distance to services. By virtue of the intervention's design, treated individuals received information about services with varying distance to their homes. Whereas at the minimum, individuals live less than 10 meters away from the treatment service, at the maximum this distance reaches almost 9 kilometers.[53] This suggests that, despite the intention to provide geographically relevant information to everyone, accessibility of these services differs largely across individuals—potentially affecting the extent to which information about service investments positively affects tax compliance.

To account for this, I define binary variables for distance to the most and least preferred type of service using below and above median distance to the respective service. Distance to services not only depends on the property location and the local government's prioritization

[53]Figure B1 in the Appendix shows the distribution of distance to the treatment service. By definition, the graph is based on data from individuals in the treatment group only, since individuals in the control group did not receive any information about service provision.

for public service investments, but also inherently on the type of service. To take this into consideration, I define median distance using service-specific distances rather than the overall distribution of distance to the treatment service independent of its type.[54]

To look at the role of all three dimensions of perceived personal benefit, I interact the match treatment with the indicator for above median distance to the most preferred service and the indicator for water being the most preferred service. Similarly, I interact the mismatch treatment with the indicator for above median distance to the least preferred service and the indicator for water being the least preferred service. In addition, as in the previous specifications, I include service preference fixed effects, linear distance to the most and least preferred service, perceived service availability at baseline, property value, and ownership status as controls.

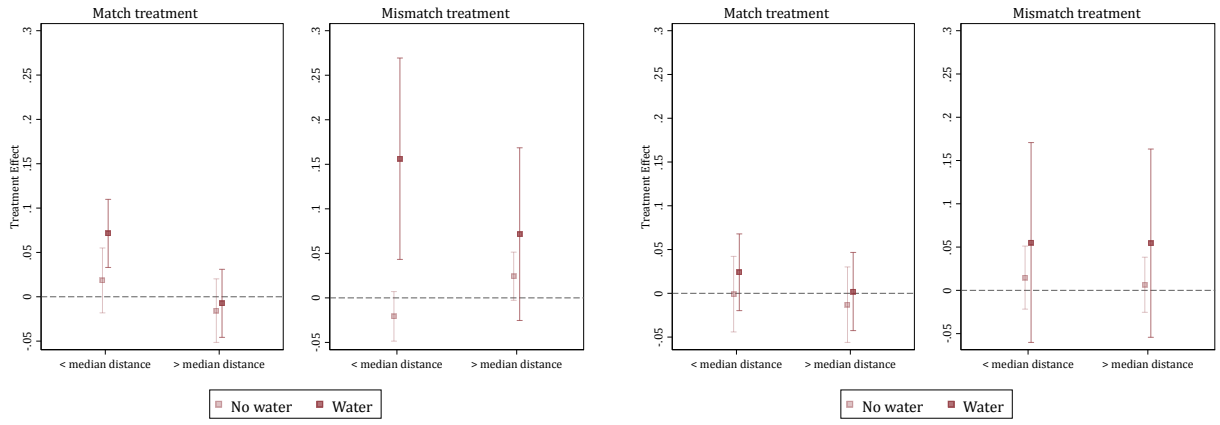
Figure 4 illustrates the interacted effects on whether any tax was paid by end of September for below and above median value properties graphically. The markers show treatment effects with 90% confidence intervals for each treatment arm (match vs. mismatch), differentiating by service accessibility ($<$ vs. $>$ median distance) and type of service (water vs. other type). While there are no significant differences in effect sizes depending on the content of the message for above median value properties, an interesting picture emerges for below median value properties. In line with the hypotheses formulated in Section 5, the point estimates are larger when individuals are informed about water as compared to any other type of service. This is true both for the match and the mismatch treatment. The point estimates are also larger in three out of four cases when information about a service at below median distance is provided.[55]

Overall, the likelihood of paying taxes among lower-value properties is affected significantly only when the treatment message informs about water at below median distance. When such information is provided to individuals with a preference for the provision of access to water, the likelihood of paying taxes increases by 7.1 percentage points or 62.1% as compared to individuals with a preference for water and below median distance to water in the control group, of which 11.5% paid taxes by the deadline. For individuals who receive the information about water at below median distance, but consider access to water their lowest

[54]Specifically, I look at the sample distribution of distance to the closest location at which a specific type of service is provided. I then define the median distance for each type of service. This allows me to categorize the minimum distance between each property and each type of service into below or above the median. Figure B2 in the Appendix illustrates the distribution of distance for each type of service. Depending on the type of service an individual received information about, I replace the indicator variable for below or above median distance to the treatment service with the property-specific indicator for that type of service. I then run regressions excluding either treated individuals with below or above median distance to the treatment service.

[55]Appendix Figure B4 shows that these patterns do not extend to the intensive margin, i.e., to effects on the amount of tax paid conditional on payment.

Figure 4: Heterogeneity of the treatment effect on whether any tax was paid depending on the content of the treatment message (September 2022)



(a) Below median property value

(b) Above median property value

Notes: The figure illustrates the effect of being treated depending on whether an individual received the match or the mismatch treatment, whether thy received information about water or another type of service, and whether the respective service has below or above median distance. Figures 4a and 4b show effects separately for low- and high-value properties. Coefficients are shown with 90% confidence intervals. All regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership. Appendix Figure B4 illustrates the effects for the entire sample.

priority, the effect amounts to 15.6 percentage points or 134.4% as compared to their control group counterpart, which has a compliance rate of 11.6%.[56] Both effects are significantly larger than the respective effects of the match or mismatch treatment when information about another type of service at below or above median distance is provided. For the match treatment, the effect is also significantly larger when information about water at below as compared to above median distance is provided.

While the effect appears to be larger for individuals in the mismatch group, the difference in coefficients between the match and mismatch group is not statistically significant. The lack of precision in the estimated effect of information about water misaligned with preferences is likely related to the small number of individuals who state water as their lowest preference, have below median distance to water, and are part of the mismatch treatment arm. Among below median value properties, only 34 individuals in the mismatch group receive information about water at below median distance, whereas this number amounts to 226 individuals in

[56]The effects on individuals informed about water at below median distance, which seem to be driven by below median value properties, are visible, although somewhat less pronounced, when looking at the entire sample as well (see Appendix Figure B3).

the match group.

A potential explanation for differences in the size of treatment effects across treatment arms could be that individuals are differentially informed about the existence of these services prior to the intervention. Differences in the level of prior information may create differences in the perceived value of the provided information, with less informed individuals perceiving the quantity of information received to be higher. Combined with the overall value of information about water with below median distance, this could explain why individuals in the mismatch group show a potentially larger reaction to the treatment than individuals in the match group. Comparing treated individuals who report water to be their least preferred service to those who indicate water to be their most preferred service indeed indicates that the former are somewhat less likely to have known about the service before the treatment call. However, these differences are—likely due to the comparably small number of observations—not significant.

These findings show that specific information about public services provided by the local government has the potential to substantially increase tax compliance. Crucially, the results underline the importance of distance to the locations at which services are provided, i.e., the importance of service accessibility (on a geographical level). They also show that information on certain types of services is valued more than others, even if these services do not match stated individual preferences. These results underline the importance of considering perceived personal benefit of public service provision along multiple dimensions.

The extensive margin effects associated with high personal benefit are driven by below median value properties. Residents of these properties have on average lower access to public services and significantly lower monthly income. This suggests that individuals living in lower-value properties rely on average more on services provided by the local government. It is therefore to be expected that these individuals are particularly sensitive to information on services they personally benefit from. As opposed to this, individuals residing in higher value properties do not respond differentially to the specific content of the treatment message, neither on the extensive nor the intensive margin. For these properties, the observed effects seem to be driven by a more general fiscal exchange channel, with specific information about any public service independent of personal benefit signaling the government's capacity and willingness to fulfill the social contract. I investigate this channel in the following subsection.

Robustness of Main Results The results for the average effect of being treated, as well as the message-specific effects on whether any tax payment was made are robust to a variety of alternative specifications, including specifications controlling for respondent baseline characteristics, specifications controlling for potential correlates of individual level preferences,

and specifications without weight-adjustments for the probability of being treated, as well as specifications excluding control variables (see Appendix Tables B7 and B8). Importantly, the results are also robust to winsorizing the two variables indicating the sum of conditional and unconditional payments made at the 95th percentile of the distribution of the share of outstanding liability paid. Finally, the results also hold true when adjusting for the oversampling of property owners.[57]This finding is of particular importance for a potential scale up, as it proxies the effect sizes to be expected when information is given to whomever is reached over the phone, independent of property ownership.

7.3 Perceived Government Capacity and Service Provision

Perceived fiscal exchange may not only be determined by the extent to which individuals personally benefit from public service provision. Receiving information about the government's efforts to increase service delivery may also improve individuals' attitudes towards the government—including perceptions of its capacity to provide public services and fairness of tax rates—and through this the willingness to pay taxes (e.g., Hofmann et al., 2008; Slemrod, 2007). The treatment may also induce individuals to update their beliefs about the provision of public services more generally, increasing the perceived availability of or satisfaction with public service provision. In line with this hypothesis, previous studies have found positive correlations between satisfaction with public services and willingness to pay taxes or the belief in the citizens' obligation to pay taxes (e.g., Van den Boogaard et al., 2020; OECD, 2019; Bodea and LeBas, 2016; Hanousek and Palda, 2004). I investigate these mechanisms using self-reported data on perceptions related to service delivery and local government performance more generally collected during baseline and endline interviews. By design, I can link all self-reported information to administrative tax outcomes and thus examine directly how these dimensions interact with the treatment effect and thus tax compliance.

Variables of Interest To investigate the treatment's effect on attitudes towards the government, I examine citizens' beliefs about the performance of the local government, as well as its responsiveness. I follow Weigel (2020) in the definition of the first outcome. *Performance of local government* is an index reflecting citizens' trust in and approval of the local government more generally (Besley and Persson, 2009).[58] *Local government responsiveness*

[57]In particular, this specification uses probability weights adjusting the weight of each property owner in the overall sample to reflect the share of owners achieved in the first part of the baseline survey when owners and tenants were interviewed randomly depending on who picked up the phone.

[58]I construct all indices by standardizing each survey question and taking the unweighted average of all index components. I use these indices as a means to reduce type I and II errors stemming from multiple hypothesis testing.

is a variable increasing in the perceived responsiveness of the local government to its citizens' needs.

Second, to better understand the role of perceived public service provision, I look at a variety of indices reflecting individuals' perceived availability of and satisfaction with public services provided by the local government:

(i) *Perceived availability of public services*: an index combining questions on perceived availability of each of the eight types of services (sanitation, health care, the improvement of market infrastructure, education, access to water, environmental management, disaster management and prevention, and road construction and maintenance) on 3-point Likert scales. These questions refer to the amount of provided services independent of quality.

(ii) *Perceived availability of most preferred service* and *Availability of least preferred service*: standardized indicators reflecting survey answers to the questions on perceived availability of the individual's most and least preferred service, respectively.

(iii) *Satisfaction with public services*: an index combining questions on satisfaction with each of the eight types of services on 5-point Likert scales. These questions refer to the quality and adequacy of the amount of services provided.

(iv) *Satisfaction with most preferred service* and *Satisfaction with least preferred service*: standardized indicators reflecting survey answers to the questions on satisfaction with the individual's most and least preferred service, respectively.

In addition, I look at heterogeneous treatment effects by perceived availability of public services at baseline. Baseline perceptions of public service provision by the local government are very negative. 44.2% of individuals report that they think the local government provides no services at all in accordance with their most preferred type, and only 3.8% report that there are a lot of services aligned with their priority. These numbers are similar, although a bit more positive for the least preferred service, with 37.6% of individuals perceiving there to be no services and 10.8% perceiving there to be lots of services aligned with their least preferred service type. To investigate how these attitudes affect the response to the treatment message, I interact a dummy variable indicating whether perceived availability of the most preferred service at baseline is low (no services) or high (a little or a lot of services) with the match treatment. Similarly, I interact the mismatch treatment with an indicator for high versus low perceived availability of the least preferred service.

Estimation As the endline survey was administered to a subset of individuals characterized by a larger share of tenants than the overall sample, and there is selective attrition between the baseline and endline sample, the self-reported endline results may not be fully representative for the overall sample.[59] Taking this into consideration, I show two different specifications for each of the self-reported outcomes of interest. In each table, I first present the results of the respective regressions for the endline subsample without any adjustments (Column 1). These results can be cautionary interpreted as average effects for a randomly selected sample of owners and tenants. Second, I use inverse probability weights adjusting for the likelihood of being interviewed at endline conditional on ownership status and ward—the two main characteristics differing between the endline and the overall sample (Column 2). Assuming that these two characteristics predict selection into the endline sample, this should adjust the regressions to make them representative for the entire sample considered in the tax compliance analysis.[60] In each table, every row presents results from a separate regression and shows the effects of being treated on the respective index or index component.

Perceived Government Performance I present the treatment effects on perceived government performance in Table 7. Column (1) shows that the effects of receiving information about public services on perceived performance of the local government are positive but insignificant, both when looking at the overall index, as well as at the five variables that feed into the index. Similarly, there are no effects of the treatment on perceived responsiveness of the government. These results are largely consistent with the specification using inverse probability weights in Column (2). In this specification, the effect on confidence in the Freetown City Council (FCC) is positive and significant, suggesting that specific, verifiable information about public services provided can indeed lead to increased trust in the local government. There are no effects on perceived government performance among properties with above median value, indicating that this channel cannot explain the positive results on payments made for these properties.[61] While the signs on all coefficients suggest a weak positive link between the treatment and perceived government performance, a message mentioning a

[59] More information on the measurement challenges related to the endline sample can be found in Appendix B.

[60] In addition, I use Horowitz-Manski bounds to look at the range of potential treatment effects in case the individuals who did not respond at endline are those with extreme outcomes. That is, for each individual who did not respond at endline, I replace the missing outcome value with the maximum value that this outcome takes for treated and the minimum for untreated individuals to calculate the upper bound of the treatment effect and vice versa for the lower bound. The results of these regressions are reported in Appendix Tables B12, B13, B14, B15, and B19. While these bounds can be useful to determine the range of potential effects, they are based on extreme cases with respect to selective attrition and are as such often very broad. Thus, they do not necessarily facilitate to draw conclusions about actual effect sizes.

[61] See Appendix Tables B13 and B14 for the effects on below and above median value properties.

single example of public service provision does not seem to be enough to significantly shift attitudes towards the government overall.

Table 7: Treatment effects on perceived government performance

	Original	Inverse prob. weights
	(1)	(2)
Performance of local government index	0.051	0.051
	(0.036)	(0.038)
	[2,684]	[2,684]
- Confidence in ward councilor	0.032	0.024
	(0.036)	(0.039)
	[2,684]	[2,684]
- Confidence in mayor	0.025	0.019
	(0.037)	(0.040)
	[2,684]	[2,684]
- Confidence in FCC	0.050	0.066*
	(0.037)	(0.039)
	[2,684]	[2,684]
- Approval of ward councilor	0.050	0.045
	(0.036)	(0.039)
	[2,684]	[2,684]
- Approval of mayor	0.028	0.029
	(0.035)	(0.037)
	[2,684]	[2,684]
Local government responsiveness (std.)	0.043	0.036
	(0.033)	(0.035)
	[2,610]	[2,610]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Appendix Table B12 presents results including Horowitz-Manski bounds.

Perceived Availability of and Satisfaction With Public Service Provision Table 8 presents the results of treatment effects on perceived availability of and satisfaction with local level public service provision. The results show that, while the information treatment does not affect perceived availability of services, it significantly increases satisfaction with the provision of services of one's most and least preferred service, as well as service provision more generally.

Table 8: Treatment effects on perceived availability of and satisfaction with public services

	Original	Inverse prob. weights
	(1)	(2)
Availability of public services index	0.034	0.034
	(0.029)	(0.031)
	[2,684]	[2,684]
Availability of most preferred service (std.)	0.037	0.054
	(0.038)	(0.040)
	[2,677]	[2,677]
Availability of least preferred service (std.)	-0.005	-0.005
	(0.038)	(0.040)
	[2,665]	[2,665]
Satisfaction with public services index	0.058**	0.057*
	(0.029)	(0.031)
	[2,683]	[2,683]
Satisfaction with most preferred service (std.)	0.065*	0.090**
	(0.037)	(0.040)
	[2,676]	[2,676]
Satisfaction with least preferred service (std.)	0.071*	0.064*
	(0.037)	(0.039)
	[2,669]	[2,669]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Appendix Table B15 presents results including Horowitz-Manski bounds.

As expected given the content of the respective treatment messages, Table 9 shows a positive effect on satisfaction with the most preferred service in the match group, and a positive effect on satisfaction with the least preferred service in the mismatch group. However, the coefficients are not significantly different across treatment arms. The treatment also significantly increases satisfaction with service provision overall, measured by an index including all eight types of services provided by the local government. This suggests that the treatment not only affects satisfaction with the type of service one was informed about, but also generates spillovers on satisfaction with other services—potentially due to increased awareness related to public service provision by the FCC post-treatment.

Interestingly, the effect of the mismatch treatment on satisfaction with one's least pre-

Table 9: Treatment effects on satisfaction with local level public service provision

	(1) Satisfaction with public services index	(2) Satisfaction with most preferred service (std.)	(3) Satisfaction with least preferred service (std.)
Panel A. Overall treatment effect			
Treated	0.066**	0.062*	0.078**
	(0.028)	(0.036)	(0.035)
Panel B. Treatment effect by arm			
Match	0.058*	0.080*	0.071*
	(0.033)	(0.042)	(0.041)
Mismatch	0.084**	0.059	0.086**
	(0.034)	(0.042)	(0.042)
p-val. Match = Mismatch	0.430	0.602	0.710
Obs.	2,664	2,660	2,659
Control mean	-0.000	-0.000	0.000
SD	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

ferred service is entirely driven by residents of above median value properties.[62] Thus, the large and significant effect of the mismatch treatment on the amount of tax paid by residents of above median value properties could be driven by an increase in satisfaction with these services. This finding is in line with the hypothesis that individuals in the mismatch treatment are ex ante less informed about the treatment service, so that the treatment affects priors, and as such tax compliance, only within the mismatch group.

Finally, I also look at the role of perceived availability of services at baseline. Whereas perceived availability correlates with elicited preferences, it does not seem to affect the reaction to the treatment (see Appendix Table B18). This result should, however, be interpreted with caution, as comparably little variation in baseline perceived availability and individuals' generally low levels of perceived availability may prevent the identification of interaction effects along this dimension.

These results suggest that giving examples of one specific service provided by the local government does not shift perceptions about the overall availability of this or any other type of service. However, it increases overall satisfaction with the provision of services by the local government. While part of this effect may be driven by new information about the

[62] See Appendix Tables B16 and B17 for effects on satisfaction by property value.

existence of a service as argued above, the treatment may also increase the salience of who provides certain services. Attributability could thus play an important role in strengthening the fiscal social contract. The salience of politicians' performance with respect to public service provision can also be a crucial predictor of re-election chances.[63] The treatment is thus of direct interest to politicians who want to increase the visibility of their policy agenda.

7.4 Alternative Explanations

Tax Awareness Given the nature of the intervention, the treatment could affect tax compliance through several other mechanisms than through the relevance of the information about public services itself. For instance, the intervention may simply raise awareness about an individual's duty to pay taxes and nudge individuals who forgot to do so or delayed payment into paying. Several pieces of evidence speak against this interpretation. First, the intervention mentioned neither the duty to pay property tax nor the payment deadline. The only sentence linked to taxes simply stated that "[t]he FCC has been raising necessary funds [to provide public services] through different sources, ranging from donor funds to property tax, local tax, and business license tax."[64] If anything, this sentence underlines the necessity of raising tax revenue to provide public services and as such emphasizes the fiscal exchange channel.

Second, I show that the treatment effects are driven by individuals who care disproportionally about the provision of public services. In addition to official property taxes, individuals in Freetown often contribute to the provision of community services via informal contributions in cash, kind, or labor. These payments should not be overlooked, as they not only increase the tax burden of households, but also crucially contribute to the development of cities and countries (Olken and Singhal, 2011). At baseline, 85% of individuals reported that at least one of their household members made such a contribution in the previous year. Looking at the heterogeneity of effects by whether an informal tax payment was made or not reveals that the treatment effects are driven by individuals in whose household an informal contribution was made (see Appendix Table B20). These results suggest that individuals who value the provision of public services more are more likely to react to the treatment, underlining once again the importance of the fiscal exchange channel as a driver of the observed effects. Rather than disrupting it, informal tax payments seem to reinforce the social

[63] For instance, Ajzenman and Durante (2023) shows that voters assigned to a ballot station in a school with poor infrastructure are less likely to vote for the incumbent politician than voters assigned to schools with higher quality infrastructure. The paper argues that this effect is driven by increased salience of the incumbent's poor performance with respect to the education sector in poor infrastructure schools.

[64] For individuals in the mismatch group, information about the provided service was further introduced by saying that "part of these funds have been used to improve" the respective service.

contract with the state.

Finally, as has been illustrated at length in Sections 7.1 and 7.2, the content of the treatment message actually matters for whether the treatment does or does not affect tax compliance. If the treatment call simply made individuals aware of their tax duty, we should see a base treatment effect independent of the message's content. Instead I find that effects among lower-value properties are driven entirely by information aligning with perceived personal benefit. Whereas the fiscal exchange mechanism is somewhat less obvious for individuals residing in above median value properties, the fact that the match treatment does not affect tax compliance at all does speak against a pure awareness effect—which should be observed equally across treatment groups. It is thus much more likely that the treatment affected tax compliance through a change in perceived fiscal exchange induced by the information about public service provision than simply by bringing attention to the tax duty independently of the message's content.

Perceived Extractive Capacity A second way of thinking about the treatment effect is through an effect on perceived extractive capacity of the government. Receiving personalized information about public services may induce people to (falsely) believe that the government now possesses more information about citizens that could facilitate the collection of tax revenue and the tracking of evaders. This makes evasion more costly and as such has the potential to increase tax compliance. Following Weigel (2020), I look at the effects of the intervention on perceived extractive capacity of the government using an index that reflects the perceived level of information the government has about taxpayers. In particular, this index increases in whether individuals believe that the Freetown City Council (FCC) knows the location of their house, whether their neighbors paid taxes, their occupation, and their income level.[65]

Table 10 illustrates the effects of being treated on the index (Row 1), as well as the standardized components of the index (Rows 2 - 5). Overall, the treatment does not affect perceived information on taxpayers. However, looking at the components separately shows that treated individuals are more likely to believe that the FCC knows whether their neighbors paid taxes as well as their occupation than individuals in the control group.

These results suggest that treated individuals may in fact, at least to a certain extent, believe the government more able to extract tax revenue than individuals in the control group. While I cannot entirely exclude that the treatment effect on tax compliance is (partly) driven by this increase, the effects on perceived FCC knowledge of one's occupation are not

[65] As before, this index is constructed by standardizing each component of the index and taking the unweighted average of all components.

Table 10: Treatment effects on perceived government information on taxpayers

	Original	Inverse prob. weights
	(1)	(2)
Information on taxpayers index	0.035	0.043
	(0.037)	(0.040)
	[2,683]	[2,683]
HH location	-0.013	0.001
	(0.032)	(0.037)
	[2,683]	[2,683]
Compliance	0.048	0.052*
	(0.029)	(0.030)
	[2,683]	[2,683]
Occupation	0.077**	0.074**
	(0.035)	(0.037)
	[2,683]	[2,683]
Income	-0.038	-0.039
	(0.031)	(0.033)
	[2,683]	[2,683]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Appendix Table B19 presents results including Horowitz-Manski bounds.

driven by the groups of individuals who increase compliance in response to the treatment. Residents of low-value properties who received information about the provision of access to water at below median distance (see Appendix Table B21) and residents of above median value properties, in particular those who received the mismatch treatment, do not react to the treatment by adjusting indicators of perceived extractive capacity (see Appendix Tables B22 and B23). This indicates that perceived extractive capacity does not drive the compliance results related to the information treatment and supports the conclusion that perceived fiscal exchange actually matters for tax compliance.

7.5 Owners Versus Tenants

Depending on who was reached at baseline, information about public services was provided either to the property owner—who is legally obliged to pay property taxes—or a resident without such obligations, such as an official tenant, a family member of the owner, or a

caretaker. Given that the link between the treatment and the payment of taxes is less strong if the latter is the case, one could expect the treatment effects to be driven entirely by property owners receiving information.

However, although tenants are not legally obliged to pay property tax, they are allowed do so on behalf of the owner once the payment deadline has passed. Payments made by the tenant can be deducted from the rent due.[66] More than that, the FCC has the right to seize the property, including the tenant's movable property, if no payment is made after receiving a warning letter. While this is unlikely to actually happen, it may give tenants reason to negotiate about tax payments with their landlord, especially upon receiving the public service treatment.

Table 11 shows that, on average, the treatment indeed affects tax compliance significantly only if the owner received information. However, looking at the two treatment arms separately shows significantly larger tax payments both among properties in which owners and tenants received information, specifically if they received information misaligned with respondent preferences. For owners, this effect may be driven by the information component of the mismatch treatment. As tenants can legally deduct property tax payments made in their landlord's stead from their rent only after the payment deadline, they are unlikely to make payments on their own prior to the deadline. Payments made for properties in which tenants received information are thus probably either made by the owner, or in accordance with the owner. If tenants communicate the provided information to their landlords and payments are made by the latter, the large reaction to the mismatch treatment could also be explained by differences in preferences between tenants and owners, with the latter responding to information about services aligned with their preferences.

7.6 Late Compliers Versus Non-Compliers

So far, this paper has shown that being treated increases tax compliance by the deadline. However, it is unclear whether these effects are driven by individuals who otherwise would have paid later (late compliers) or individuals who would not have paid at all (non-compliers). To shed light on this question, I look at the effects of information about public services on tax compliance throughout the 2022 tax cycle. Table 12 shows that the effects on the amount of tax paid persist over time. At the end of the year, individuals living in below median value properties who received information about their most preferred service pay on average 21% more taxes than residents of below median value properties in the control group. The effect is on average positive, but insignificant among treated individuals in the mismatch group. As

[66] See https://fcc.gov.sl/property-reform-faq/

Table 11: Treatment effects on tax compliance by property ownership (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) Owner	(2) Tenant	(3) Owner	(4) Tenant	(5) Owner	(6) Tenant
Treated	16.717*	34.830	0.008	0.019	96.785**	159.050
	(9.873)	(25.258)	(0.009)	(0.017)	(45.513)	(188.848)
p-val T S1=S2	0.304		0.321		0.472	
Match	8.169	2.101	0.007	0.014	30.286	45.707
	(10.181)	(27.246)	(0.011)	(0.020)	(45.357)	(199.276)
Mismatch	23.212*	68.896*	0.010	0.030	149.210**	253.190
	(14.044)	(41.098)	(0.011)	(0.020)	(61.444)	(225.926)
p-val. Match = Mismatch	0.289	0.132	0.775	0.437	0.029	0.337
p-val. Match S1 = S2	0.725		0.689		0.102	
p-val. Mismatch S1 = S2	0.121		0.205		0.915	
Obs.	4,279	1,190	4,279	1,190	928	223
Control mean	115.167	131.709	0.214	0.155	537.448	848.794
SD	373.943	546.626	0.410	0.363	653.100	1154.713

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

before, the effects are largely driven by individuals who receive information about water at below median distance and decide to pay taxes in response to this information (see Appendix Figure B5). This suggests that the treatment does not only induce these individuals to pay taxes earlier, but actually motivates taxpayers who otherwise would not pay taxes to comply.

Similarly, the positive effect of the mismatch treatment among residents of above median value properties persists over time. By the end of the 2022 tax cycle, these individuals pay on average 29% more taxes—an effect that has substantial influence on overall tax revenue. The increase in payments seems to be driven by effects on both the extensive and intensive margin. However, the effects cannot be identified separately at conventional levels of significance. Overall, this suggests that individuals residing in above median value properties who receive the mismatch treatment pay significantly more taxes than individuals in the control group instead of simply paying (more) earlier.

Finally, it is also worth looking at whether the treatment effect persists beyond the tax year in which the intervention took place. Looking at tax payments for the 2023 tax cycle until June 20, 2023 shows that the treatment does not affect tax payments more than a year after the treatment (see Appendix Table B24). Taken together, these results suggest that light-touch interventions like the one implemented in this paper have important effects on tax compliance in the short- and medium run. However, a one time intervention is not enough to change individual tax compliance behavior in the long run. Whether and how repeated information treatments could shape tax compliance in the long run deserves further investigation.

8 Conclusion

This paper examines the effects of specific, targeted information about public service provision on tax compliance. I find that treated individuals pay on average 20% more taxes on time than individuals in the control group. The effect is driven by increases on both the extensive and intensive margin of tax compliance. Individuals residing in low-value properties, who depend heavily on public service provision, react to information that signals personal benefit. in particular, these individuals increase their likelihood of paying taxes on time by 7–16 percentage points when receiving information about an easily accessible service with unambiguous value for all citizens, namely water. The effect of public service information persists throughout the tax cycle, suggesting that messages emphasizing fiscal exchange have the potential to not only shift the timing of tax payments, but actually increase compliance.

These results emphasize the importance of the underlying level of public service supply, illustrating that interventions showcasing government effort to provide services do not increase

Table 12: Treatment effects on tax compliance by property value (December 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1)	(2)	(3)	(4)	(5)	(6)
	< median	> median	< median	> median	< median	> median
Treated	4.167**	27.533	0.011	0.002	9.641	52.642
	(1.917)	(28.944)	(0.011)	(0.014)	(7.705)	(93.100)
p-val T S1=S2	0.233		0.808		0.994	
Match	5.252**	-8.395	0.015	-0.002	12.367	-15.680
	(2.385)	(28.146)	(0.013)	(0.016)	(9.654)	(99.559)
Mismatch	3.306	73.971*	0.009	0.009	4.283	112.792
	(2.291)	(39.132)	(0.013)	(0.016)	(9.905)	(107.910)
p-val. Match = Mismatch	0.449	0.016	0.663	0.487	0.465	0.143
p-val. Match S1 = S2	0.692		0.994		0.673	
p-val. Mismatch S1 = S2	0.127		0.677		0.719	
Obs.	2,736	2,733	2,736	2,733	468	773
Control mean	25.084	258.728	0.163	0.273	153.596	946.319
SD	62.022	741.296	0.370	0.446	61.771	1167.218

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

tax compliance among those who most depend on public services when services are too far away. Hence, increasing the density of public services, and as such service provision more generally, is a crucial condition for fiscal exchange messages to affect tax compliance among those most in need. By itself, increasing public service provision has been shown to positively affect tax compliance. Informing individuals about the latter can have complementary effects, thus increasing the returns to the provision of public services for local governments.

Residents of high-value properties, on the other hand, increase tax compliance in response to information about their least preferred service. These effects are not driven by the specific content of the message, suggesting that benefiting personally from the service one is informed about may be less relevant for residents of properties with better average access to public services. However, individuals in the mismatch group update their priors about local level public service provision and are ex-post more satisfied with the provision of their least preferred type of service. The effect of the mismatch treatment can thus be interpreted as reaction to a positive signal of government performance with respect to public service provision that is independent of personal benefit.

My results also show that individuals residing in high-value properties drive the effects on overall tax revenue. This is an important finding, suggesting that policy makers with the goal of increasing tax revenue should not only focus on providing information about public services to those most in need of the latter. By contrast, providing information that signals the government's capacity to deliver public services to richer individuals has the potential to generate substantially higher overall returns for the government.

References

Abbiati, L., A. Antinyan, and L. Corazzini (2020). A survey experiment on information, taxpayer preferences, and perceived adequacy of the tax burden. *Heliyon 6*(3).

Ajzenman, N. and R. Durante (2023). Salience and accountability: School infrastructure and last-minute electoral punishment. *The Economic Journal 133*(649), 460–476.

Alm, J., B. R. Jackson, and M. McKee (1993). Fiscal exchange, collective decision institutions, and tax compliance. *Journal of Economic Behavior & Organization 22*(3), 285–303.

Antinyan, A. and Z. Asatryan (2020). Nudging for tax compliance: A meta-analysis. CESifo Working Paper No. 8500.

Balán, P., A. Bergeron, G. Tourek, and J. L. Weigel (2022). Local elites as state capacity: How city chiefs use local information to increase tax compliance in the Democratic Republic of the Congo. *American Economic Review 112*(3), 762–97.

Bergeron, A., G. Z. Tourek, and J. L. Weigel (2023). The state capacity ceiling on tax rates: Evidence from randomized tax abatements in the DRC. NBER Working Paper No. w31685.

Bergolo, M., R. Ceni, G. Cruces, M. Giaccobasso, and R. Perez-Truglia (2023, February). Tax audits as scarecrows: Evidence from a large-scale field experiment. *American Economic Journal: Economic Policy 15*(1), 110–53.

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

Besley, T. and T. Persson (2009). The origins of state capacity: Property rights, taxation, and politics. *American Economic Review 99*(4), 1218–44.

Besley, T. and T. Persson (2013). Taxation and development. In *Handbook of Public Economics*, Volume 5, Amsterdam, pp. 51–110. Elsevier.

Blumenthal, M., C. Christian, and J. Slemrod (2001). Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal 54*(1), 125–138.

Bodea, C. and A. LeBas (2016). The origins of voluntary compliance: Attitudes toward taxation in urban Nigeria. *British Journal of Political Science 46*(1), 215–238.

Bott, K. M., A. W. Cappelen, E. Ø. Sørensen, and B. Tungodden (2020). You've got mail: A randomized field experiment on tax evasion. *Management Science 66*(7), 2801–2819.

Brockmeyer, A., A. Estefan, K. R. Arras, and J. C. S. Serrato (2023). Taxing property in developing countries: Theory and evidence from Mexico. NBER Working Paper No. w28637.

Casal, S., C. Kogler, L. Mittone, and E. Kirchler (2016). Tax compliance depends on voice of taxpayers. *Journal of Economic Psychology 56*, 141–150.

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

Cohen, I. (2023). Technology and the state: Building capacity to tax via text. SSRN Working Paper No. 4523287.

Collin, M. E., V. Di Maro, D. K. Evans, and F. Manang (2022). Property tax compliance in Tanzania: Can nudges help? CGD Working Paper 621.

De Neve, J.-E., C. Imbert, J. Spinnewijn, T. Tsankova, and M. Luts (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy 129*(5), 1425–1463.

Dom, R., A. Custers, S. Davenport, and W. Prichard (2022). Taxing at the local level. In *Innovations in Tax Compliance: Building Trust, Navigating Politics, and Tailoring Reform*, Chapter 6, pp. 141–170. World Bank Publications.

Fernandez, M., M. Gonzalez-Navarro, and C. Quintana-Domeque (2023). Local public goods and property tax compliance: Evidence from residential street pavement. Unpublished Manuscript, presented at the EEA-ESEM congress 2023.

Fish, P. (2018). Practical guidance note: Training manual for implementing property tax reform with a points-based valuation. ICTD African Tax Administration Paper 2.

Freetown City Council (2019). *Transform Freetown: An Overview — 2019 – 2022*. FCC, Freetown.

Giaccobasso, M., B. C. Nathan, R. Perez-Truglia, and A. Zentner (2023). Where do my tax dollars go? Tax morale effects of perceived government spending. NBER Working Paper No. w29789.

Grieco, K., A. Bakarr Kamara, N. F. Meriggi, J. Michel, W. Prichard, and G. Stewart-Wilson (2019). Simplifying property tax administration in africa: Piloting a points-based valuation in Freetown, Sierra Leone. IGC Policy Brief 39427.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics 148*, 14–31.

Hanousek, J. and F. Palda (2004). Quality of government services and the civic duty to pay taxes in the Czech and Slovak republics, and other transition countries. *Kyklos 57*(2), 237–252.

Hernandez, M., J. Jamison, E. Korczyc, N. Mazar, and R. Sormani (2017). *Applying behavioral insights to improve tax collection: Experimental Evidence from Poland*. World Bank, Washington, DC.

Hobbes, T. (1651). *Leviathan or the Matter, Forme, & Power of a Common-wealth Ecclesiasticall and Civill*. Andrew Crooke, London.

Hofmann, E., E. Hoelzl, and E. Kirchler (2008). Preconditions of voluntary tax compliance: Knowledge and evaluation of taxation, norms, fairness, and motivation to cooperate. *Zeitschrift für Psychologie/Journal of Psychology 216*(4), 209–217.

Jibao, S. S. (2017). Sierra Leone. In R. Franzsen and M. McCluskey (Eds.), *Property tax in Africa: Status, challenges, and prospects*, pp. 264–376. Lincoln Institute of Land Policy, Cambridge, Massachusetts.

Kaldor, N. (1963). Taxation for economic development. *The Journal of Modern African Studies 1*(1), 7–23.

Kamara, A. B., N. Meriggi, and W. Prichard (2020). Freetown just implemented a new property tax system that could quintuple revenue. International Centre for Tax and Development Blog [May 22, 2020].

Khan, A., A. Khawja, B. Olken, and M. Shaukat (2022). Rebuilding the social compact: Urban service delivery and property taxes in Pakistan. IGC Final Report PAK-22003.

Khan, A. Q., A. I. Khwaja, and B. A. Olken (2019). Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. *American Economic Review 109*(1), 237–70.

Knebelmann, J. and V. Pouliquen (2023). Discretion versus algorithms: Bureaucrats and tax equity in Senegal. Unpublished Manuscript, Massachusetts Institute of Technology.

Krause, B. (2020). Balancing purse and peace: Tax collection, public goods and protests. Unpublished Manuscript, University of California Berkley.

Lamberton, C., J.-E. De Neve, and M. I. Norton (2018). The power of voice in stimulating morality: Eliciting taxpayer preferences increases tax compliance. *Journal of Consumer Psychology 28*(2), 310–328.

Levi, M. (1988). *Of Rule and Revenue*. University of California Press, Berkeley, CA.

Locke, J. (1690). *Two Treatises of Government: In the Former, The False Principles and Foundation of Sir Robert Filmer, And His Followers, are Detected and Overthrown. The Latter is an Essay concerning The True Original, Extent, and End of Civil-Government*. Awnsham Churchill, London.

Luttmer, E. F. and M. Singhal (2014). Tax morale. *Journal of Economic Perspectives 28*(4), 149–68.

Mascagni, G. (2018). From the lab to the field: A review of tax experiments. *Journal of Economic Surveys 32*(2), 273–301.

Mascagni, G. and C. Nell (2022). Tax compliance in Rwanda: Evidence from a message field experiment. *Economic Development and Cultural Change 70*(2), 587–623.

Moore, M., W. Prichard, and O.-H. Fjeldstad (2018). *Taxing Africa: Coercion, reform and development*. Zed Books Ltd, London.

OECD (2019). *Tax Morale: What Drives People And Businesses To Pay Tax?* OECD Publishing, Paris.

OECD (2022). Revenue statistics in Africa 2022 – Sierra Leone. Available online [last accessed April 23, 2024].

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

Prichard, W. (2017). Linking property tax revenue and public services. ICTD Summary Brief 13.

Rousseau, J. J. (1762). *Du contrat social; ou, Principes du droit politique*. Marc-Michel Rey, Amsterdam.

Sjoberg, F. M., J. Mellon, T. C. Peixoto, J. Z. Hemker, and L. L. Tsai (2019). Voice and punishment: A global survey experiment on tax morale. World Bank Policy Research Working Paper No. 8855.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic perspectives 21*(1), 25–48.

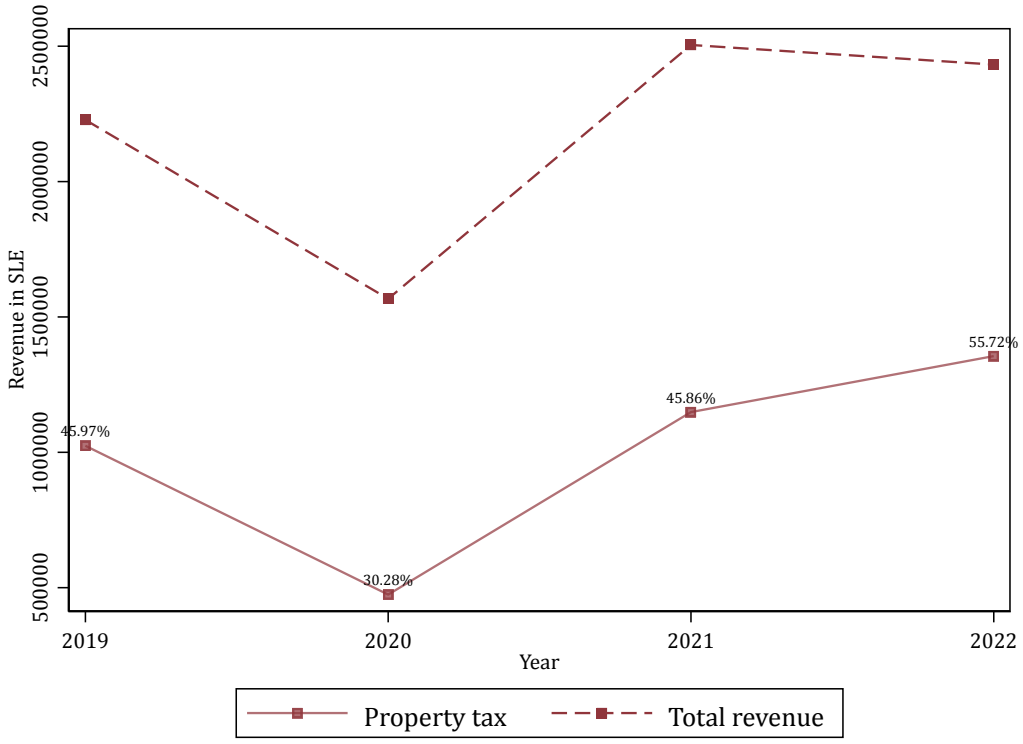
United Nations, Department of Economic and Social Affairs, Population Division (2019). *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*. United Nations, New York.

Van den Boogaard, V., W. Prichard, R. Beach, and F. Mohiuddin (2020). Strengthening tax-accountability links: Fiscal transparency and taxpayer engagement in Ghana and Sierra Leone. ICTD Working Paper 114.

Weigel, J. L. (2020). The participation dividend of taxation: How citizens in Congo engage more with the state when it tries to tax them. *The Quarterly Journal of Economics 135*(4), 1849–1903.

Appendix A

Figure A1: Property tax revenue as compared to overall FCC budget



Notes: The figure illustrates total municipality level revenue as well as revenue from property tax between 2019 and 2022.

Appendix A.1 Data

The following paragraphs explain the sampling and data collection procedures in more detail. Figure A3 at the end of the section illustrates the process graphically.

Sampling The target population of this project are owners and residents of registered domestic properties across Freetown. To sample a subset of these individuals, I use the database of registered properties generated during the recent implementation of the property tax reform. In January 2022, the database contained information about 103,407 registered properties, out of which 97,846 are domestic properties.[67] The database contains information about the property value, annual and outstanding tax liabilities, and the exact, geo-referenced location of the property, as well as phone numbers of at least one owner or resident for 79,860 properties. Restricting the database to domestic properties with at least one phone number available, I am left with 75,687 properties.

[67] This includes properties used for both domestic and commercial purposes as well as purely domestic properties.

To test the effect of information about public services among individuals for whom public service provision is considered most relevant, I restrict the sample to properties which are owned by an individual who does not own more than one property. Ownership of more than one registered property is used as an indicator for disproportionally wealthy individuals within the sample. Since wealthier citizens in Freetown tend to rely less on public service provision, and instead often make use of private providers or provide services for themselves, information on public service delivery is likely to be less relevant for these individuals. I proxy ownership of multiple properties using overlaps in owner names and phone numbers across properties.[68] Excluding properties owned by a person who possesses several properties, I am left with 46,844 properties.

To have sufficient power to test all outcomes of interest, I aimed at a final sample of 7,000 individuals. Given projections for response rates based on similar projects in the same setting, I randomly sampled 15,000 properties from the above-described final database as the respondent pool for this project.

Baseline During the baseline data collection, enumerators conducted phone interviews with a total of 5,827 individuals.[69] The collection of baseline data had three main goals: (1) to collect individual level characteristics used for a stratified randomization of the treatment assignment, (2) to elicit public service preferences in order to provide targeted information during the intervention, and (3) to collect information on baseline attitudes towards public service provision, taxation, and the local government more broadly.

Due to budgetary constraints, the full baseline survey could be conducted only with a subset of individuals. The baseline data collection was therefore done in two parts. In the first part (batch 1), enumerators interviewed individuals who either own or reside in a sampled property. These respondents received the full length survey, including a variety of questions about the local government and their interaction with the latter. It was planned that these individuals would receive an endline survey that would allow to assess the effects of the treatment on the respective variables. In total, 3,581 individuals were interviewed during this stage.

In the second part (batch 2), respondents received a shorter version of the baseline survey, which included questions about individual level characteristics, baseline public service preferences, and a small subset of relevant baseline attitudes, e.g., perceived availability of public services. For these individuals, focus was put on assessing the effects of the treatment on tax compliance using administrative data, and no endline survey was planned. Since effects on tax compliance (rather than attitudes) were expected to arise predominantly through an interaction with property

[68] Given that names were recorded manually by different enumerators when assessing the properties, names are likely spelled differently across properties even if they belong to the same individual. I therefore generate a variable quantifying the extent to which two names with the same phone number overlap using continuous values from 0 to 1. If this variable is larger than or equal to 0.7, I consider two observations with the same phone number to belong to the same individual.

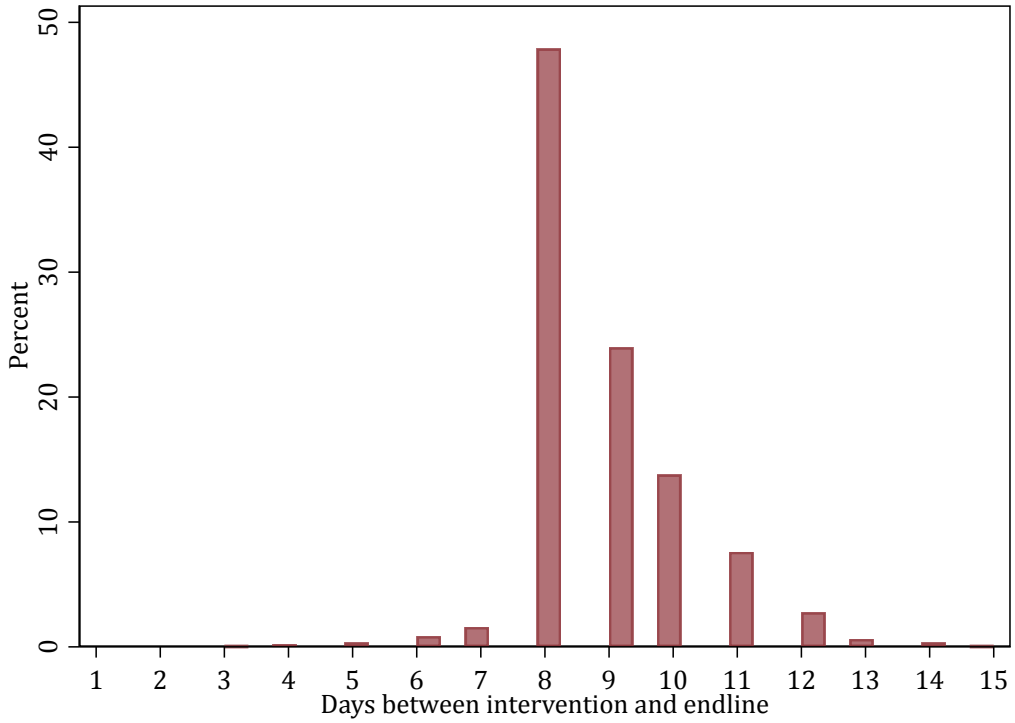
[69] Due to uncertainties concerning the funding at this stage, it was not feasible to continue the data collection until the envisioned 7,000 individuals were reached.

owners—i.e., those legally obliged to pay property tax—the second part of the baseline data collection was restricted to property owners only. A total of 2,246 individuals were interviewed during this stage.

Given the design of the baseline data collection, property owners are oversampled with respect to the natural rate of reaching an owner rather than a tenant when calling a contact number provided to the Freetown City Council. Whereas 65% of respondents in the first batch are property owners, the final sample of individuals interviewed at baseline consists of 78% property owners.

I link the baseline data to the administrative tax data including property locations, as well as the dataset of pubic service locations. Prior to randomization, I exclude the 3% of individuals with the largest distance to their most preferred service and the 3% of individuals with the largest distance to their least preferred service. Similarly, individuals with missing data on their most or least preferred public service are excluded. This leaves me with a final sample of 5,494 individuals, 3,389 in batch 1 and 2,105 in batch 2.

Figure A2: Time span between intervention and endline (treatment group, days)

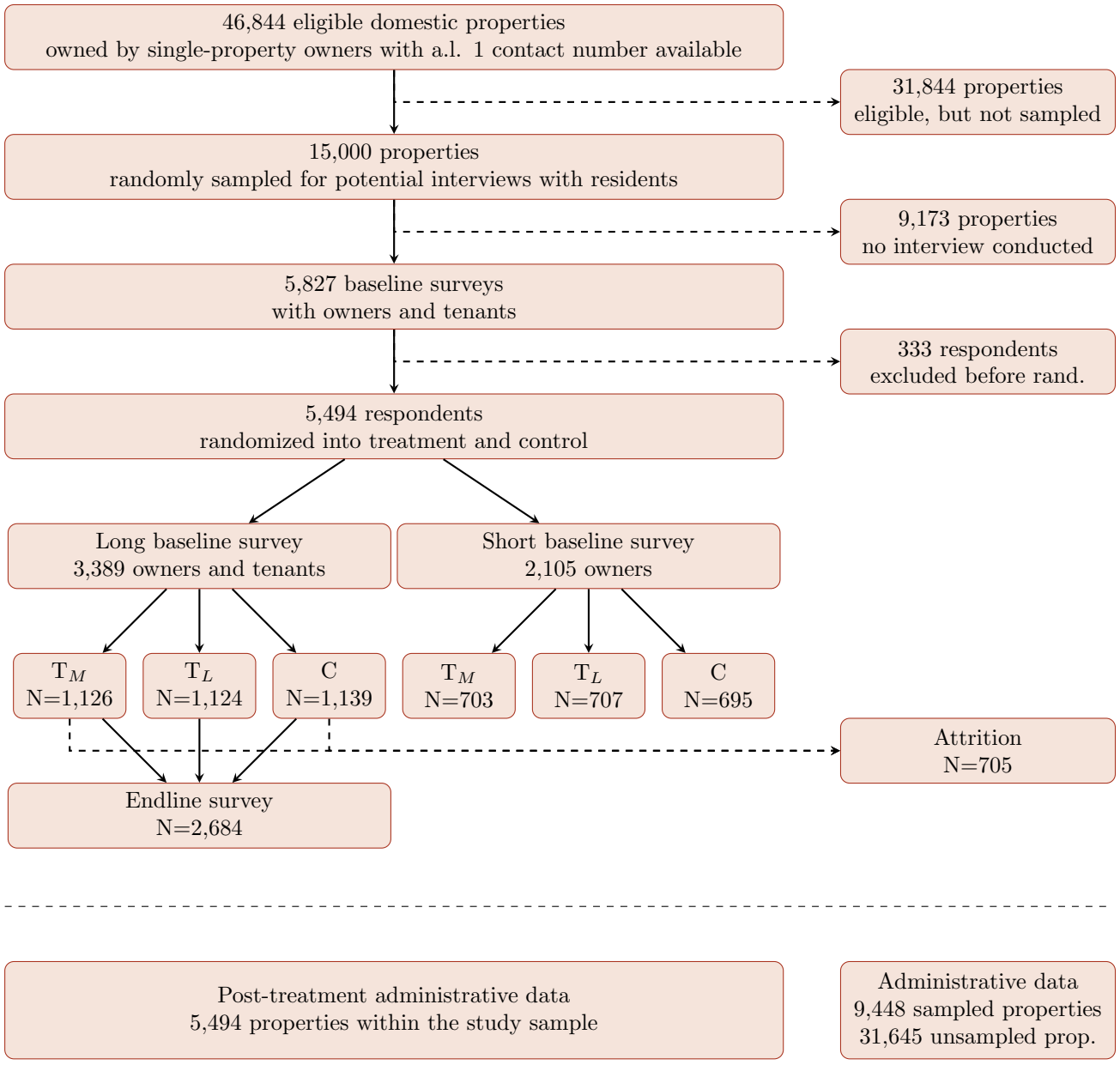


Notes: The figure illustrates the amount of time in days between the intervention and the endline survey for individuals who both received treatment and participated in the endline survey.

Endline The endline survey was administered roughly 8 to 12 days after the intervention. Figure A2 illustrates the exact distribution of days between intervention and endline survey in the treatment group. Individuals in the control were interviewed parallel to those in the treatment

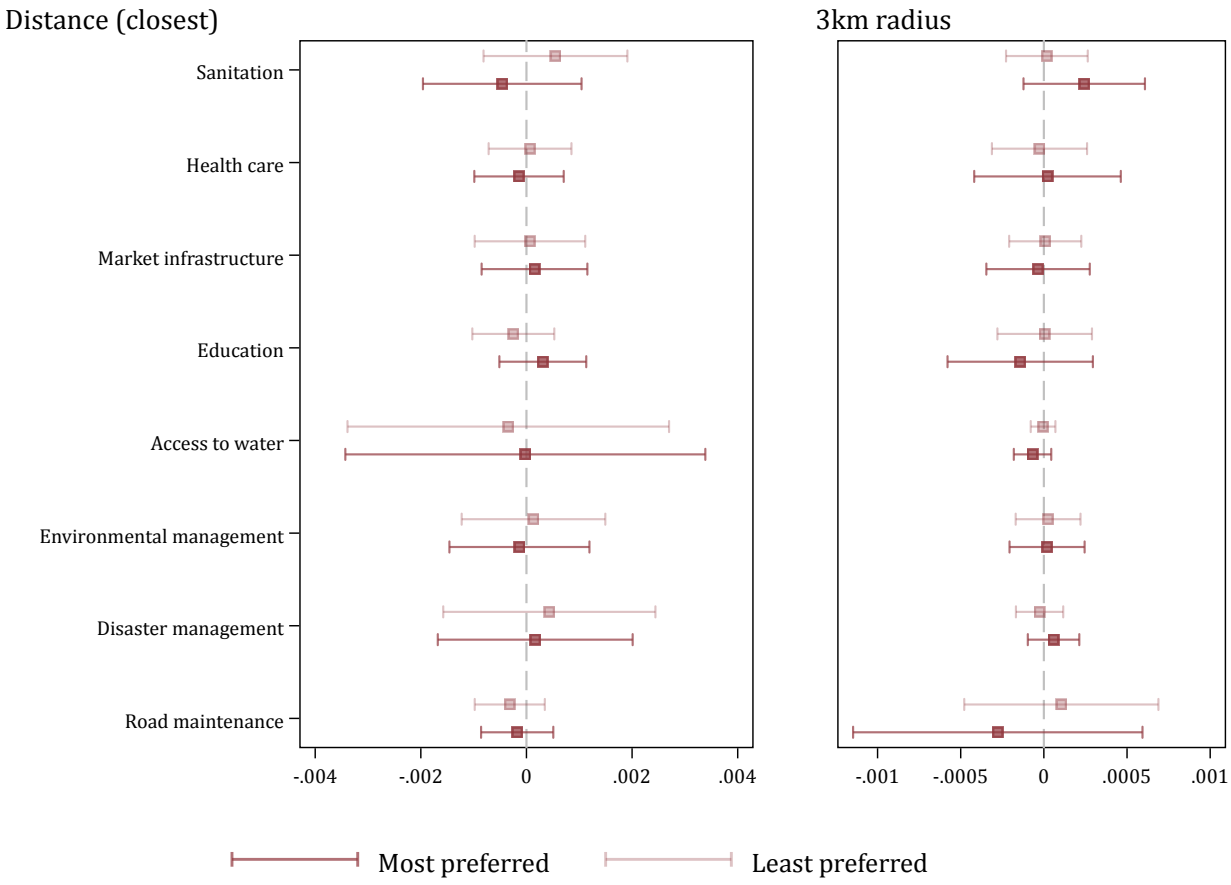
group, with the aim of keeping the share of treated and control group individuals interviewed constant over the course of the data collection. The endline survey predominantly contained questions about attitudes towards the local government, taxation, and public service provision. In total, 2,684 individuals were interviewed. This represents 79% of individuals in batch 1.

Figure A3: Research Design and Data Collection Diagram



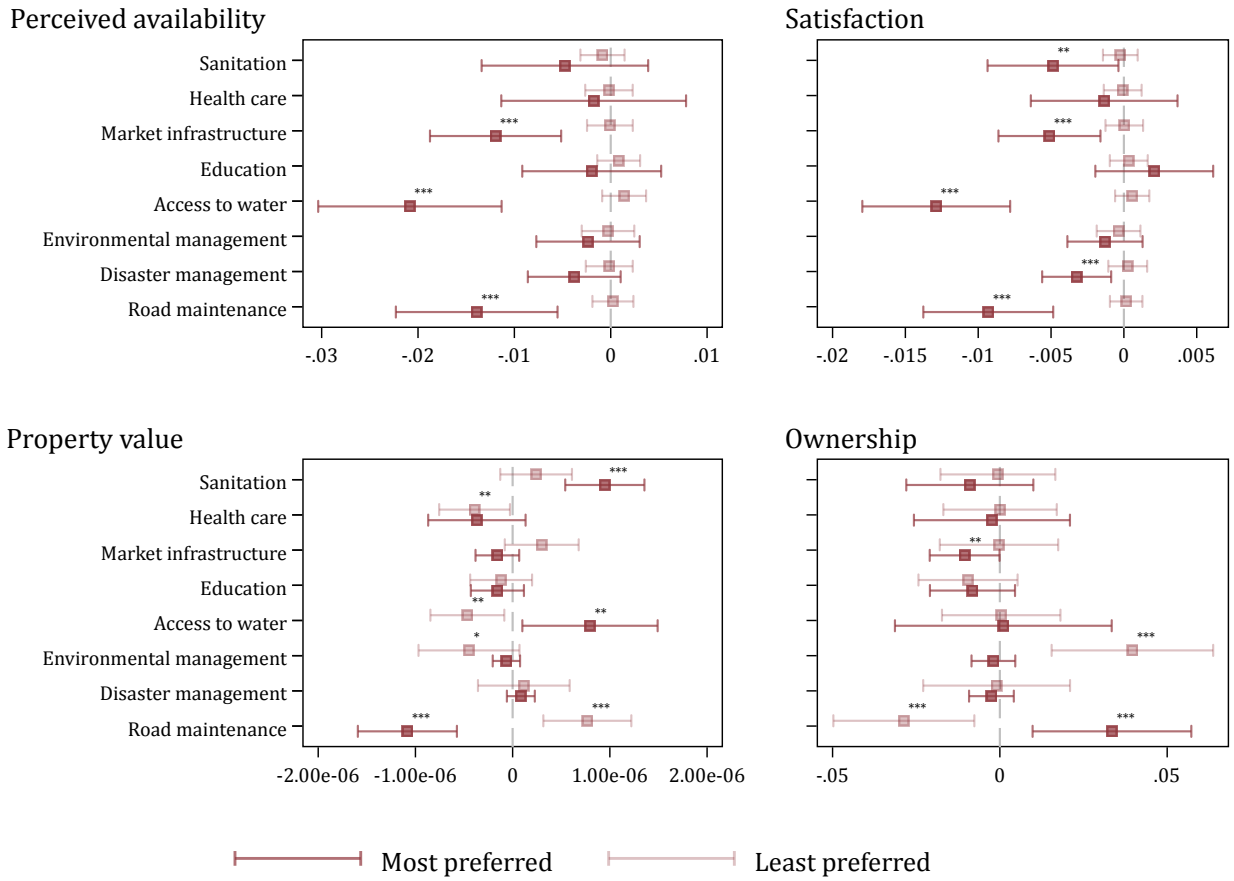
Notes: This figure illustrates the data collection procedure from sampling, over baseline data collection, randomization, and endline data collection, including information on the respective group sizes.

Figure A4: Regressions of service priorities on perceived distance to the closest service of each type and the amount of services provided within a 3km radius



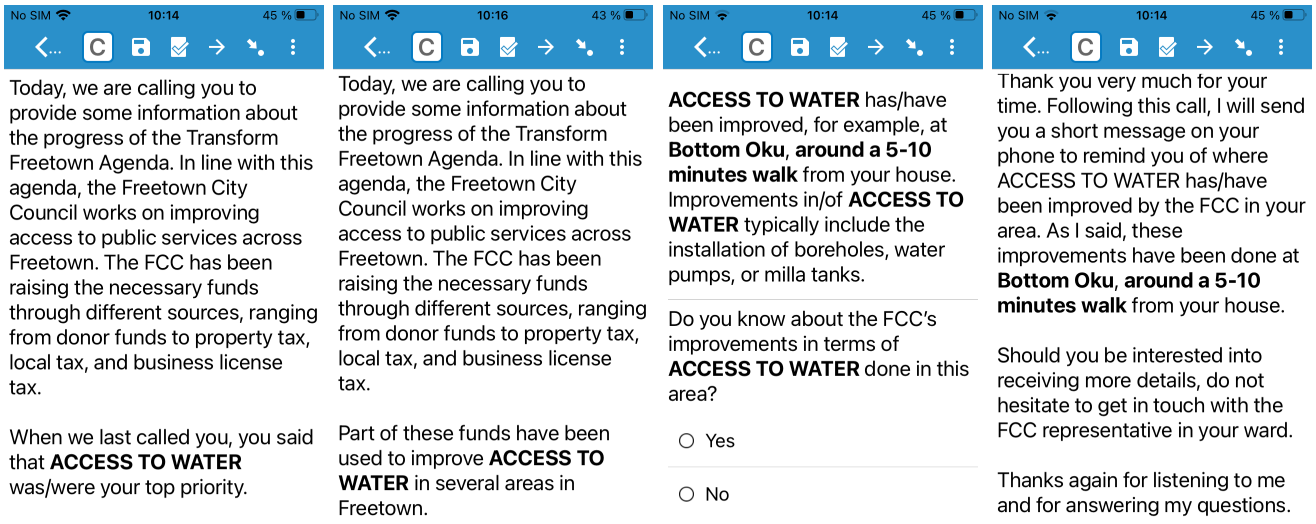
Notes: The figure shows correlations between preferences for a specific type of service and distance to the closest service of a that type recently provided by the FCC (left), as well as the amount of services of that type recently provided (right). Coefficients stem from simultaneous regressions and are shown with 95% confidence intervals. All regressions include baseline surveyor fixed effects.

Figure A5: Regressions of service priorities on perceived availability of services, satisfaction with services, property value, and property ownership



Notes: The figure shows correlations between preferences for a specific type of service and perceived availability of that type of service (upper left), satisfaction with that type of service (upper right), the value of the property in which respondents live (lower left), and ownership of the respective property (lower right). Coefficients stem from simultaneous regressions and are shown with 95% confidence intervals. All regressions include baseline surveyor fixed effects.

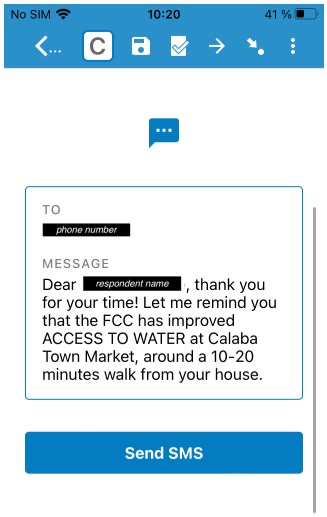
Figure A6: Intervention phone call script (SurveyCTO)



(a) Part I - Match (b) Part I - Mismatch (c) Part II (d) Part III

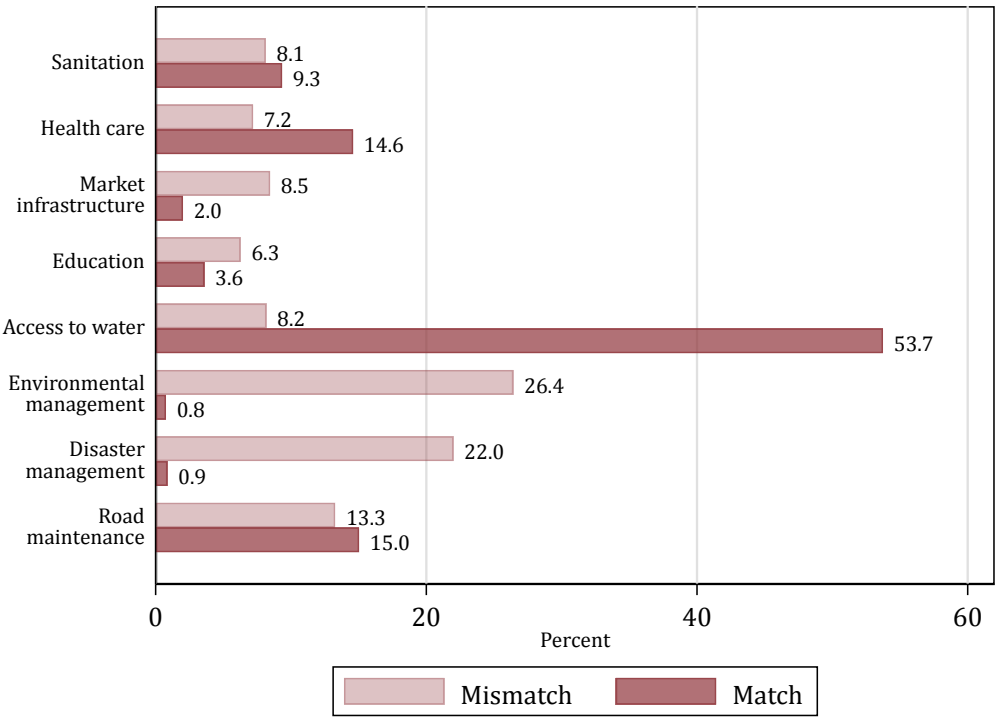
Notes: The figure shows the script used for the intervention call. Figures A6a and A6b illustrate the call's introduction for the match and mismatch treatment, respectively. Figure A6c shows the main part of the message. Figure A6d illustrates the final part of the call.

Figure A7: Follow-up short message (SMS)



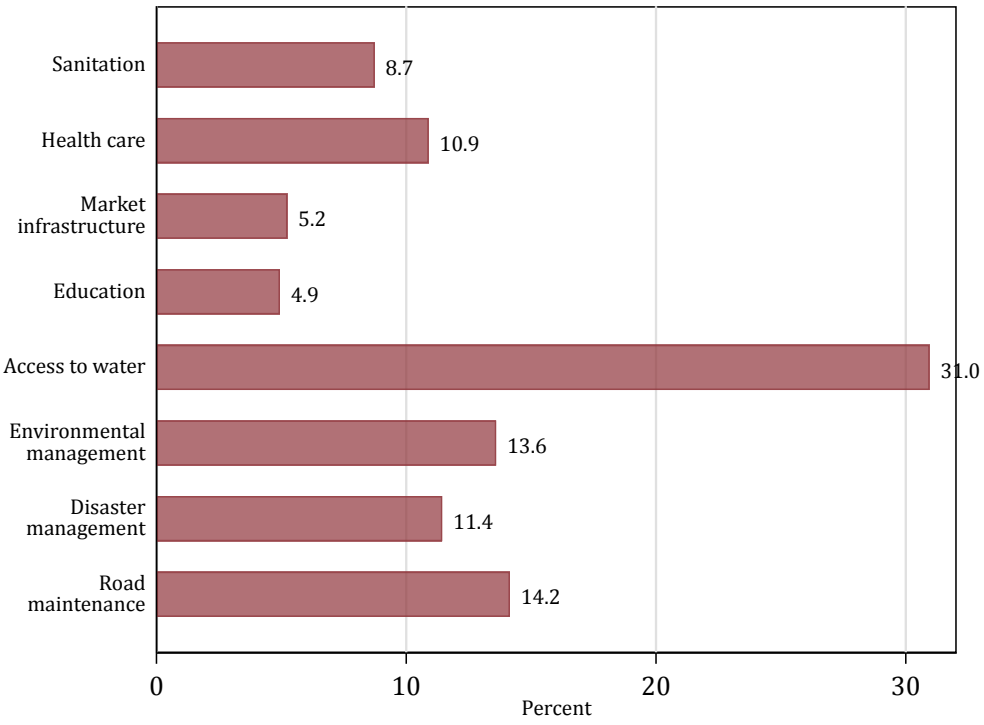
Notes: The figure shows an example of an SMS sent to treated respondents after the treatment call.

Figure A8: Type of service individuals where informed about by treatment arm



Notes: The figure shows the share of individuals in the match and mismatch group informed about a specific type of service.

Figure A9: Type of service individuals where informed about



Notes: The figure shows the share of treated individuals informed about a specific type of service.

Table A1: Characteristics of eligible as compared to sampled properties

Variable	N	(1) Eligible Mean/SE	N	(2) Sampled Mean/SE	t-test Difference (1)-(2)
Assessed property value	31845	13818.404	14999	14070.650	-252.246
		[107.970]		[157.484]	
Property is both domestic and commercial/institutional	31845	0.050	14999	0.055	-0.005**
		[0.001]		[0.002]	
Ward group 1	31844	0.184	14999	0.193	-0.009**
		[0.002]		[0.003]	
Ward group 2	31844	0.135	14999	0.139	-0.004
		[0.002]		[0.003]	
Ward group 3	31844	0.133	14999	0.130	0.003
		[0.002]		[0.003]	
Ward group 4	31844	0.134	14999	0.128	0.005
		[0.002]		[0.003]	
Ward group 5	31844	0.192	14999	0.192	-0.000
		[0.002]		[0.003]	
Ward group 6	31844	0.222	14999	0.217	0.005
		[0.002]		[0.003]	

Notes: The value displayed for t-tests are the differences in the means across groups. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Ward groups refer to groupings of neighboring wards.

Table A2: Characteristics of eligible as compared to interviewed properties

Variable	N	(1) Eligible Mean/SE	N	(2) Interviewed Mean/SE	t-test Difference (1)-(2)
Assessed property value	41350	13914.932	5494	13780.543	134.389
		[95.247]		[250.256]	
Property is both domestic and commercial/institutional	41350	0.050	5494	0.058	-0.007**
		[0.001]		[0.003]	
Ward group 1	41349	0.192	5494	0.153	0.038***
		[0.002]		[0.005]	
Ward group 2	41349	0.134	5494	0.159	-0.025***
		[0.002]		[0.005]	
Ward group 3	41349	0.130	5494	0.145	-0.015***
		[0.002]		[0.005]	
Ward group 4	41349	0.133	5494	0.122	0.012**
		[0.002]		[0.004]	
Ward group 5	41349	0.190	5494	0.206	-0.015***
		[0.002]		[0.005]	
Ward group 6	41349	0.221	5494	0.215	0.006
		[0.002]		[0.006]	

Notes: The value displayed for t-tests are the differences in the means across groups. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Ward groups refer to groupings of neighboring wards.

Table A3: Baseline characteristics of individuals interviewed at endline compared to those not interviewed at endline

Variable	(1) No N	(1) No Mean/SE	(2) Yes N	(2) Yes Mean/SE	T-test Difference (1)-(2)
Female	2810	0.356	2684	0.349	0.006
		(0.009)		(0.009)	
Age	2804	5.813	2673	5.447	0.366***
		(0.050)		(0.050)	
Education level	2785	4.269	2682	4.208	0.061
		(0.044)		(0.045)	
In relationship	2806	0.870	2683	0.873	-0.003
		(0.006)		(0.006)	
Owner	2810	0.913	2684	0.644	0.269***
		(0.005)		(0.009)	
Monthly income	2399	3.351	2577	3.423	-0.071
		(0.037)		(0.036)	
Ward group 1	2810	0.148	2684	0.159	-0.011
		(0.007)		(0.007)	
Ward group 2	2810	0.150	2684	0.168	-0.018*
		(0.007)		(0.007)	
Ward group 3	2810	0.143	2684	0.147	-0.003
		(0.007)		(0.007)	
Ward group 4	2810	0.123	2684	0.120	0.002
		(0.006)		(0.006)	
Ward group 5	2810	0.214	2684	0.197	0.016
		(0.008)		(0.008)	
Ward group 6	2810	0.222	2684	0.208	0.013
		(0.008)		(0.008)	
Distance to most preferred service (km)	2810	1.476	2684	1.573	-0.098**
		(0.033)		(0.035)	
Distance to least preferred service (km)	2810	1.632	2684	1.680	-0.048
		(0.028)		(0.030)	
Most preferred service: water	2810	0.542	2684	0.509	0.033**
		(0.009)		(0.010)	
Least preferred service: water	2810	0.085	2684	0.082	0.003
		(0.005)		(0.005)	
Property value (SLL)	2810	1.45e+07	2684	1.31e+07	1.38e+06***
		(3.64e+05)		(3.42e+05)	

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A4: Balance of baseline characteristics across treatment groups

	(1) Match		(2) Mismatch		(3) Control		T-test Difference		
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)
Female	1829	0.353	1831	0.353	1834	0.352	-0.000	0.000	0.001
		(0.011)		(0.011)		(0.011)			
Age	1822	5.581	1827	5.732	1828	5.590	-0.152*	-0.010	0.142
		(0.061)		(0.063)		(0.060)			
Education level	1817	4.232	1821	4.273	1829	4.212	-0.041	0.021	0.061
		(0.055)		(0.054)		(0.054)			
In relationship	1826	0.867	1830	0.866	1833	0.883	0.001	-0.016	-0.017
		(0.008)		(0.008)		(0.008)			
Owner	1829	0.781	1831	0.785	1834	0.779	-0.004	0.003	0.006
		(0.010)		(0.010)		(0.010)			
Monthly income	1660	3.400	1659	3.374	1657	3.391	0.026	0.009	-0.017
		(0.045)		(0.046)		(0.044)			
Ward group 1	1829	0.153	1831	0.152	1834	0.155	0.000	-0.003	-0.003
		(0.008)		(0.008)		(0.008)			
Ward group 2	1829	0.158	1831	0.161	1834	0.158	-0.003	0.000	0.004
		(0.009)		(0.009)		(0.009)			
Ward group 3	1829	0.146	1831	0.147	1834	0.142	-0.001	0.004	0.006
		(0.008)		(0.008)		(0.008)			
Ward group 4	1829	0.121	1831	0.122	1834	0.122	-0.001	-0.001	-0.000
		(0.008)		(0.008)		(0.008)			
Ward group 5	1829	0.205	1831	0.205	1834	0.207	-0.000	-0.002	-0.002
		(0.009)		(0.009)		(0.009)			
Ward group 6	1829	0.218	1831	0.212	1834	0.216	0.006	0.002	-0.004
		(0.010)		(0.010)		(0.010)			
Distance to most preferred service (km)	1829	1.518	1831	1.537	1834	1.515	-0.019	0.003	0.022
		(0.042)		(0.041)		(0.041)			
Distance to least preferred service (km)	1829	1.644	1831	1.681	1834	1.643	-0.038	0.001	0.039
		(0.035)		(0.036)		(0.035)			
Most preferred service: water	1829	0.537	1831	0.522	1834	0.519	0.015	0.019	0.004
		(0.012)		(0.012)		(0.012)			
Least preferred service: water	1829	0.077	1831	0.082	1834	0.092	-0.005	-0.015	-0.010
		(0.006)		(0.006)		(0.007)			
Property value (SLL)	1829	1.40e+07	1831	1.34e+07	1834	1.40e+07	5.93e+05	-1.45e+04	-6.08e+05
		(4.55e+05)		(4.17e+05)		(4.27e+05)			

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A5: Balance of baseline characteristics across treatment arms among individuals who participated in the endline survey

	(1) Match		(2) Mismatch		(3) Control		T-test Difference		
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)
Female	900	0.342	883	0.350	901	0.356	-0.008	-0.014	-0.006
		(0.016)		(0.016)		(0.016)			
Age	898	5.362	880	5.619	895	5.363	-0.257**	-0.001	0.256**
		(0.085)		(0.091)		(0.084)			
Education level	898	4.259	883	4.227	901	4.138	0.033	0.122	0.089
		(0.078)		(0.078)		(0.077)			
In relationship	899	0.867	883	0.866	901	0.887	0.000	-0.020	-0.020
		(0.011)		(0.011)		(0.011)			
Owner	900	0.652	883	0.644	901	0.636	0.008	0.016	0.008
		(0.016)		(0.016)		(0.016)			
Monthly income	868	3.472	848	3.439	861	3.357	0.034	0.116	0.082
		(0.062)		(0.063)		(0.062)			
Ward group 1	900	0.151	883	0.153	901	0.173	-0.002	-0.022	-0.020
		(0.012)		(0.012)		(0.013)			
Ward group 2	900	0.160	883	0.176	901	0.169	-0.016	-0.009	0.007
		(0.012)		(0.013)		(0.012)			
Ward group 3	900	0.144	883	0.151	901	0.145	-0.006	-0.001	0.005
		(0.012)		(0.012)		(0.012)			
Ward group 4	900	0.122	883	0.123	901	0.115	-0.001	0.007	0.008
		(0.011)		(0.011)		(0.011)			
Ward group 5	900	0.204	883	0.195	901	0.193	0.010	0.011	0.002
		(0.013)		(0.013)		(0.013)			
Ward group 6	900	0.218	883	0.203	901	0.204	0.015	0.014	-0.001
		(0.014)		(0.014)		(0.013)			
Distance to most preferred service (km)	900	1.536	883	1.560	901	1.625	-0.024	-0.089	-0.065
		(0.060)		(0.059)		(0.062)			
Distance to least preferred service (km)	900	1.705	883	1.654	901	1.682	0.050	0.023	-0.027
		(0.051)		(0.050)		(0.052)			
Most preferred service: water	900	0.508	883	0.519	901	0.501	-0.011	0.007	0.018
		(0.017)		(0.017)		(0.017)			
Least preferred service: water	900	0.069	883	0.077	901	0.100	-0.008	-0.031**	-0.023*
		(0.008)		(0.009)		(0.010)			
Property value (SLL)	900	1.30e+07	883	1.30e+07	901	1.33e+07	-8446.520	-3.51e+05	-3.43e+05
		(5.80e+05)		(6.06e+05)		(5.94e+05)			

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table A6: Balance of baseline service preferences and distance to services across treatment groups

	(1) Match		(2) Mismatch		(3) Control		T-test Difference		
Variable	N	Mean/SE	N	Mean/SE	N	Mean/SE	(1)-(2)	(1)-(3)	(2)-(3)
most = sanitation	1829	0.093	1831	0.092	1834	0.089	0.002	0.004	0.002
		(0.007)		(0.007)		(0.007)			
most = health	1829	0.146	1831	0.149	1834	0.148	-0.003	-0.002	0.001
		(0.008)		(0.008)		(0.008)			
most = markets	1829	0.020	1831	0.031	1834	0.026	-0.010**	-0.005	0.005
		(0.003)		(0.004)		(0.004)			
most = education	1829	0.036	1831	0.040	1834	0.039	-0.004	-0.003	0.001
		(0.004)		(0.005)		(0.005)			
most = water	1829	0.537	1831	0.522	1834	0.519	0.015	0.019	0.004
		(0.012)		(0.012)		(0.012)			
most = environment	1829	0.008	1831	0.010	1834	0.011	-0.003	-0.004	-0.001
		(0.002)		(0.002)		(0.002)			
most = disaster	1829	0.009	1831	0.008	1834	0.014	0.001	-0.005	-0.007*
		(0.002)		(0.002)		(0.003)			
most = roads	1829	0.150	1831	0.149	1834	0.154	0.001	-0.004	-0.005
		(0.008)		(0.008)		(0.008)			
least = sanitation	1829	0.083	1831	0.081	1834	0.081	0.001	0.002	0.001
		(0.006)		(0.006)		(0.006)			
least = health	1829	0.077	1831	0.072	1834	0.077	0.005	-0.000	-0.005
		(0.006)		(0.006)		(0.006)			
least = markets	1829	0.083	1831	0.085	1834	0.091	-0.002	-0.008	-0.006
		(0.006)		(0.007)		(0.007)			
least = education	1829	0.068	1831	0.063	1834	0.059	0.006	0.009	0.004
		(0.006)		(0.006)		(0.005)			
least = water	1829	0.077	1831	0.082	1834	0.092	-0.005	-0.015	-0.010
		(0.006)		(0.006)		(0.007)			
least = environment	1829	0.267	1831	0.264	1834	0.260	0.002	0.007	0.004
		(0.010)		(0.010)		(0.010)			
least = disaster	1829	0.215	1831	0.220	1834	0.213	-0.005	0.002	0.007
		(0.010)		(0.010)		(0.010)			
least = roads	1829	0.131	1831	0.133	1834	0.128	-0.002	0.003	0.005
		(0.008)		(0.008)		(0.008)			
Approximate walking distance in km: sanitation	1829	1.439	1831	1.469	1834	1.486	-0.031	-0.047	-0.017
		(0.026)		(0.027)		(0.028)			
Approximate walking distance in km: health	1829	3.182	1831	3.191	1834	3.190	-0.009	-0.008	0.002
		(0.037)		(0.037)		(0.037)			
Approximate walking distance in km: markets	1829	2.297	1831	2.331	1834	2.335	-0.035	-0.039	-0.004
		(0.053)		(0.057)		(0.057)			
Approximate walking distance in km: education	1829	2.991	1831	3.004	1834	3.004	-0.013	-0.014	-0.000
		(0.044)		(0.045)		(0.046)			
Approximate walking distance in km: water	1829	0.486	1831	0.497	1834	0.484	-0.011	0.003	0.014
		(0.008)		(0.008)		(0.008)			
Approximate walking distance in km: environment	1829	1.232	1831	1.253	1834	1.250	-0.020	-0.018	0.002
		(0.020)		(0.021)		(0.020)			
Approximate walking distance in km: disaster	1829	0.908	1831	0.939	1834	0.938	-0.030	-0.030	0.000
		(0.017)		(0.019)		(0.019)			
Approximate walking distance in km: roads	1829	3.577	1831	3.665	1834	3.641	-0.088	-0.064	0.024
		(0.073)		(0.075)		(0.075)			
Approximate walking distance in km: most	1829	1.518	1831	1.537	1834	1.515	-0.019	0.003	0.022
		(0.042)		(0.041)		(0.041)			
Approximate walking distance in km: least	1829	1.644	1831	1.681	1834	1.643	-0.038	0.001	0.039
		(0.035)		(0.036)		(0.035)			
Above median distance (most)	1829	0.483	1831	0.487	1834	0.481	-0.004	0.002	0.006
		(0.012)		(0.012)		(0.012)			
Above median distance (least)	1829	0.484	1831	0.507	1834	0.495	-0.023	-0.010	0.013
		(0.012)		(0.012)		(0.012)			

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Appendix B

Table B1: Treatment effects on the amount of tax paid including different control variables (September 2022)

	Amount paid (SLE)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	23.784**	22.880**	17.946*	22.114**
	(11.322)	(10.993)	(9.799)	(9.627)
Panel B. Treatment effect by arm				
Match	17.226	15.324	4.064	7.894
	(14.639)	(14.378)	(10.222)	(9.857)
Mismatch	31.115**	30.503**	30.568**	35.413**
	(14.913)	(14.704)	(14.390)	(14.277)
Service type FEs (m, l)	No	Yes	Yes	Yes
Distance linear (m, l)	No	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	No	No	Yes	Yes
Wealth FEs	No	No	No	Yes
p-val. Match = Mismatch	0.431	0.393	0.076	0.057
Obs.	5,494	5,494	5,469	5,469
Control mean	118.829	118.829	118.829	118.829
SD	418.244	418.244	418.244	418.244

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B2: Treatment effects on any payment made including different control variables (September 2022)

	Paid any tax (d)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	0.012	0.012	0.012	0.013
	(0.008)	(0.008)	(0.008)	(0.008)
Panel B. Treatment effect by arm				
Match	0.012	0.011	0.011	0.012
	(0.009)	(0.009)	(0.009)	(0.009)
Mismatch	0.013	0.013	0.014	0.015
	(0.009)	(0.009)	(0.009)	(0.009)
Service type FEs (m, l)	No	Yes	Yes	Yes
Distance linear (m, l)	No	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	No	No	Yes	Yes
Wealth FEs	No	No	No	Yes
p-val. Match = Mismatch	0.949	0.866	0.817	0.778
Obs.	5,494	5,494	5,469	5,469
Control mean	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B3: Treatment effects on the amount of tax paid conditional on payment including different control variables (September 2022)

	Amount paid (SLE, conditional)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	74.983	65.295	54.497	84.546**
	(55.252)	(51.904)	(49.515)	(42.135)
Panel B. Treatment effect by arm				
Match	26.249	2.016	-29.021	13.145
	(64.738)	(60.195)	(51.647)	(43.045)
Mismatch	110.366	104.502	121.975*	143.640**
	(74.488)	(73.076)	(69.479)	(56.784)
Service type FEs (m, l)	No	Yes	Yes	Yes
Distance linear (m, l)	No	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	No	No	Yes	Yes
Wealth FEs	No	No	No	Yes
p-val. Match = Mismatch	0.288	0.190	0.027	0.011
Obs.	1,159	1,159	1,151	1,151
Control mean	590.605	590.605	590.605	590.605
SD	769.365	769.365	769.365	769.365

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B4: Treatment effects on the amount of tax paid including different controls for distance (September 2022)

	Amount paid (SLE)			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	22.114**	22.303**	21.909**	22.437**
	(9.627)	(9.601)	(9.703)	(9.615)
Panel B. Treatment effect by arm				
Match	7.894	8.360	7.654	8.141
	(9.857)	(9.809)	(9.881)	(9.829)
Mismatch	35.413**	35.341**	35.304**	35.472**
	(14.277)	(14.249)	(14.347)	(14.269)
Distance linear (m, l)	Yes	No	No	No
Above median distance (m, l)	No	Yes	No	No
Distance quartiles (m, l)	No	No	Yes	No
Distance linear (all services)	No	No	No	Yes
p-val. Match = Mismatch	0.057	0.060	0.056	0.057
Obs.	5,469	5,469	5,469	5,469
Control mean	118.829	118.829	118.829	118.829
SD	418.244	418.244	418.244	418.244

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B5: Treatment effects on any payment made including different controls for distance (September 2022)

	Any payment made by September 2022			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	0.013	0.013	0.013	0.013
	(0.008)	(0.008)	(0.008)	(0.008)
Panel B. Treatment effect by arm				
Match	0.012	0.011	0.012	0.012
	(0.009)	(0.009)	(0.009)	(0.009)
Mismatch	0.015	0.014	0.014	0.014
	(0.009)	(0.009)	(0.009)	(0.009)
Distance linear (m, l)	Yes	No	No	No
Above median distance (m, l)	No	Yes	No	No
Distance quartiles (m, l)	No	No	Yes	No
Distance linear (all services)	No	No	No	Yes
p-val. Match = Mismatch	0.778	0.773	0.778	0.790
Obs.	5,469	5,469	5,469	5,469
Control mean	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401

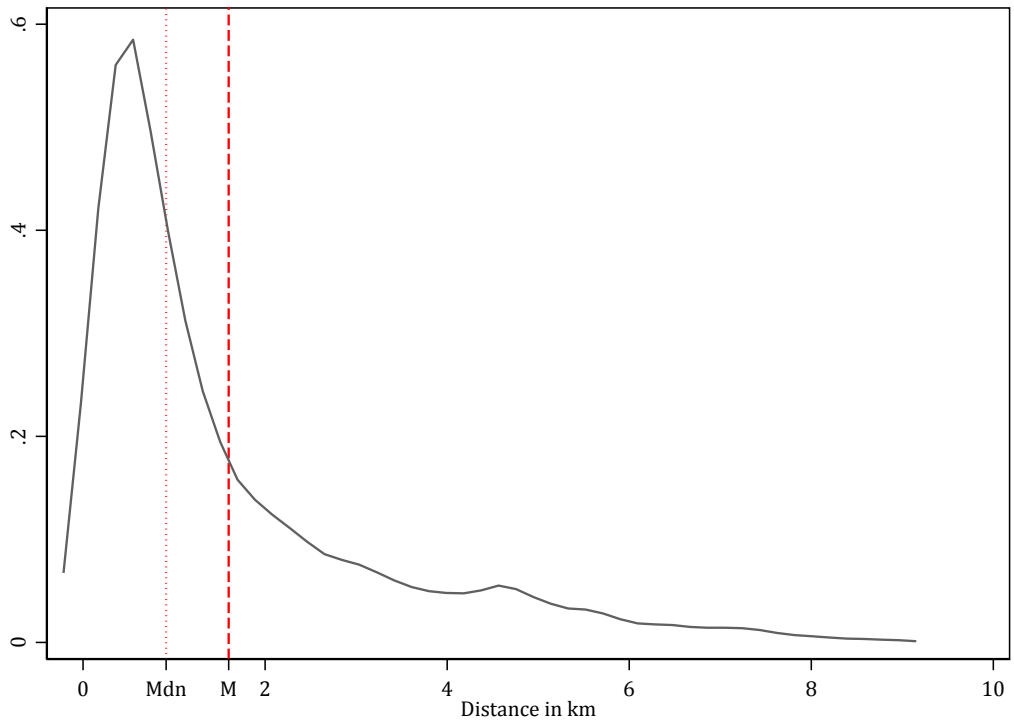
Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B6: Treatment effects on the amount of tax paid conditional on payment including different controls for distance (September 2022)

	Any payment made by September 2022			
	(1)	(2)	(3)	(4)
Panel A. Overall treatment effect				
Treated	84.546**	79.746*	81.361*	85.316**
	(42.135)	(41.546)	(42.101)	(42.388)
Panel B. Treatment effect by arm				
Match	13.145	8.646	8.809	13.484
	(43.045)	(42.790)	(43.551)	(42.835)
Mismatch	143.640**	139.160**	142.944**	147.082**
	(56.784)	(55.779)	(56.230)	(57.223)
Distance linear (m, l)	Yes	No	No	No
Above median distance (m, l)	No	Yes	No	No
Distance quartiles (m, l)	No	No	Yes	No
Distance linear (all services)	No	No	No	Yes
p-val. Match = Mismatch	0.011	0.011	0.009	0.009
Obs.	1,151	1,151	1,151	1,151
Control mean	590.605	590.605	590.605	590.605
SD	769.365	769.365	769.365	769.365

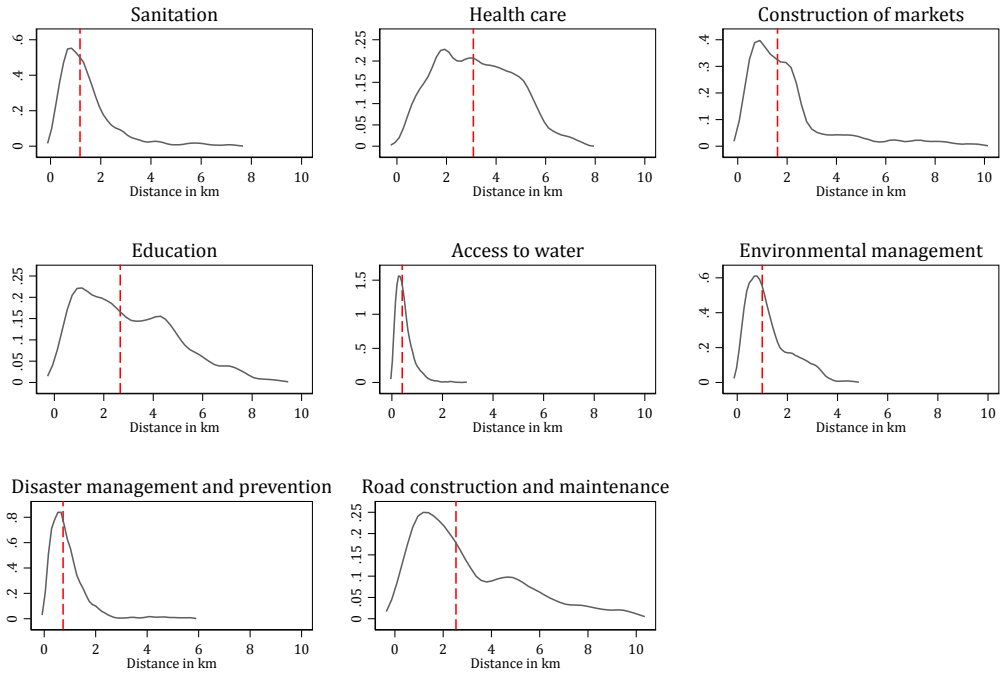
Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Figure B1: Distance to treatment service



Notes: The figure shows the distribution of distance to the service treated individuals were informed about. Mdn marks the median distance, M marks the mean distance.

Figure B2: Minimum distance to services by type



Note: Red lines mark type-specific median distances.

Notes: The figure shows the distribution of distance to the closest service of each type. The dotted line marks the median distance.

Figure B3: Heterogeneity of the treatment effects on whether any tax was paid depending on the content of the treatment message (September 2022)

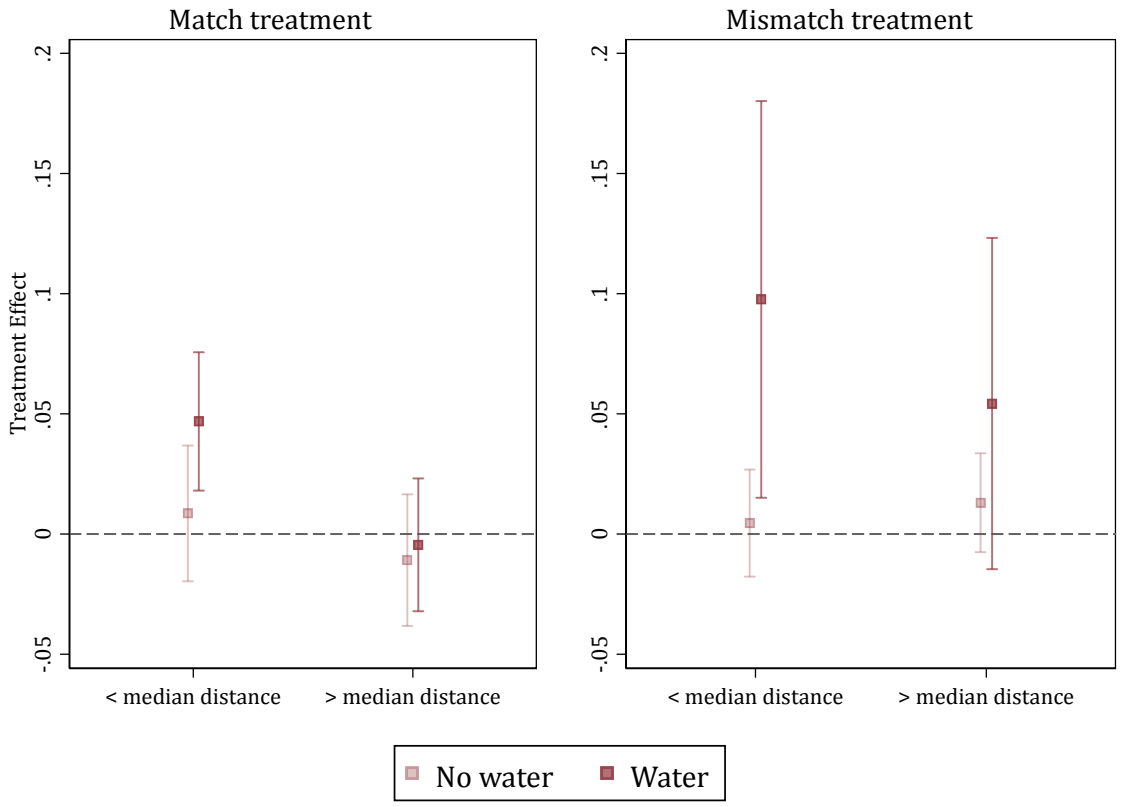
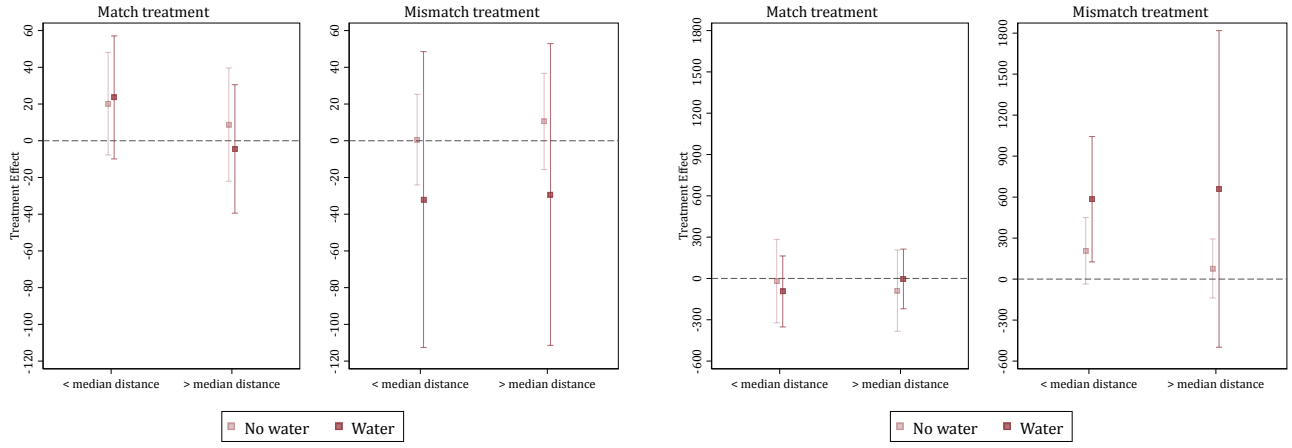


Figure B4: Heterogeneity of the treatment effects on the amount of tax paid conditional on payment depending on the content of the treatment message (September 2022)



(a) Below median property value

(b) Above median property value

Table B7: Robustness checks—Treatment effects on tax compliance (September 2022)

	Tax compliance by September 2022						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Amount paid (SLE) (September)	23.565**	23.784**	21.464*	22.114**	20.127**	19.717**	20.459**
	(11.826)	(11.322)	(11.154)	(9.627)	(9.558)	(9.796)	(10.016)
Amount paid (SLE) (win 95, share) (September)	22.824*	23.106**	20.865*	21.359**	19.449**	19.031*	19.742**
	(11.769)	(11.255)	(11.084)	(9.551)	(9.480)	(9.714)	(9.940)
Paid any tax (d) (September)	0.013	0.012	0.011	0.013	0.012	0.012	0.013
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Amount paid (SLE, cond.) (September)	90.164*	74.983	95.681*	84.546**	90.001**	87.242**	86.058*
	(53.291)	(55.252)	(56.998)	(42.135)	(42.597)	(44.095)	(44.338)
Amount paid (SLE, cond.) (win 95, share) (September)	86.623	72.431	93.055	82.479**	87.614**	84.714*	83.501*
	(53.024)	(54.873)	(56.643)	(41.634)	(42.136)	(43.716)	(44.003)
Base controls		Yes	Yes	Yes	Yes	Yes	Yes
Indiv. controls			Yes		Yes	Yes	Yes
Service type FEs (m, l)				Yes	Yes	Yes	Yes
Distance linear (m, l)				Yes	Yes	Yes	Yes
Perceived availabilty at BL (m, l)				Yes	Yes	Yes	Yes
Wealth FEs				Yes	Yes	Yes	Yes
Obs.	1,159	1,159	1,159	1,151	1,151	1,151	1,151
Control mean	587.467	587.467	587.467	587.467	587.467	587.467	587.467
SD	762.420	762.420	762.420	762.420	762.420	762.420	762.420

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Column (1) shows results from the main specification. Column (2) additionally controls for respondent gender, education level, marital status, and income level. Column (3) is based on the main specification and additionally controls for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership. Column (4) combines Columns (2) and (3). Column (5) uses the same controls as Column (4), but does not employ any weights. Column (6) uses the same controls as Column (4) and employs weights adjusting for oversampling of property owners.

Table B8: Robustness checks—Heterogeneity of the treatment effects on the amount of tax paid conditional on payment depending on the content of the treatment message (September 2022)

	Any tax paid by September 2022						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Match	0.009	0.011	0.010	0.009	0.008	0.006	0.008
	(0.018)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Match × Water (m)	0.034	0.036	0.037	0.038	0.039	0.040*	0.040*
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
Match × > median distance (m)	-0.026	-0.026	-0.025	-0.019	-0.020	-0.019	-0.021
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Match × Water (m) × > median distance (m)	-0.016	-0.025	-0.026	-0.032	-0.031	-0.032	-0.031
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
Mismatch	0.008	0.004	0.005	0.005	0.006	0.006	0.008
	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
Mismatch × Water (l)	0.087*	0.091*	0.087*	0.093*	0.089*	0.089*	0.086*
	(0.051)	(0.052)	(0.051)	(0.052)	(0.051)	(0.051)	(0.051)
Mismatch × > median distance (l)	0.004	0.007	0.003	0.008	0.005	0.004	0.001
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Mismatch × Water (l) × > median distance (l)	-0.046	-0.045	-0.037	-0.052	-0.045	-0.044	-0.043
	(0.066)	(0.068)	(0.067)	(0.068)	(0.067)	(0.067)	(0.067)
Base controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Indiv. controls	Yes		Yes		Yes	Yes	Yes
Service type FEs (m, l)	Yes			Yes	Yes	Yes	Yes
Distance linear (m, l)	Yes			Yes	Yes	Yes	Yes
Perceived availabilty at BL (m, l)	Yes			Yes	Yes	Yes	Yes
Wealth FEs	Yes			Yes	Yes	Yes	Yes
p-val. Match = Mismatch	0.945	0.694	0.780	0.838	0.927	0.975	0.980
Match + water (m) = 0	0.014	0.007	0.007	0.007	0.008	0.008	0.007
Mismatch + water (l) = 0	0.054	0.058	0.064	0.052	0.056	0.056	0.058
Match + > median distance (m) = 0	0.319	0.394	0.364	0.514	0.463	0.428	0.425
Mismatch + > median distance (l) = 0	0.335	0.411	0.517	0.297	0.388	0.428	0.453
Match + water (m) + > median distance (m) = 0	0.932	0.847	0.811	0.789	0.782	0.776	0.761
Mismatch + water (l) + > median distance (l) = 0	0.193	0.170	0.167	0.195	0.195	0.187	0.201
Obs.	5,494	5,494	5,494	5,469	5,469	5,469	5,469
Control mean	0.201	0.201	0.201	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. All regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. Column (1) shows results from the main specification. Column (2) additionally controls for respondent gender, education level, marital status, and income level. Column (3) is based on the main specification and additionally controls for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership. Column (4) combines Columns (2) and (3). Column (5) uses the same controls as Column (4), but does not employ any weights. Column (6) uses the same controls as Column (4) and employs weights adjusting for oversampling of property owners.

Endline Data Measurement Challenges As explained in Section 3, the endline survey was administered to the subsample of respondents containing both property owners and tenants in the proportion in which they were reached during the baseline interview. As opposed to this, in the overall sample, property owners are oversampled. Robustness checks suggest that the tax compliance results are robust to the use of weights adjusting for the share of property owners in the randomly selected subsample of owners and tenants (see Tables B7 and B8). This is important, as scaling up this intervention to a larger set of properties will likely yield a similar distribution between property owners and tenants being contacted, given that phone numbers do not necessarily belong to the owner (alone), but often to a tenant, family member, or caretaker. Based on these results, the self-reported endline information can be used to draw conclusions about potential mechanisms in this population.

An additional challenge, however, is selective attrition within the sample of individuals to be interviewed at endline. As compared to those who participated in the endline survey, attrited individuals are more likely to be female and more educated, have higher monthly income, live in different areas of the city, are more likely to indicate water as their least preferred service, and live in more valuable properties (see Table B9). Whereas the sample is still well-balance across treatment groups, it is not representative for the population of individuals interviewed at baseline (see Table A5).

Despite these differences in characteristics, the tax compliance results presented above largely hold true for the subsample of individuals who participated in the endline interview.[70] This does not guarantee that the channels leading to these effects are the same as for the remainder of the sample, but can serve as indicative evidence for the comparability of both samples with respect to their tax compliance behavior despite differences in observables.

[70] See Table B10 for a summary of results using the complete sample and Table B11 for the sample of individuals interviewed at endline only.

Table B9: Baseline characteristics of attrited (left) vs. non-attrited (right) individuals

Variable	N	(1) No Mean/SE	N	(2) Yes Mean/SE	T-test Difference (1)-(2)
Female	705	0.417	2684	0.349	0.068***
		(0.019)		(0.009)	
Age	701	5.399	2673	5.447	-0.048
		(0.096)		(0.050)	
Education level	696	4.409	2682	4.208	0.202**
		(0.086)		(0.045)	
In relationship	702	0.886	2683	0.873	0.013
		(0.012)		(0.006)	
Owner	705	0.652	2684	0.644	0.008
		(0.018)		(0.009)	
Monthly income	589	3.635	2577	3.423	0.212**
		(0.078)		(0.036)	
Ward group 1	705	0.128	2684	0.159	-0.031**
		(0.013)		(0.007)	
Ward group 2	705	0.106	2684	0.168	-0.062***
		(0.012)		(0.007)	
Ward group 3	705	0.123	2684	0.147	-0.023
		(0.012)		(0.007)	
Ward group 4	705	0.119	2684	0.120	-0.001
		(0.012)		(0.006)	
Ward group 5	705	0.251	2684	0.197	0.054***
		(0.016)		(0.008)	
Ward group 6	705	0.272	2684	0.208	0.064***
		(0.017)		(0.008)	
Distance to most preferred service (km)	705	1.587	2684	1.573	0.013
		(0.066)		(0.035)	
Distance to least preferred service (km)	705	1.658	2684	1.680	-0.023
		(0.058)		(0.030)	
Most preferred service: water	705	0.504	2684	0.509	-0.005
		(0.019)		(0.010)	
Least preferred service: water	705	0.106	2684	0.082	0.024**
		(0.012)		(0.005)	
Property value (SLL)	705	1.73e+07	2684	1.31e+07	4.20e+06***
		(8.71e+05)		(3.42e+05)	

Notes: The value displayed for t-tests are the differences in the means across the groups. ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

Table B10: Treatment effects on whether any tax was paid—entire sample

	Any tax paid by September 2022				
	(1)	(2)	(3)	(4)	(5)
Treated	0.012	0.013			
	(0.008)	(0.008)			
Match			0.012	0.012	0.009
			(0.009)	(0.009)	(0.017)
Mismatch			0.013	0.015	0.005
			(0.009)	(0.009)	(0.014)
Match × Water (m)					0.038
					(0.024)
Match × > median distance (m)					-0.019
					(0.023)
Match × Water (m) × > median distance (m)					-0.032
					(0.033)
Mismatch × Water (l)					0.093*
					(0.052)
Mismatch × > median distance (l)					0.008
					(0.017)
Mismatch × Water (l) × > median distance (l)					-0.052
					(0.068)
Service type FEs (m, l)	No	Yes	No	Yes	Yes
Distance linear (m, l)	No	Yes	No	Yes	Yes
Perceived availabilty at BL (m, l)	No	Yes	No	Yes	Yes
Wealth FEs	No	Yes	No	Yes	Yes
p-val. Match = Mismatch			0.949	0.778	0.838
Match + water (m) = 0					0.007
Mismatch + water (l) = 0					0.052
Match + > median distance (m) = 0					0.514
Mismatch + > median distance (l) = 0					0.297
Match + water (m) + > median distance (m) = 0					0.789
Mismatch + water (l) + > median distance (l) = 0					0.195
Obs.	5,494	5,469	5,494	5,469	5,469
Control mean	0.201	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B11: Treatment effects on whether any tax was paid—endline sample

	Any tax paid by September 2022				
	(1)	(2)	(3)	(4)	(5)
Treated	0.011	0.012			
	(0.011)	(0.011)			
Match			0.020	0.021	0.016
			(0.013)	(0.013)	(0.023)
Mismatch			0.004	0.005	0.008
			(0.013)	(0.013)	(0.019)
Match × Water (m)					0.064*
					(0.034)
Match × > median distance (m)					-0.026
					(0.029)
Match × Water (m) × > median distance (m)					-0.061
					(0.045)
Mismatch × Water (l)					0.086
					(0.070)
Mismatch × > median distance (l)					-0.013
					(0.024)
Mismatch × Water (l) × > median distance (l)					-0.069
					(0.086)
Service type FEs (m, l)	No	Yes	No	Yes	Yes
Distance linear (m, l)	No	Yes	No	Yes	Yes
Perceived availabilty at BL (m, l)	No	Yes	No	Yes	Yes
Wealth FEs	No	Yes	No	Yes	Yes
p-val. Match = Mismatch			0.246	0.230	0.766
Match + water (m) = 0					0.002
Mismatch + water (l) = 0					0.170
Match + > median distance (m) = 0					0.624
Mismatch + > median distance (l) = 0					0.785
Match + water (m) + > median distance (m) = 0					0.760
Mismatch + water (l) + > median distance (l) = 0					0.771
Obs.	2,684	2,664	2,684	2,664	2,664
Control mean	0.201	0.201	0.201	0.201	0.201
SD	0.401	0.401	0.401	0.401	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B12: Treatment effects on perceived government performance

	Original	Inverse prob.	Horowitz-Manski bounds	
		weights	Upper	Lower
	(1)	(2)	(3)	(4)
Performance of local government index	0.051	0.051	1.129***	-1.021***
	(0.036)	(0.038)	(0.050)	(0.049)
	[2,684]	[2,684]	[3,389]	[3,389]
- Confidence in ward councilor	0.032	0.024	0.727***	-0.648***
	(0.036)	(0.039)	(0.038)	(0.040)
	[2,684]	[2,684]	[3,389]	[3,389]
- Confidence in mayor	0.025	0.019	0.857***	-0.813***
	(0.037)	(0.040)	(0.046)	(0.043)
	[2,684]	[2,684]	[3,389]	[3,389]
- Confidence in FCC	0.050	0.066*	1.005***	-0.886***
	(0.037)	(0.039)	(0.047)	(0.045)
	[2,684]	[2,684]	[3,389]	[3,389]
- Approval of ward councilor	0.050	0.045	0.770***	-0.685***
	(0.036)	(0.039)	(0.039)	(0.041)
	[2,684]	[2,684]	[3,389]	[3,389]
- Approval of mayor	0.028	0.029	0.936***	-0.872***
	(0.035)	(0.037)	(0.049)	(0.045)
	[2,684]	[2,684]	[3,389]	[3,389]
Local government responsiveness (std.)	0.043	0.036	0.803***	-0.683***
	(0.033)	(0.035)	(0.042)	(0.041)
	[2,610]	[2,610]	[3,304]	[3,304]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B13: Treatment effects on perceived government performance (below median value properties)

	Original	Inverse prob.	Horowitz-Manski bounds	
		weights	Upper	Lower
	(1)	(2)	(3)	(4)
Performance of local government index	0.057	0.078	0.936***	-0.838***
	(0.054)	(0.056)	(0.071)	(0.070)
	[1,395]	[1,395]	[1,693]	[1,693]
- Confidence in ward councilor	0.041	0.065	0.644***	-0.489***
	(0.055)	(0.058)	(0.055)	(0.059)
	[1,395]	[1,395]	[1,693]	[1,693]
- Confidence in mayor	0.040	0.045	0.696***	-0.681***
	(0.058)	(0.060)	(0.067)	(0.064)
	[1,395]	[1,395]	[1,693]	[1,693]
- Confidence in FCC	0.043	0.057	0.818***	-0.744***
	(0.056)	(0.060)	(0.068)	(0.066)
	[1,395]	[1,395]	[1,693]	[1,693]
- Approval of ward councilor	0.041	0.046	0.646***	-0.555***
	(0.054)	(0.058)	(0.057)	(0.059)
	[1,395]	[1,395]	[1,693]	[1,693]
- Approval of mayor	0.043	0.064	0.751***	-0.741***
	(0.053)	(0.052)	(0.070)	(0.065)
	[1,395]	[1,395]	[1,693]	[1,693]
Local government responsiveness (std.)	0.057	0.062	0.673***	-0.537***
	(0.048)	(0.050)	(0.060)	(0.060)
	[1,368]	[1,368]	[1,659]	[1,659]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B14: Treatment effects on perceived government performance (above median value properties)

	Original	Inverse prob. weights	Horowitz-Manski bounds	
			Upper	Lower
	(1)	(2)	(3)	(4)
Performance of local government index	0.001	-0.011	1.332***	-1.251***
	(0.056)	(0.060)	(0.077)	(0.073)
	[1,289]	[1,289]	[1,696]	[1,696]
- Confidence in ward councilor	-0.000	-0.035	0.815***	-0.837***
	(0.057)	(0.060)	(0.057)	(0.059)
	[1,289]	[1,289]	[1,696]	[1,696]
- Confidence in mayor	-0.017	-0.022	1.027***	-0.979***
	(0.056)	(0.061)	(0.069)	(0.064)
	[1,289]	[1,289]	[1,696]	[1,696]
- Confidence in FCC	0.044	0.059	1.211***	-1.062***
	(0.058)	(0.061)	(0.070)	(0.067)
	[1,289]	[1,289]	[1,696]	[1,696]
- Approval of ward councilor	0.022	0.000	0.891***	-0.857***
	(0.056)	(0.059)	(0.059)	(0.060)
	[1,289]	[1,289]	[1,696]	[1,696]
- Approval of mayor	-0.031	-0.035	1.134***	-1.038***
	(0.053)	(0.058)	(0.074)	(0.068)
	[1,289]	[1,289]	[1,696]	[1,696]
Local government responsiveness (std.)	0.046	0.040	0.960***	-0.857***
	(0.050)	(0.055)	(0.063)	(0.061)
	[1,242]	[1,242]	[1,645]	[1,645]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B15: Treatment effects on perceived availability of and satisfaction with public services

	Original	Inverse prob. weights	Horowitz-Manski bounds	
			Upper	Lower
	(1)	(2)	(3)	(4)
Availability of public services index	0.034	0.034	1.261***	-1.216***
	(0.029)	(0.031)	(0.052)	(0.055)
	[2,684]	[2,684]	[3,389]	[3,389]
Availability of most preferred service (std.)	0.037	0.054	0.826***	-0.741***
	(0.038)	(0.040)	(0.041)	(0.044)
	[2,677]	[2,677]	[3,385]	[3,385]
Availability of least preferred service (std.)	-0.005	-0.005	0.772***	-0.770***
	(0.038)	(0.040)	(0.041)	(0.043)
	[2,665]	[2,665]	[3,371]	[3,371]
Satisfaction with public services index	0.058**	0.057*	1.106***	-0.955***
	(0.029)	(0.031)	(0.047)	(0.048)
	[2,683]	[2,683]	[3,388]	[3,388]
Satisfaction with most preferred service (std.)	0.065*	0.090**	0.753***	-0.627***
	(0.037)	(0.040)	(0.039)	(0.040)
	[2,676]	[2,676]	[3,384]	[3,384]
Satisfaction with least preferred service (std.)	0.071*	0.064*	0.738***	-0.589***
	(0.037)	(0.039)	(0.039)	(0.039)
	[2,669]	[2,669]	[3,376]	[3,376]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B16: Treatment effects on satisfaction with local level public service provision (below median value properties)

	(1)	(2)	(3)
	Satisfaction with public services index	Satisfaction with most preferred service (std.)	Satisfaction with least preferred service (std.)
Panel A. Overall treatment effect			
Treated	0.079*	0.074	0.043
	(0.044)	(0.054)	(0.052)
Panel B. Treatment effect by arm			
Match	0.071	0.090	0.062
	(0.050)	(0.064)	(0.060)
Mismatch	0.088*	0.052	0.019
	(0.053)	(0.063)	(0.063)
p-val. Match = Mismatch	0.737	0.526	0.468
Obs.	1,385	1,382	1,382
Control mean	-0.000	-0.000	0.000
SD	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B17: Treatment effects on satisfaction with local level public service provision (above median value properties)

	(1) Satisfaction with public services index	(2) Satisfaction with most preferred service (std.)	(3) Satisfaction with least preferred service (std.)
Panel A. Overall treatment effect			
Treated	0.044	0.038	0.101*
	(0.041)	(0.053)	(0.054)
Panel B. Treatment effect by arm			
Match	0.030	0.036	0.058
	(0.049)	(0.064)	(0.064)
Mismatch	0.063	0.042	0.163**
	(0.049)	(0.062)	(0.064)
p-val. Match = Mismatch	0.499	0.927	0.092
Obs.	1,279	1,278	1,277
Control mean	-0.000	-0.000	0.000
SD	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B18: Heterogeneous treatment effects by perceived availability at baseline

Dep. var.: Any tax paid by September 2022	(1)
Match	0.021
	(0.013)
A little/lot (m)	0.009
	(0.012)
Match × A little/lot (m)	-0.017
	(0.017)
Mismatch	0.010
	(0.014)
A little/lot (l)	0.013
	(0.012)
Mismatch × A little/lot (l)	0.007
	(0.018)
p-val. Match = Mismatch	0.474
p-val. Match + Match × Perceived availability (m) = 0	0.730
p-val. Mismatch + Mismatch × Perceived availability (l) = 0	0.159
Obs.	5,469
Control mean	0.201
SD	0.401

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), property value, and property ownership.

Table B19: Treatment effects on perceived government information on taxpayers

	Original	Inverse prob.	Horowitz-Manski bounds	
		weights	Upper	Lower
	(1)	(2)	(3)	(4)
Information on taxpayers index	0.035	0.043	1.755***	-1.644***
	(0.037)	(0.040)	(0.070)	(0.066)
	[2,683]	[2,683]	[3,387]	[3,387]
HH location	-0.013	0.001	1.389***	-1.351***
	(0.032)	(0.037)	(0.079)	(0.063)
	[2,683]	[2,683]	[3,387]	[3,387]
Compliance	0.048	0.052*	0.701***	-0.566***
	(0.029)	(0.030)	(0.041)	(0.039)
	[2,683]	[2,683]	[3,387]	[3,387]
Occupation	0.077**	0.074**	0.744***	-0.589***
	(0.035)	(0.037)	(0.040)	(0.045)
	[2,683]	[2,683]	[3,387]	[3,387]
Income	-0.038	-0.039	0.738***	-0.841***
	(0.031)	(0.033)	(0.043)	(0.049)
	[2,683]	[2,683]	[3,387]	[3,387]

Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B20: Heterogeneous treatment effects on tax compliance by informal public good contribution (September 2022)

	Amount paid (SLE)		Paid any tax (d)		Amount paid (SLE, cond.)	
	(1) Did not pay	(2) Paid	(3) Did not pay	(4) Paid	(5) Did not pay	(6) Paid
Treated	30.815	20.326**	-0.044*	0.019**	56.633	72.092
	(41.345)	(9.793)	(0.026)	(0.009)	(242.370)	(45.995)
p-val T S1=S2	0.093		0.051		0.247	
Match	-19.921	13.739	-0.048	0.019*	-377.200	19.028
	(40.313)	(10.742)	(0.031)	(0.010)	(468.303)	(46.805)
Mismatch	74.746	24.878*	-0.037	0.019*	241.131	115.693*
	(55.623)	(13.840)	(0.030)	(0.010)	(281.165)	(62.255)
p-val. Match = Mismatch	0.047	0.438	0.710	0.967	0.134	0.087
p-val. Match S1 = S2	0.177		0.075		0.502	
p-val. Mismatch S1 = S2	0.172		0.126		0.200	
Obs.	835	4,623	835	4,623	201	947
Control mean	150.260	111.937	0.249	0.191	603.125	584.729
SD	509.186	398.036	0.433	0.394	880.174	743.296

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B21: Heterogeneity of the treatment effects on perceived extractive capacity of the government depending on the content of the treatment message

	(1) Information on taxpayers index	(2) HH location	(3) Compliance	(4) Occupation	(5) Income
Match	0.073	0.052	0.040	0.047	0.012
	(0.069)	(0.052)	(0.059)	(0.071)	(0.056)
Match × > median distance (m)	-0.033	-0.052	0.036	0.055	-0.106
	(0.102)	(0.092)	(0.079)	(0.104)	(0.082)
Match × Water (m)	-0.040	-0.030	0.035	-0.087	-0.005
	(0.099)	(0.081)	(0.075)	(0.096)	(0.081)
Match × > median distance (m) × Water (m)	0.025	0.028	-0.203*	0.162	0.069
	(0.149)	(0.133)	(0.113)	(0.148)	(0.118)
Mismatch	-0.009	-0.024	0.041	0.039	-0.075
	(0.054)	(0.046)	(0.044)	(0.055)	(0.048)
Mismatch × > median distance (l)	0.074	0.008	0.027	0.086	0.032
	(0.074)	(0.070)	(0.060)	(0.076)	(0.062)
Mismatch × Water (l)	0.126	0.339**	-0.089	-0.043	0.074
	(0.182)	(0.160)	(0.151)	(0.174)	(0.107)
Mismatch × > median distance (l) × Water (l)	0.200	-0.242	0.251	0.114	0.258
	(0.250)	(0.194)	(0.201)	(0.258)	(0.207)
p-val. Match = Mismatch	0.240	0.184	0.984	0.910	0.144
p-val. Match + Match × > median distance (m) = 0	0.615	0.998	0.190	0.210	0.148
p-val. Mismatch + Mismatch × > median distance (l) = 0	0.293	0.795	0.152	0.037	0.390
p-val. Match + Match × Water (m) = 0	0.672	0.734	0.159	0.573	0.917
p-val. Mismatch + Mismatch × Water (l) = 0	0.506	0.039	0.743	0.981	0.998
p-val. Match + > median distance (m) + Water (m) = 0	0.751	0.980	0.165	0.031	0.619
p-val. Mismatch + > median distance (l) + Water (l) = 0	0.014	0.376	0.065	0.274	0.089
Obs.	2,663	2,663	2,663	2,663	2,663
Control mean	0.000	-0.000	0.000	0.000	0.000
SD	1.000	1.000	1.000	1.000	1.000

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Table B22: Treatment effects on perceived extractive capacity of the government (above median value properties)

	Original	Inverse prob.	Horowitz-Manski bounds	
		weights	Upper	Lower
	(1)	(2)	(3)	(4)
Information on taxpayers index	0.040	0.067	2.112***	-1.978***
	(0.056)	(0.060)	(0.108)	(0.103)
	[1,288]	[1,288]	[1,695]	[1,695]
HH location	-0.035	-0.016	1.716***	-1.579***
	(0.051)	(0.056)	(0.121)	(0.099)
	[1,288]	[1,288]	[1,695]	[1,695]
Compliance	0.053	0.076	0.837***	-0.690***
	(0.047)	(0.049)	(0.061)	(0.059)
	[1,288]	[1,288]	[1,695]	[1,695]
Occupation	0.058	0.058	0.839***	-0.765***
	(0.054)	(0.058)	(0.060)	(0.067)
	[1,288]	[1,288]	[1,695]	[1,695]
Income	0.008	0.021	0.908***	-0.991***
	(0.046)	(0.049)	(0.066)	(0.073)
	[1,288]	[1,288]	[1,695]	[1,695]

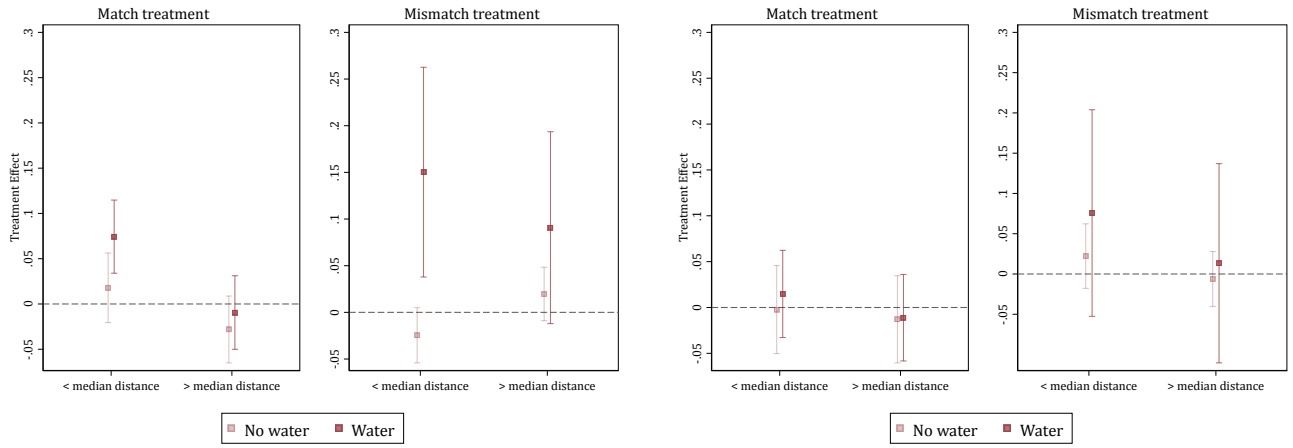
Notes: Robust standard errors in parentheses. Number of observations in brackets. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.

Table B23: Treatment effects on perceived extractive capacity of the government by treatment arm (above median value properties)

	(1) Information on taxpayers index	(2) HH location	(3) Compliance	(4) Occupation	(5) Income
Panel A. Original					
Match	0.051	-0.023	0.049	0.060	0.021
	(0.067)	(0.058)	(0.055)	(0.065)	(0.053)
Mismatch	0.026	-0.050	0.060	0.066	-0.022
	(0.066)	(0.064)	(0.055)	(0.067)	(0.057)
p-val. Match = Mismatch	0.745	0.734	0.511	0.516	0.769
Obs.	1,278	1,278	1,278	1,278	1,278
Panel B. Inverse probability weights					
Match	0.052	-0.018	0.061	0.032	0.033
	(0.072)	(0.065)	(0.057)	(0.071)	(0.056)
Mismatch	0.079	-0.011	0.093	0.092	-0.010
	(0.068)	(0.069)	(0.057)	(0.069)	(0.059)
p-val. Match = Mismatch	0.500	0.963	0.259	0.408	0.740
Obs.	1,278	1,278	1,278	1,278	1,278

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline. In addition, regressions control for public service preferences (most, least), linear distance to closest services (most, least), perceived availability of services at baseline (most, least), property value, and property ownership.

Figure B5: Heterogeneity of the treatment effects on whether any payment was made depending on the content of the treatment message (December 2022)



(a) Below median property value

(b) Above median property value

Table B24: Treatment effects on tax compliance (January – June 2023)

	(1) Amount paid (SLE)	(2) Paid any tax (d)	(3) Amount paid (SLE, cond.)
Treated	-5.668	0.003	-55.269
	(17.451)	(0.009)	(101.327)
Obs.	5,494	5,494	833
Control mean	134.538	0.149	900.524
SD	613.328	0.357	1354.032

Notes: Robust standard errors in parentheses. Significance levels are indicated by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Regressions include strata, enumerator, and ward fixed effects and control for respondent age, multiple property use, and the respective outcome variable at baseline.