The Limits of Government Outsourcing: Property Tax Re-Assessments in India

Esther Duflo, Rema Hanna, Benjamin A. Olken, and Shreya Tandon*

August 2026

Abstract

Theories of outsourcing often presume that private firms are more efficient producers of goods and services, but that contracting frictions mean that firms do not internalize the government's objective function. But perhaps, for some tasks, the government can actually be more productively efficient? We study this question through an experiment in Chennai, India, where the government randomly selected neighborhoods to have property tax assessments determined by private firms contracted by the government, rather than government tax inspectors. We show that inspections by government tax inspectors yielded almost double the increase in tax revenue as those done by private firms. They were also more accurate, as judged by comparing new assessments to independent assessments done by third-party surveyors we hired. Difficulties contracting with the government discouraged many competent firms from bidding. Perhaps as a result, government inspectors had higher skills and put in more effort than the workers at the private firms. Importantly, the government inspectors also had more authority to enter properties, which may be hard to transfer even to higher quality firms. The results suggest that the limits to government outsourcing may go beyond multitasking issues.

*This project was a collaboration involving many people. We thank our many colleagues in the Government of Tamil Nadu and the Greater Chennai Corporation, particularly, Minister Palanivel Thiaga Rajan, J. Radhakrishnan (then chairman of GCC), Gagandeep Singh Bedi, Vishu Mahajan and R. Lalitha, as well as many colleagues from the Revenue Department and Information Technology Department of the Greater Chennai Corporation for their support with project implementation, data collection, and data transfers. We thank Youssef Assarssah, Clotaire Boyer, Smitha Cunigal, Abhinav Kumar, Karan Makkar, Murali Mallikarjunan, Hasto Narendra, Santiago Andres Valdivieso Puertas, Shrddha Rajesh, and Varshini Sridhar for outstanding research assistance in Chennai and in Cambridge. We thank the Government of Tamil Nadu / J-PAL partnership fund, and the National Science Foundation (award number 2315488) for financial support.

1 Introduction

Governments face a tradeoff between performing tasks in-house with their own employees, or procuring these services from private firms. Classic theories, such as Hart, Shleifer, and Vishny (1997), frame this as a tradeoff between efficiency and multitasking concerns: even though private firms may be more efficient, have strong motives to cut costs, and have a greater ability to impose high powered incentives on their employees, contracting frictions imply that they do not face the full set of multi-dimensional incentives as the government, generating problems.

This is indeed a real tension in many circumstances (such as prisons or school meals)[1], but this setup sidesteps other issues that may be important in some settings. First, challenges in information transfer may play a role: while the government can relatively easily share information that is recorded in official data, sharing 'soft' information – e.g., knowledge held by government employees – with private firms may be much harder, making some tasks hard to outsource. Second, there are a wide range of private firms with varying quality, and it is not clear that all private firms are more efficient than the government. Identifying and contracting the right firms may be difficult. Moreover, the complexity and rigidity of the government procurement process, and the fact that payment schedules are often uncertain may make it less attractive for the best firms to bid. Third, certain types of tasks require the authority of the government to carry out, and it may be hard for the government to delegate this authority to private sector firms.

We study these issues in the context of outsourcing property tax assessments in Chennai, India, a common function that many local governments worldwide perform and fairly commonly outsource.[2] Determining the characteristics of a property for valuation is a common task: in the public sector, tax inspectors assess property characteristics to determine taxes, while private sector actors complete the equivalent task to facilitate commercial transactions and bank underwriting. At its core, the task is relatively simple and formulaic: measure

[1] The example in the original paper is prisons, where the private sector may be more efficient than the government, but it may be hard to contract over the full set of human rights outcomes the government may want to achieve.

[2] For example, municipal governments in Nigeria, Tanzania, and Uganda have all outsourced various aspects of tax collection; for a qualitative review of the experience of outsourcing tax in poor countries, see Haas and Manwaring (2017). Moreover, outsourcing property tax assessments is not limited to lower income countries; in rich countries as well, including the US, Australia and New Zealand, some municipalities outsource property tax assessment to private firms. For examples of outsourcing firms, see, for example, vsgi.com (a vendor based in Massachusetts) and www.qv.co.nz (a vendor based in New Zealand); see also the New South Wales, Australia property tax outsourcing website (https://www.nsw.gov.au/housing-and-construction/land-values-nsw/why-land-values/service-contracts).

the property and determine its characteristics and usage – how many square feet, how much retail space, and so on. However, in practice, it is often more complex because the task also involves high-stakes interactions with owners – owners may prefer to keep their taxes low and may try a variety of strategies to do so, from denying access to properties to assessors to attempting to bribe them (Khan, Khwaja, and Olken, 2016).

In our setting, the Chennai city government was interested in reassessing property taxes for about 280,000 properties, which they had identified as likely undervalued through a coarse GIS-based assessment done by drones. Such a reassessment exercise was long overdue. Chennai is a rapidly growing city and many smaller properties have been converted to much more valuable multi-storied buildings or changed their use. The government was concerned, however, that doing so would take a long time with their existing staff because their staffing levels are based on the typical year-to-year changes in properties that need to be checked, rather than a wholesale survey of every property. Therefore, they decided to experiment with outsourcing some of the reassessments to private firms to see whether this would be effective. They ran the procurement for these firms using standard Government of India procedures, which are similar to procurement procedures in many developing countries. Moreover, they experimented with two potential contracts: a fixed-fee contract, in which firms would be paid a fixed fee per property assessed, and an incentive contract, in which firms would be paid a small fixed fee plus a percentage of the increase in property tax found.

We worked with the government to randomize which of the 199 areas of the city (wards) would be reassessed by these firms and which would be conducted by regular government assessors, and among wards randomized to be outsourced, which wards would be done by firms with fixed vs. incentive-based contracts. Within each ward, we randomized individual streets into three waves for reassessment, or into a control group. We use this experiment to study whether the government and private firms had different levels of productivity.

The experiment ended up lasting for 8 weeks (it stopped earlier than initially planned for reasons we describe below). We therefore use the randomization assignment from the first reassessment wave as our randomized treatment variable. We examine both the number of properties that they re-assessed in a given period, and even more importantly, through an instrumental variables approach (using assignment to treatment as an instrument) the differential impact on property values conditional on being re-assessed by either government or private firms.

Overall, the government was much more effective than the private firms. The government inspectors re-assessed 25 percent more properties than the private firms, despite the

fact that they dedicated about half the man-days to the task (449 versus 808 man-days). Perhaps even more importantly, they also raised more revenue from the properties that they did re-assess. On average, the properties re-assessed by the government surveyors saw their property valuations increase by 84 percent, whereas those re-assessed by the private firms saw valuations increase by only 54 percent (p-value < 0.01). This then translated into more revenues collected: the increase in tax revenues – per property inspected – from inspections by government inspectors was almost double (87 percent higher) that from inspections done by private firms.

With such large increases in property valuations, one might worry that the government assessors were systematically over-valuing properties. Thus, for a random subset of properties that were inspected, we conducted our own independent assessment by sending our own set of independently trained surveyors to carefully measure properties. Comparing our independent measure and the official valuation, we find that the assessments by government surveyors are in fact *more* accurate than those by private firms, and that they are not systematically overvaluing re-assessed properties.

Part of why the government was more effective in raising revenue than the private firms is that they were better at assessing more complex and larger properties. We examine this in two ways. First, we examine the details of the property assessments. Property measurement requires both full access to the property and technical skill. Specifically, a team of two inspectors must enter the property, access the roof, and then measure the size of the roof with a tape measure.[3] For buildings that are just rectangular boxes, one can estimate the total floor area by multiplying the roof measurement by the number of floors. However, to the extent that the building is irregularly shaped, with different floors of different sizes, or has a mix of residential and commercial properties (which are differentially taxed), tax inspectors need to manually measure each room, draw a sketch of each floor of the property, and then add up the area of each room on each floor of each usage type. We find that while the government assessors do better on most dimensions relative to the private workers, they are particularly good at finding commercial usage and identifying buildings with irregular measurements – i.e. measurements where each floor is a different size. The ability in particular to correctly assess and tax the pieces of a mixed-use property that are commercial is an area where government inspectors have a particular advantage.

Second, we explore whether there is treatment effect heterogeneity based on properties' baseline tax valuations. We find that there is – treatment effects for government assessors

[3] Buildings in Chennai typically have flat, accessible roofs, which makes this possible.

are almost twice as large for the properties with the largest baseline valuations as for smaller properties, whereas treatment effects for private firms are essentially flat with respect to baseline property valuation. Note that the government assessors were slightly more likely to reassess these large properties. However, we show that differential composition of which properties were reassessed explains at most 4 percent of the difference between government and private reassessments, i.e. this difference is being driven by government assessors finding more previously untaxed square footage when they arrive.

The finding that government workers do better than private firms is striking, given most economists' priors, but not unheard of: in a very different setting (placement services for hard-to-place unemployed workers), Behaghel, Crépon, and Gurgand (2014) also find that government workers are more effective and cost effective than the private job placement firms that they hired. A nice feature of our setting and design is that we have rich data to explore why it is the case.

We first rule out two key theories of why the private sector was less efficient: (1) that the government had a temporary advantage in "soft" information that would go away as the firms got more experience and (2) that the firms would have been more effective if the government could have offered them higher powered incentives.

Two types of soft information – knowing which properties are most undervalued, and finding the locations– could give the government an advantage, at least in the short-run before the private firms get more experience. We do not believe this to be a substantial explanation of the difference in performance we observe. Importantly, given that the biggest issue was finding untaxed square footage conditional on visiting a site, we show that the government officials did not appear to have more private information than the assessors about which properties were undervalued. To examine this, we use the fact that, as discussed above, the government had previously commissioned an independent drone-based survey of properties. For this exercise, prior to our study, a separate firm had conducted an aerial survey of all of Chennai, demarcated properties, and created estimates of property demand for each property.[4] None of these data were shared with either the government assessors or the private assessors. We do not find that the government was systematically targeting highly undervalued properties relative to the private sector. On finding properties, the private assessors complained more than the government assessors in our surveys that properties were hard to

[4] These surveys were intended to be a diagnostic exercise at the aggregate level. At the individual property level, they have substantial signal but are noisy. While we can use them here as a statistical proxy for soft information - i.e. likely increases in tax assessment - the government could not use them as a substitute for a legal tax assessment.

find. But, they were also given GPS locations, and were more likely to use machines to find the properties. Thus, overall, it is unlikely that the government had an advantage in soft information that explained the large differential increase in assessed property values.

The evidence also suggests that the strong incentives for cost-savings in the private sector are not driving the difference, which is the type of concern raised by Hart, Shleifer, and Vishny (1997). For firms paid a piece rate for each audit they conduct, a cost-saving firm's incentive may have been to instruct their workers to do the job as quickly as they can, which would potentially explain the low accuracy, limited ability to detect complex properties, and lower overall assessments.[5] To examine how incentives could affect the behavior of the firms, the government recruited firms under two types of contracts – a fixed fee per-property contract, and an explicit output-based incentive contract, where the firms were also paid a percentage of the increase in revenue collected, and thus have more high powered incentives to find larger property area. Wards were randomly assigned to either the fixed fee per property or incentive contract firms. In general, we find no statistically significant differences between the fixed and incentive firms on the number of properties re-assessed, the amount of revenue raised, the accuracy of assessments, or the knowledge or effort dimensions that we observe. This suggests that to the extent that outsourcing is less effective at collecting revenue, it is not a problem that can be solved by a simple fix in the incentives provided in the contract.

What then drives the difference between the government and the private firms? We offer two explanations. First, a broader set of issues around procurement and contracting may have limited the types of firms that were attracted and ultimately contracted to undertake the activity. In selecting firms, the government followed the standard government procurement procedures, which essentially chose the lowest-price bid among firms meeting a set of qualification requirements.[6] Consistent with the idea that this is a common task, more than 20 firms initially expressed interest in the project and attended informational meetings. The vast majority of them, however, did not submit bids.

Challenges with the procurement process may suggest why: qualitatively, the firms reported that the project information was complicated and that the government often did

[5] We also examine whether there is some evidence that the private assessors may have engaged in straight *dereliction of duty*, for example by creating fake records or accepting bribes to lower the tax assessments. In a small taxpayer survey carried out with taxpayers, however, we found that those who were assessed were very likely to report being visited, and that reported bribes are very low in this setting.

[6] One aspect to note is that in their standard procurement rules, all contracted firms are paid the lowest price bid from any firm, and the government also negotiates to attempt to further reduce this price; we discuss the details of the procurement regime in more detail below.

not respond to their queries. In our firm survey, firms also reported frequent payment delays for past government projects, and only bigger firms –perhaps those that could take the risk– bid. Perhaps the firms were right to be worried – as discussed above, ex-post the reassessment contracts were canceled part-way through and firms were only paid much, much later.

In addition, the government demanded very low prices: firms that did not bid reported that the price at which they would have been willing to accept the contract was anywhere from 2 to 10 times what the government was willing to pay. Moreover, because the government procurement rules state that the government pays the lowest amount of any firm that bids – and then further attempts to negotiate this lowest price down – even for the firms that bid, the final contract price was about 75 percent less than the average initial bid. This kind of rigid procurement system with little discretion is very common. In an attempt to reduce the likelihood of corruption, bureaucrats are often not given much discretion to trade off price and quality (Decarolis et al., 2025). In the end, despite the government aiming to select 6 firms, only 4 qualified firms bid on the contract, and only 3 firms were ultimately willing to accept the prices offered by the government and agreed to the contract.

And finally, partway through the contract period (and in the midst of the first randomized wave of streets being reassessed by both government and private firms), the concerns of the firms that were hesitant to bid were proven to be prescient. After some citizen protests about new tax assessments, and faced with a looming election, the government changed course and abruptly decided to cancel the reassessment project entirely. The contracts with the outsourcing firms were suspended and all reassessment work, by both government and private firms, was halted. Payments for the work were then substantially delayed, with private firms only being paid 16 months later.

Consistent with having to complete the project at a very low price, the firms hired less skilled employees than the government. Through our survey, we find that government workers have more prior property measurement experience – an average of 11 years of experience, compared to just 1.4 years for private workers. They are also more educated – 43 percent of the outsourced inspectors do not have a college degree, compared with just 15 percent of government workers – so even with more years of experience, these lower education levels would represent persistent differences in skill. Given that calculating property areas of different types requires some mathematical ability and spatial training (i.e. making sketches of complex property layouts), this difference in skills may be related to the accuracy results.

Moreover, we find that government workers expended more effort overall. It is not that they are just more socially minded: We first examine (through our surveys) the public service

motivation of public and private sector workers through the Perry Public Service Module, and find no difference in public service motivation. In a lab effort task, we also find no greater willingness to exert effort. Despite this, we find that the government workers exert substantially more effort on the job: for example, conditional on entering a property, they spend about 35 percent more time measuring a property – about 31 minutes for a government worker, compared to about 23 minutes for a private firm worker. This may reflect taking the time to measure each room separately, for example, and may be related to why government workers are more accurate and uncovered more complex-shaped properties. This perhaps reflects both the different types of workers hired by the government and the private sector, as well as management practices of both.

Our second answer to the question of why the government assessors were more efficient is that even if the firms wanted to do more, the nature of the particular type of task made it difficult. Tax assessment is an official government function, which requires obtaining access to citizens' properties. Private firms were issued official letters from the government authorizing them to access properties, as well as formal identification cards and appropriate attire (caps, jackets). However, the government workers may have more confidence exercising this authority, and this may be hard to delegate – particularly in a potentially adversarial context such as tax collection, where property owners typically want to minimize their tax bills. In fact, we find that workers from outsourced firms report being refused entry to properties substantially more than government workers. And, related to the idea of effort above, conditional on being refused entry, government workers do not give up – they do almost 4 times as many revisits of refused properties compared to private sector workers.

All told, the experience here suggests a different set of constraints on outsourcing. The classic tradeoff set up by Hart, Shleifer, and Vishny (1997) is a principal-agent problem: private firms are more efficient, but shirk on non-measurable aspects of quality. Here, private firms are worse even on raising tax assessments, which the government can easily measure (and in fact, for a subset of private firms, explicitly contracted on), so this is not about non-contractable quality dimensions. We find a different set of issues instead, related to the difficulty of procuring government services of these kinds, and of delegating authority on a task that involves a conflict between the government and the citizens. This suggests that when governments are deciding which tasks to outsource, those that require authority over citizens (e.g. tax collection, pollution monitoring) may be less beneficial to outsource than those where government objectives are aligned with citizens (e.g. providing free meals to all, etc.). We believe these difficulties are not transitory, but fundamental to the task at hand.

On net however, while the government is substantially more cost-effective than the private firms, both bring significant revenue, for a fairly small outlay.[7] Even assuming that the increase in revenues from the reassessments lasts only for a single year, we calculate a net benefit per property of INR 5,758 (for the government) and INR 3,749 (for the firm) from reassessments. Since higher tax assessments likely last for years if not decades, returns on investment are commensurately higher – if the higher assessments last for a decade, this implies net benefits of INR 60,146 (government) and INR 39,556 (private firms). Costs per property are about INR 230 (private) and 285 (government), so in both cases, these are astronomical rates of return – between 1,600% - 2,000% if benefits last only one year, and between 17,200% - 21,100% if benefits last a full 10 years. This is because a tax reassessment takes about 30 minutes on average of an inspector's time (plus travel time), but increases tax payments by 30-50% for years. If the government is unable to assess all properties on their own, as they have other activities to do, rather than randomly outsourcing properties or wards, we also show that the government could do much better by allocating larger properties to the government assessors (who do much better with larger and more complex properties), and outsourcing smaller and less complex properties to the private firms.

This project contributes to both the literature on outsourcing public service provision as well as the literature on tax administration in developing countries. Prior empirical research on outsourcing public service delivery has largely focused on understanding which services are likely to be outsourced and where (e.g., Levin and Tadelis, 2010) and comparing the performance of outsourcing to government provision of various types of services (Galiani, Gertler, and Schargrodsky, 2005; Behaghel, Crépon, and Gurgand, 2014; Romero, Sandefur, and Sandholtz, 2020; Banerjee et al., 2019; Coube, Fontes, and Rocha, 2026). However, less is known about contexts that may be more adversarial such as taxation, where the tax assessor is tasked by the government with fairly assessing tax liabilities among many parties (who may want to minimize their tax payments). Theory suggests such 'adversarial' cases could be quite different (Banerjee, Hanna, and Mullainathan, 2012). For example, in the context of pollution audits, Duflo et al. (2013) find that the performance of private auditors is catastrophic, due to the perverse incentive they face to collude with the firms they audit, but there is no direct comparison of their performance with government workers.

With regard to taxation, much of the previous economics literature on property tax collection in developing countries has focused on how to make government tax inspectors

[7] For private firms, we know the monetary costs the government paid the firm; for government inspectors, who have responsibilities beyond assessing properties (i.e. issuing notices and collecting revenue), we know their wage and, for a subsample, how many properties they can do in a day.

more efficient: whether they should be incentivized (Khan, Khwaja, and Olken, 2016, 2019), the optimal assignment and team composition of tax collectors (Bergeron et al., 2022), the extent to which tax inspectors should have discretion in their assessments as opposed to being constrained by rules (Knebelmann, Pouliquen, and Sarr, 2026), improving tax compliance by engaging local elites (Balán et al., 2022), and the effects of varying property tax and enforcement rates (Brockmeyer et al., 2021; Bergeron, Tourek, and Weigel, 2024); see Besley et al. (2022) for an overview of the role of bureaucrats and development more generally. Much less, however, is known about outsourcing parts of the tax assessment and collection system entirely. Yang (2008)'s quasi-experimental study of pre-shipment inspection in customs suggests the potential for outsourcing to be effective, but it is not clear how the results from that context – an international, high-reputation firm being hired by a domestic customs agency to perform customs inspections in a different country – inform the question of whether a government should outsource its own regular tax assessments to domestic firms. There is also a small qualitative literature, outside of economics, which focuses specifically on government outsourcing of tax collection in Sub-Saharan Africa (see Fjeldstad, Katera, and Ngalewa, 2008; Iversen et al., 2006; Cheeseman and de Gramont, 2017). Outside of Lagos, Nigeria, which is generally held as a success in the qualitative literature (but where the subcontracted firm was associated with the mayor), these experiences have also been short-lived and relatively unsuccessful, for reasons having to do with contracting, legitimacy, and excessive prices (see Haas and Manwaring, 2017). Our results are consistent with these qualitative insights, but establish the results much more systematically, and in a context with higher government capacity.

The remainder of the paper is organized as follows. Section 2 describes the setting, formal procurement process, the experimental design, and the data we use. Section 3 explores the impact of outsourcing on tax assessments and tax revenues, and Section 4 explores mechanisms in more detail. Section 5 discusses cost-effectiveness. Section 6 concludes.

2 Setting, Experimental Design, and Data

2.1 Setting

We study urban property taxes in the city of Chennai, India. Property taxes are assessed and collected by the municipal government, officially known as the Greater Chennai Corporation (GCC). The GCC is part of the Government of the State of Tamil Nadu, and therefore their procurement practices and regulations follow local state law and procedures.

Property taxes are determined by a formula that considers property characteristics – particularly the floor area and land area – and usage. Each square foot of property in a given usage category (residential or non-residential) is assessed at a fixed street rate. Since the per-square-foot street rate is fixed exogenously by the government, the main determinant of a given property's tax is therefore its area and usage type.[8]

Tax inspectors measure property areas manually. Since most buildings in Chennai have flat, accessible roofs, a team of two inspectors usually begins by measuring the roof's dimensions with a tape measure. If the building's footprint is identical on all floors, inspectors can multiply the roof area by the number of floors to compute the total covered area.

While this process sounds straightforward, it is complex in practice. First, this procedure requires access to the property, and while inspectors are legally allowed to enter a property, the owners may not be home, or they may be resistant. Second, the properties may be irregularly shaped, requiring detailed measurements room by room on each floor. Third, the usage may not be obvious and may also be concealed since commercial areas are taxed at a higher rate. For example, in one visit that we conducted, inspectors discovered a workshop behind a closed door on the second floor of an otherwise residential building; a lax inspector may miss this. Fourth, there are a variety of edge cases that require detailed knowledge of the law – at what point does an awning or non-permanent roof become 'covered area,' how does one account for light shafts in the middle of buildings, or courtyards on an intermediate floor that are open to the sky, and so on. Thus, a complete property assessment takes some combination of authority to enter the property, effort, and skill.

Prior to our project, the GCC had concerns that many properties may have been under-assessed. The GCC does not re-assess properties at regular intervals to update its database, largely relying on property owners to report any changes to their properties.[9] In 2018, it commissioned an external firm to assess the degree of under-reporting and identify any properties that had not been assessed yet. To do so, this firm used drones to take aerial photographs of the city and then used GIS tools to demarcate properties (henceforth the 'drone survey'). Staff from the drone survey company conducted some field visits to supplement this information with the number of floors for each property and the property usage. While this survey is not usable for legally assessing taxes – the area of the property is based on an

[8] Small discounts are also given based on property age (3% for buildings built between 2007 – 2017, 6% for buildings built between 1997 – 2007, and 10% for older buildings) and for 'semi-permanent' structures, such as areas covered by awnings.

[9] In 2016, the GCC was contemplating a switch to a self-assessment based tax system, and offered taxpayers an opportunity to declare changes to their properties. However, the data were not used to update any of the property records as the GCC decided not to adopt the self-assessment based system.

approximation of the overall footprint, rather than a detailed hand-measurement accounting for irregularities – it revealed that 340,000 properties were likely substantially undervalued or not valued.

The large number of properties to be re-assessed meant that using regular government surveyors only to conduct the assessment could take a long time. Thus, the GCC was keen to test whether outsourcing the assessment to private firms could be an effective alternative.

2.2 Official Procurement Process

We worked with the Chennai government to experimentally test this idea. To do so, the government sought to contract firms through its standard procurement process, as per Government of India regulations. The government aimed to hire six firms, three with a fixed per-property contract, and three with an incentive-based contract based on the amount of additional tax assessments generated.

The procurement process began with an open call for bids. This started by listing a preliminary notice for tender on the Government of Tamil Nadu's website (posted on November 2, 2022). This preliminary notice had limited information, such as the earnest money deposit required to submit a bid, and stated a bid submission date of December 20, 2022. The tender document itself, which specified the full set of requirements, was posted several weeks later, during the last week of November 2022. In addition to posting this information on the city government's e-tendering website, the government publicized the tender in one Tamil and one English newspaper. They also identified (through the internet) and directly contacted firms that had specialized in property measurement, topographical surveys, GIS surveys, and property valuation about the tender.

For each contract, firms were then selected through a standard two-stage process. First, the government would conduct a technical review of each firm's bid to ensure they were qualified. Second, the government would choose firms with the lowest financial bid from among qualified bidders. Bidders were required to submit a technical proposal, including balance sheets, evidence of prior experience, and CVs of key personnel. In addition, they had to present their proposal in an in-person presentation to GCC officials. Only firms that had completed at least two similar assignments in the last ten years were eligible to submit bids.[10] Technical proposals were scored out of 100 based on the evaluation criteria in Appendix Table 1, and the GCC only opened financial bids submitted by technically-qualified

[10] Similar projects included enumeration, assessment, or collection of property tax for municipal governments, topographical, cadastral, land surveys for public- or private-sector clients etc.

bidders that received a technical score of 80% or higher. The financial bid depended on the contract type: for the fixed-price tender, the contract template asked firms to bid a fixed price per property assessed; for the incentive-based tender, applicants were required to bid a percentage of the additional tax demand (i.e. assessed values) that they would be paid on top of a fixed 50 rupee per-property fee, up to a cap of 6% of additional tax demand. As per the standard procurement rules, the GCC then negotiated with the firm with the lowest financial bid to reduce the price further. Once the final price was determined, all technically-qualified bidders were offered the opportunity to match the bid.

The chosen firms were responsible for visiting the list of properties supplied by the government and measuring the area and usage of each property. GCC assessors and private firm surveyors were provided the list of properties that they were supposed to re-assess, including the street address and GPS coordinates collected during the drone survey. To confer legitimacy to the private firms, and to let people know the private assessors were not scammers, etc., the GCC's decision to outsource the reassessment to private firms was announced in major newspapers as well as on the GCC's Instagram page. Surveyors hired by private firms were provided caps and jackets with the GCC logo, ID cards, as well as an official letter confirming they had been authorized by the GCC to enter and measure properties as part of a reassessment drive.

As described above, the government procured firms under two different contract types, with bids corresponding to the payment mechanism. In the *fixed fee* version, the government would pay the firm a fixed fee based on each successfully surveyed property. For this contract, firms' financial bids consisted of the per-property fee that they would charge the government. In contrast, in the *incentive* version, the government would pay the firm a fixed fee of 50 rupees per property, plus a fixed percentage of the increase of the tax liability uncovered. For this contract, firms' financial bids consisted of the percentage of the revenue increase that they would be paid. Firms could bid under either or both contracts.

In the end, the government selected two firms under the fixed fee contract and one firm under the incentive contract. Once chosen, the government conducted a training for the private firms' assessors in the second week of April 2023, consisting of an in-class part, where the measurement process was explained in full by the government staff, and a practical training, where the private firm assessors visited specific properties with the city government staff and practiced taking measurements. The contracts were finalized in the second week of February 2023, and firms began work in the second week of April.

2.3 Experimental Design and Sample Selection

The government used several approaches to construct the sample of properties for possible reassessment. First, as described above, it used the drone survey to flag properties where there were differences in covered area of at least 20 percent, or differences in usage. Second, it used a self-declaration exercise, where property owners were asked to declare the official size of their property; any deviations of 20 percent or more between self-declarations and official tax liabilities were also flagged. The government then removed government buildings, large apartment buildings, and several other categories that it deemed were too complex for reassessment by private firms and thus could not plausibly be randomized to private or government assessment.[11] This resulted in about 279,000 properties for possible reassessment that enter our sample.

We worked with the government to randomize which properties would be re-assessed, and whether they would be reassessed by government surveyors or private firms.[12] The randomization was conducted in two steps (Figure 1). First, at the ward level (i.e. geographic unit), we randomized each of the city's 199 wards to be reassessed by one of the three private surveying firms or by government surveyors, stratified by the 3 regions of the city (north, central, south); hence, 50% of wards were assigned to the fixed-fee private reassessment (25% to each of the two firms), 25% to the incentive-based private reassessment, and 25% to the government surveyors (see Appendix Figure 1). Second, within each ward, we randomized individual street segments to be assessed or not.[13] Within each selected ward, 40% of the street segments were randomly selected to be assessed, with the remaining 60% of street segments serving as a control group. These treatment street segments were further randomly divided, within each ward, into three lists, lists A, B, and C. Firms and GCC surveyors were both instructed to begin with properties on the first randomized list of treatment streets provided to them, 'list A'.

As described above, the government unexpectedly halted the reassessment process – by both private firms and government surveyors – on June 3, 2023, due to political sensitivities.

[11] The reassessment was supposed to also include un-assessed properties identified by the GIS. However, spot checks conducted by the city assessors concluded that many of these were vacant lots, or they could not be easily linked to the databases of property owners, so they decided not to include them.

[12] The project was originally designed to have a 'study phase' – to determine whether private outsourcing worked, and if so, which type of contract was more effective – and eventually a scale-up phase, whereby the most effective approach was scaled up. We examine the 'study phase' in this paper, and the scale up phase eventually did not take place.

[13] Note that if a street contained more than 20 properties, we sequentially subdivided the streets geographically until we were left with street segments of 20 properties or fewer each.

An election was coming up, and some local politicians started to complain about the new tax payments. At this time, both the firms and the government surveyors had only seen treatment list A. In our analysis, we therefore compare the properties on list A – which were given to the firms to reassess during the period the study was underway – to the 60% of streets in the control group, yielding a total sample of approximately 206,000 properties.[14] A small number of properties were reassessed by the government after the project halted; we discuss these below and account for them in the analysis using an intent-to-treat design.

2.4 Data

We use several data sources. First, we use *administrative data* on property tax assessments and payments. These data have information on tax assessments, including the detailed property characteristics (i.e. floor area by usage type, street rate) used in the tax formula, as well as the overall tax assessment (known as the 'tax demand'). We have data on property characteristics for both the baseline year and the year following the experiment. We also have administrative data on the actual tax demand and tax payments between 1999 and 2024, the year after the intervention.[15]

Second, we use data from four different surveys that we conducted. First, to assess tax assessment accuracy, for a random subsample of 2,150 properties that had been re-assessed by either the private firms or government enumerators, we conducted a *property survey*. In this survey, our survey team (two people per property) obtained an independent assessment of the property's valuation by measuring each property using the same procedures as the tax staff (i.e. start on the roof with a tape measure, then for irregularly shaped properties, measure each room separately, and so on). They also noted property usage (and the square feet of each usage type for mixed properties) and collected some demographic characteristics of the property occupant. We were able to find and survey about 70 percent of these properties, for a total of 1,499 properties; as shown in Appendix Table 2, we observe no evidence of differential response. These data allow us to compare the accuracy of private

[14] In our main analysis, we drop streets on study phase lists B and C, which would have been treated but were not due to the unexpected study halt, which yields a total sample of about 206,000 properties. In the appendix, we report alternative analysis including lists B and C as well; results are similar.

[15] We observe tax payments in two datasets. The Receipts Dataset records each individual tax payment, including the amount, payment date, and the tax year to which it is applied. The Demand, Collection, and Balance (DCB) Dataset is dynamically updated and reflects the current levels of tax demand, tax collection, and unpaid balances for each tax year as of the time the data extract is generated. For example, if a taxpayer makes a large advance payment in 2024, the DCB dataset may show positive tax collection for 2025, even though the Receipts Dataset would show no payment made in that year. Since we aim to measure payments made after the intervention, we rely on the Receipts Dataset for our analysis.

and government tax assessments.[16]

Second, we conducted an *assessor survey*. This survey was administered to both the government enumerators and the assessors employed by private firms. The survey includes demographic characteristics (e.g. age, education, prior experience), and information about the surveying process itself (e.g., the number of minutes spent on each property, refusals, repeated tries after refusals). The survey also collected a standard measure of public service motivation (Perry, 1996), and an experimental measure of effort – the 'A/B' task from DellaVigna and Pope (2018) – in which respondents are asked to alternately press the A and B keys on a computer as many times as possible during 1 minute. The survey was conducted between June and September 2023, after the project was halted. The GCC and private firms provided the names and contact information of employees who had been involved in re-assessing properties. We were able to successfully survey 82% of GCC assessors and 66% of firm employees. When we were unable to contact firm employees, it was largely due to incorrect phone numbers provided by the private firms.

Third, we conducted a *firm survey*. We attempted to survey all 34 firms that the GCC actively invited to bid on the tender – both those that ultimately bid on the project and those that did not. We surveyed 16 firms, including the 4 that submitted qualified bids. This allows us to capture additional information about the background of the firms, as well as to understand why those firms that chose not to bid did not do so.

Fourth, we have evidence from a small pilot survey that we conducted with 191 taxpayers, most of whom had been assigned to be assessed by private firms, where we asked them about their experience with the reassessment, including bribes (whether they had paid them themselves or had asked of others) and whether they had been visited at all by an officer. This was meant to provide inputs into the design of a larger taxpayer survey, but the government took the decision that surveying taxpayers after the project had stopped was too sensitive.

Finally, we have the approximate tax assessments from the *drone survey*. While these are not intended to be as accurate as an official firm assessment, on average, the difference between valuations on the drone survey and official property valuations may be indicative of which properties were most undervalued, and thus the highest priority for reassessment. Properties were short-listed for reassessment by comparing the property characteristics in

[16] We had initially aimed to collect data on satisfaction with the tax authority and city services, as well as any extra payments made by citizens during the reassessments. However, given the sensitivity around tax collection by the end of our endline survey, we were not given approval to ask these questions. Nonetheless, we have some information on these topics based on a survey of 191 pilot properties that we conducted.

the drone survey to the official tax records. Properties were categorized for reassessment (and hence included in our sample) if the property usage recorded during the drone survey placed them in a higher tax category than the usage recorded in the official data (mixed-use or non-residential vs. residential and non-residential vs. mixed), or if the estimated built up area in the drone survey was larger than the official records.

3 Impacts on Tax Assessments

3.1 Estimation

To estimate the impact of outsourcing, we use the fact that we worked with the government to randomly assign a) wards to be re-assessed by private firms (either under fixed or incentive-based contracts) or by government assessors, and b) within wards, to randomly choose which streets would be assessed.

Since the reassessment process was unexpectedly halted partway through the implementation, not all properties on streets that were randomly assigned to be re-assessed were completed. We first assess whether government workers and private firms differentially progressed through their assigned list for reassessment:

$$Reassessed_{isw} = \alpha_w + \beta_G^F AssignedGovernment_{sw} + \beta_P^F AssignedPrivate_{sw} + \gamma^F y_{isw}^0 + \upsilon_{isw} \quad (1)$$

Then we compute the intention to treat on tax collection and tax demand:

$$y_{isw} = \alpha_w + \beta_G^R AssignedGovernment_{sw} + \beta_P^R AssignedPrivate_{sw} + \gamma^R y_{isw}^0 + \omega_{isw} \quad (2)$$

Since only some streets within each ward were randomized to be re-assessed, we can include ward fixed effects α_w. We control for the pre-randomization value of the outcome variable, indicated by y_{isw}^0, when possible. Since the randomization to the government or private firms occurred at the ward level, we cluster the standard errors by ward.

Equations (1) and (2) are of independent interest: Equation (1) tells us whether the government workers were more or less effective at conducting the tax reassessment than the private companies that were contracted specifically for this purpose in the same period of time. Equation (2) tells us how much the government and the private firms manage to increase demand and collection. In addition, they form the first stage and reduced form for an instrumental variable strategy of the effect, for a particular property, of being re-assessed,

which we estimate as:

$$y_{isw} = \alpha_w + \beta_G ReassessedGovernment_{isw} + \beta_P ReassessedPrivate_{isw} + \gamma y^0_{isw} + \epsilon_{isw} \quad (3)$$

where $ReassessedGovernment_{isw}$ indicates that property i located on street s in ward w was re-assessed by the government surveyors and $ReassessedPrivate_{isw}$ indicates that the property was re-assessed by the private surveyors. We instrument for $ReassessedGovernment_{isw}$ and $ReassessedPrivate_{isw}$ with dummies for whether the street was randomly chosen to be re-assessed by the government or a private firm, i.e. $AssignedGovernment_{sw}$ and $AssignedPrivate_{sw}$. The control variables are the same as for the reduced form and first stages.

The results from Equation (3) show the impact of being reassessed by the government or a private firm, compared to not being reassessed at all. However, we are also interested in whether the government or private firms are more effective. We, therefore, report statistical tests of whether $\beta_G = \beta_P$ in all specifications.

To estimate the impacts of the two different types of private contracts (fixed vs. incentive), we also re-estimate equation (3) with separate variables for reassessment by fixed fee and incentive firms, instrumented by the randomization of wards and streets into private fixed, private incentive, or government reassessment. For these regressions, we report statistical tests for the difference between fixed and incentive reassessment.

We report tests for statistical balance based on baseline administrative variables in Appendix Table 3. We examine 10 variables, and regress each baseline variable on the government and private firm treatment indicators and ward fixed effects, clustering the standard errors by ward. We also report the p-value of the difference between the treatment effect for the private firms and the government. Out of the 30 differences computed (10 for government versus control, 10 for private versus control and 10 for government versus private), we find that 4 of them are significant at the 10 percent level, as expected by chance.

3.2 Completion Rates

We begin by examining what share of properties were re-assessed, and whether this varied between the government and private firms. Not only is this one measure of productivity, but this also serves as the first stage when we estimate the differential impact of a reassessment using equation (3).

We examine two measures of reassessments in Table 1. In Panel A, Column 1, we regress

an indicator for whether a property had been re-assessed as of August 2023, right after the experiment was halted, on the government and private assignment dummies. We find that the government reassessed 47 percent of the properties assigned as compared to 34 percent of properties assigned to private firms (p-value of the difference is 0.006).

In Panel A Column 2, we replicate Column 1 using data updated as of April 2024, which corresponds to when we observe updated tax demand and collections. We do so for several reasons. First, the government suggested that there was a delay in the reassessment data entry from the private firms, and so that some of the observed differences as of August 2023 may have been driven by lags; all of this lagged data entry was completed by April 2024. Second, as discussed above, the government completed a small number of reassessments after the outsourcing was halted; this appears to be very small (at most 313 reassessments, or less than 6.5 percent of the total).[17] Finally, since April 2024 is when we observe tax demand and tax payments, this serves as our first stage when we estimate equation (3) using instrumental variables. Using the updated April 2024 data, we find that the government surveyors re-assessed 47 percent of the properties they had been assigned, compared to 38 percent for private firms (column 2). This implies that the government was able to complete about 25 percent more properties (p-value 0.050), despite also having other responsibilities. This provides some initial suggestive evidence that the government workers may have been more productive.

Importantly, the GCC managed to complete reassessments for more properties despite dedicating fewer man-days to this task than the private firms. In our survey, GCC assessors and private surveyors report their daily working hours and the number of properties they re-assess. We use this to compute average time per reassessment, scale it by the total number of re-assessed properties (by GCC and private firms), and convert total man-hours into man-days by dividing by 8. Despite collectively spending 808 man-days on the reassessment project, the private firms managed to complete fewer properties than the GCC which only dedicated 449 man-days to the task. We explore the mechanisms underlying these productivity differences in more detail below.

[17] The April 2024 data reflects 313 additional properties measured by government assessors. Although our current data do not allow us to distinguish between reassessments completed by the government after the project halt and lagged data entry, this serves as an upper bound.

3.3 Overall Impacts on Tax Assessments and Collections

We next investigate whether this led to differences in actual tax assessments. We focus on two key outcome variables, tax demand for the 2023-2024 year (measured as of April 2024), i.e., the annual tax assessment of a property (Column 3), and actual tax collections for the same period (Column 4). Panel A of Table 1 shows that the additional tax demand from properties assigned to government workers was about twice as large as that from properties assigned to private firms (INR 2835 vs INR 1451), and tax collection was correspondingly much larger.

Using our instrumental variables strategy, shown in Panel B of Table 1, reassessment – by either government or private firms – leads to substantial increases in tax assessments and collections.[18] For those properties re-assessed by government workers, tax demand increases by an average of 6,014 rupees, or about 84 percent. For those properties re-assessed by private firms, tax demand increases by an average of 3,849 rupees, or about 54 percent. Thus, government workers increase tax assessments by an additional 30 percentage points (p-value 0.001). These differential increases in tax demand, in turn, translate into differential increases in tax revenue collected by the government (Column 4 of Table 1). Tax collections increase by 69 percent when re-assessed by government surveyors, but by only 37 percent when re-assessed by private firms. Thus, in short, a reassessment by government workers increased tax collections in the 2023-2024 tax year by 87 percent more than if a private firm reassessed the property.

In principle, the tax assessed is formulaic, so all differences in the tax assessed should come from differences in the input variables measured (i.e., square feet for residential use, square feet for commercial use). As an additional check, however, we also examine whether government assessors and private firms applied the tax formulas differently conditional on the areas they measured. As described above, the tax is supposed to be the residential and non-residential areas measured multiplied by a fixed street rate for each. We therefore recalculate the tax demand implied by applying the tax formula to measured areas in the administrative data (rather than using the official assessed value directly), and re-estimate equation (3). The results are shown in column 2 of Appendix Table 6; column 1 repeats the actual tax demand results for ease of comparison. The results are broadly similar, with very slightly higher demand for government assessors (85% increase in demand when we apply

[18] We replicate Table 1 omitting the baseline controls in Appendix Table 4, and with adding the B and C treatment lists as controls in Appendix Table 5. The conclusions are similar with these alternative specifications.

the formula, vs. 84% for official tax demand) and a more substantial increase for private assessors (66% when we apply the formula vs. 54% for official tax demand), suggesting that private assessors may have slightly under-applied the formulas even conditional on the areas measured. Nevertheless, even when we recalculate the tax demand using measured areas, the difference between government and private assessments remains large and statistically significant (p-value 0.04), so this is not the primary factor driving the differences.[19]

3.4 Accuracy

Just because the government assessors have higher assessments and raise more revenues does not necessarily mean that they are better. After all, the goal is that these higher assessments should be the result of government surveyors assessing properties more accurately, rather than just systematically overcharging.

To investigate this, we use the fact that we have independent measures of a random subsample of re-assessed properties that we obtained from our own survey. Following Khan, Khwaja, and Olken (2016), we define *accuracy* as

$$Inaccuracy_i = \frac{|\hat{y}_i - y_i^*|}{\hat{y}_i + y_i^*} \times 2 \tag{4}$$

where y_i^* is our own independent assessment of a property's valuation from our survey and $\hat{y}_i$ is the official assessment by the government or private sector surveyors. This captures the percent difference between the official valuation and our independent valuation. Note that $Inaccuracy_i$ measures the absolute difference between the independent property valuation and the official valuation, but it is non-directional. We also compute the directional equivalent, which measures systematic over- or under-valuation, as follows:

$$TaxGap_i = \frac{\hat{y}_i - y_i^*}{\hat{y}_i + y_i^*} \times 2 \tag{5}$$

[19] We also check whether each type of tax inspector differentially applied a change in tax formulas. Prior to our experiment, the government used to give a 25% reduction in tax for owner-occupied residential areas and a 10% reduction for owner-occupied non-residential areas. These reductions were canceled prior to the start of this experiment, but anecdotally may only have applied once reassessments took place. If this change in tax formula was applied equally by government and private assessors, as it should be, it would not explain the difference between them, though it may be related to why the overall increase in tax demand is larger in percent terms than the increases in area in Table 3. To examine this, in column 3 of Appendix Table 6, we recalculate the tax demand given measured areas, but this time we include these now-canceled rebates, so that un-assessed and re-assessed properties both receive rebates. Again government assessors increase demand much more than private assessors (p-value 0.02), but this brings the percent increase down from 84% to 59% for government assessors and from 54% to 41% for private assessors.

This captures the average directional percent difference between the official valuation and our independent valuation.

The results are presented in Table 2. Note that since we only surveyed re-assessed properties, in Table 2 (unlike in previous analysis), we are comparing properties reassessed by the private firms vs. those re-assessed by the government workers, with government reassessments as the omitted category, rather than control properties.

We find that the government assessors are substantially *more* accurate than assessments done by private firms, not less. Specifically, inaccuracy is about 14 percentage points higher (49%) for assessments done by private sector firms compared to government officials (p-value < 0.01). There is also no evidence of overcharging – in fact, column 1 of Table 2 indicates that both government and private inspectors are still somewhat undervaluing properties compared to what our independent survey concludes, though the private sector undervalues them by substantially more. On net, the key point here is that the differential increases in tax demand and government revenue we saw from government assessors compared to private assessors are associated with a higher rate of accuracy of assessments, not less, without any evidence of systematic overcharging.[20]

Importantly, the lower accuracy is not due to private assessors not going; in our small pilot survey for 191 properties, Appendix Table 8, we find that most properties assessed by private firms indeed had a visit for the assessment. Moreover, bribery does not appear to be a large channel either as reported bribes are low in this setting: only 9.5 percent of respondents had heard of officials demanding unofficial payments, and only 8.9 percent had heard of anyone making an unofficial payment.[21]

It is worth noting that private firms were less accurate despite having explicit incentives to maintain the accuracy of assessments. Contracts signed by firms stated that at each payment milestone, government assessors would cross-check a 10% random sample of properties re-assessed by firms. If the area deviation between government and firm estimates exceeded 10%, the firm would have to pay a penalty, determined by the share of properties with substantial deviations.[22] Moreover, the government reserved the right to terminate the

[20] Appendix Table 7 shows similar results using an alternative formula for tax gap and inaccuracy.

[21] Unfortunately, we have very little direct evidence on whether government officials or private sector employees differentially asked for informal payments or bribes to change the level of the tax assessment; these are sensitive questions, and we were not authorized to ask them in our larger endline, so we cannot compare these rates between government and private employees.

[22] Penalties were based on the share of randomly sampled properties with a discrepancy >10% at each payment milestone. For each milestone, the GCC would deduct: 2% of the value of the payment due if 1–2% of properties had a discrepancy; 4% of the value if the share was 2–4%; 6% if the share was between 4–6%, and so on. If over 10% of randomly sampled properties had discrepancies, the government would deduct

contract if a firm's assessment quality fell below the required standard.

3.5 Complexity and Size

Part of why the government assessors may be more efficient at raising revenue is that they are better at assessing complex properties or larger properties. We examine these hypotheses below.

3.5.1 Are government assessors better at more complex properties?

Recall that tax assessments are determined by a formula that essentially multiplies the floor area in square feet for each usage type by a fixed multiplier based on the property location. For reassessments to change valuations, they therefore must change area, usage, or some combination. Some of these changes are relatively easy to observe – walking up to a property, one can easily count the number of floors the property has, for example. Some changes are harder – capturing differences in area requires carefully measuring each room one by one, especially if there are different use types in the building which must be measured separately, or if the building is irregularly shaped (particularly with different shapes on different floors).

Table 3 therefore decomposes the changes in tax assessments.[23] We first examine total area, and find that it increases by 572 sq feet if reassessed by government, whereas it increases by 476 sq feet if reassessed by private firms (p-value of difference = 0.074). In addition, both the government and private firms find more floors–0.57 floors for government vs 0.49 floors for private assessed properties (p-value of difference 0.15; Column 4).

We then disaggregate area by residential area (which is taxed at a lower rate) and non-residential area (which is taxed at the higher rate). We find that residential property area increases for both government assessors (506 sq feet, or about 46% of the control area surface) and the private firms (438 sq feet), with a p-value of the difference of 0.157 (Column 2). However, as shown in Column 3, we observe a large difference for the higher taxed non-residential property: government assessors find about 40 more square feet (about 20%), while we observe no effect for the private firms (we can reject that these are the same with a p-value of 0.008). Since non-residential property is taxed at a much higher rate, this difference is an important driver of the difference in tax assessments between government and private assessors.

20% of the payment owed to the firm.

[23] Appendix Table 9 provides the reduced form estimates, while Appendix Table 10 provides the estimates using the B and C lists additionally as controls. The conclusions are the same.

Consistent with these findings, it is also worth noting that changes in the type of usage within a given property are more likely with government assessors. Columns 5 and 6 show if usage type is 'upgraded' or 'downgraded,' respectively, where 'upgraded' indicates moving to a usage type with a higher tax rate (typically commercial or mixed use) and downgraded means moving to a usage type with a lower tax rate (typically residential). Column (5) shows that while reassessments of both types lead to properties being reclassified at higher rates, this is almost twice as likely for government compared to private sector assessors (4.2 percent vs. 2.5 percent; p-value 0.04). Downgrades also happen with reassessments, but these are less likely with government than private assessors (0.9 percent vs. 1.7 percent; p-value 0.021).

Together, these facts suggest that some of the differences between government and private assessors come from discovering and reporting features that are more highly taxed. This typically requires a more in-depth assessment than single-use buildings, as each room needs to be measured separately.

To probe this idea further, we define an "irregular-shaped assessment" as one in which different floors of the building are recorded with different areas (see Column 7 of Table 3). Recall that for a standard, rectangular building with the same size floors, assessors just measure the area of the rooftop, and use that area for each floor; by contrast, an "irregular-shaped assessment" requires measuring each room separately and constructing the area of each floor from the room-by-room assessments. In the control group (i.e. prior to reassessments), only 45 percent of properties have irregular-shaped, i.e. non-uniform-by-floor, assessments. This increases substantially with reassessments, but more so when done by government surveyors (24pp vs. 19pp; p-value 0.057). Combined, this suggests that government surveyors are not only finding more area, but they are also doing more sophisticated assessments that result in higher taxed features.

3.5.2 Targeting and Heterogeneity by Baseline Property Value

The results presented thus far show that government assessors appear to generate larger treatment effects from reassessments than private assessors, based on the random assignment of wards to government vs. private assessors. At every step, it seems that the private assessors err in the direction of a lower tax bill for the taxpayers: they measure slightly smaller property areas, are less likely to identify commercial real estate, and even conditional on this information, slightly under-apply the official formula.

To understand the sources of this difference, it is important to understand whether

there is important heterogeneity in treatment effects across different types of properties. If government and private firms chose to focus on different sets of properties within a given ward, and there are heterogeneous treatment effects based on which type of properties were chosen for reassessment, then part of the difference we are capturing could reflect the fact that government vs. private assessors could lead to different properties being assessed, and therefore different local average treatment effects. The reduced form would still reflect the causal effect of assigning a property to the government rather than to the firm, since the choice of where to work is obviously itself an outcome, but the IV could not be interpreted as what would happen if the government vs the firms were assigned to assess a complete set of properties and did (as was originally planned) complete the assignment.

Moreover, if there are important heterogeneous treatment effects, that could suggest that an ideal assignment mechanism could consider using each type of surveyor where their comparative advantage is highest. This section explores these issues.

To begin, we first examine whether, in fact, government and private assessors differentially focus on different types of properties. Figure 2, Panel A plots the results from a series of locally-weighted regressions of the probability of reassessment by both government and private firms against log baseline tax demand; the right hand column also plots the difference between the two. To do so, we divide the x-axis into 100 equally sized points. For each point, we run a separate weighted regression, with weights given from a triangular kernel with width $\frac{1}{4}$ of the total width of the x-axis centered at that point. We then plot the results of each of these regressions as a function of the x point. We overlay a histogram of the data in the background.

These non-parametric regressions show that the government surveyors do appear to differentially target larger properties to some degree. Specifically, the difference in the probability of reassessment between government and private firms is negligible for the smallest properties at baseline, but rises to about 15 percentage points for the largest properties.

To see whether this differential targeting explains part of the treatment effects we saw above, we then estimate whether there are also heterogeneous treatment effects for different types of properties. If treatment effects were homogeneous, then the fact that different types of properties were targeted would not matter, but if the difference in treatment effects between government and private surveyors is also heterogeneous, then we could be capturing different local-average treatment effects.

To investigate this, Panel B of Figure 2 repeats the same locally-weighted regression procedure for our main instrumental variables treatment regressions. That is, for each point

on the x axis, we re-run our regression from equation (3), using the same local kernels as in Panel A. In this regression, one can think of the non-parametric regression results shown in Panel A as capturing the differential first stage as we move across different values of baseline log tax demand, ensuring that we are capturing the different treatment-on-treated estimates at each point on the x-axis. As before, the left panel shows the results separately for government and private reassessment compared with the control group; the right panel shows the difference between treatment effects for government vs. private firms.

Panel B shows that indeed, there are substantial heterogeneous treatment effects. While both government and private sector assessors find substantially more tax demand everywhere, their effects are similar for the smallest properties. But while treatment effects for private reassessment are essentially flat with respect to baseline tax demand, they rise dramatically for government assessors, so that the treatment effects of government reassessment are more than double for the largest properties. Note that since these are all locally weighted IV regressions, these are capturing heterogeneous treatment-on-treated effects at different levels of baseline demand levels, and do not reflect the differential reassessment rates shown in Panel A. We will return to the possible interpretation of this profile (in term of what it tells us about firm behavior) in the next section.

These heterogeneous treatment effects have several implications. Qualitatively, it is important to note that Figure 2 Panel B shows that the treatment effects for government assessors lie substantially (and statistically significantly) above the treatment effects for private assessors for almost the entire mass of the distribution. This means that qualitatively, the result that government assessors are more effective holds true essentially everywhere, and hence also holds true on average as well.

The combination of heterogeneous treatment effects shown in Panel B *and* the heterogeneous targeting shown in Panel A could mean that the average differences we observe could slightly overstate the differences we would observe if all properties were reassessed. We can, however, quantitatively calculate average treatment effects that we would obtain if both government and private firms investigated the same set of properties by integrating the heterogeneous treatment effects from Panel B of Figure 2 over the distribution of properties shown in the histogram, holding the distribution the same for both private firms and government assessors. This moves us from estimating two different local average treatment effects over potentially two different complier sets, which we have in Table 1, to estimating average treatment effects holding the set of compliers fixed. When we do this, we find impacts that are in fact broadly similar to the LATEs we estimated in Table 1 – we find that the average

impact of a government reassessment is Rs 6,043 and a private impact of Rs. 3,979. This shows that while there are heterogeneous treatment effects, because the differential probabilities of reassessment in Panel A of Figure 2 are not that large, on net, differences in selection of who is reassessed do not substantially affect our results.[24]

These results are broadly similar because the probabilities of reassessment are fairly similar across the two treatment groups. But the heterogeneous treatment effects also suggest that the government assessors have a *comparative advantage* at reassessing high value properties, and so an intentional assignment rule could take advantage of this to increase average treatment effects overall. That is, if the government has a limited number of government assessors available and wants to supplement with private firms, the results suggest having the private firms be assigned to take on properties with low baseline values, and government assessors be assigned take on those with higher baseline values, could lead to much higher revenue allocations.

To investigate this, we consider instead what would happen if we assigned the smallest 50% of properties (based on baseline value) to the private firms, and the largest 50% of properties (again based on baseline value) to government re-assessors. This 'optimal assignment' would yield a combined average treatment effect of Rs. 5,350 per property, compared with an average treatment effect of Rs. 5,011 if we randomly assigned half of the properties to government and half to private firms. Thus, if the government is constrained in how many government assessors it has and needs to use some private firms to supplement, it could obtain an additional 7 percent more revenue (5350/5011) by optimally assigning the larger properties to government assessors, at no extra cost. We will come back to this when we discuss cost-effectiveness.

4 Mechanisms

The results thus far show that reassessments by government workers appear to dominate those by private firms: they complete more properties, are more likely to correctly identify commercial real estate, handle more complex cases, and generate higher tax revenues – and are more accurate, not less, when judged by independent surveyors.

[24] Note that this calculation in fact changes two things compared to the LATEs estimated by the IV regressions: a) it holds the set of properties assessed constant across both treatment groups and b) it considers all properties, not just those that were reassessed. If we consider (a) alone – i.e. holding the set of properties constant but not switching from actually reassessed properties to all properties, we find that the differential selection shown in Figure 2 Panel A combined with the heterogeneous treatment effects shown in Figure 2 Panel B explains only about 4 percent of the difference between government and private assessors.

In this section, we explore a series of possible mechanisms to understand why government surveyors appear to have been more effective. We first rule out two key explanations for why the government assessors are more productive: that they have more soft information that gives them an advantage in the short run, or that their incentive structure incentivizes cost minimization over increasing tax assessments. We then offer two explanations that drive the observed differences between the government and the private sector: broader procurement and contracting challenges, and the nature of the task being outsourced.

4.1 A (Perhaps Temporary) Advantage of Soft Information?

We begin by exploring the role of soft knowledge. Local information may be important to decide which properties are under-assessed and hence could have the highest returns to target, as well as to locate properties. This advantage may, of course, change over time as firms gain more experience.

Given that government assessors were able to find more square footage, it is possible that the government assessors had more soft knowledge than the private assessors of what properties would have the highest returns, and thus were able to figure out which properties had higher expected returns and to target their efforts accordingly.

To examine this, we examine whether the assessors are differentially assessing those with a higher potential *gain* in demand. We can construct this using the estimate of property demand from the drone survey. To isolate the role of 'soft information' we regress this estimate of demand from the GIS drone survey on baseline demand and consider the residuals; this captures that portion of expected increase in demand that cannot be predicted based on the 'hard information' available in the tax information. Recall that neither the GCC staff nor the private companies had access to the predicted demand from the GIS survey, so this is a proxy for soft information that assessors may have had from, for example, knowing the neighborhoods and hence knowing that a property listed as residential actually had some commercial activities.

The results are shown in Figure 3, which shows the results from the type of non-parametric regressions used in Figure 2, but where we consider heterogeneity in soft information (residualized demand from the GIS drone surveys) as the key dimension of heterogeneity. As shown in Panel A, we do not find any evidence that properties that the drone survey suggests were under-valued were more likely to be re-assessed by either the government or private surveyors. In Panel B, we do show that indeed those properties with higher residual demand have higher increase when reassessed – so this measure has content and does in

fact predict gains – but this is not systematically different between government and private assessors.

Second, we examine whether the government had an advantage finding locations. From our qualitative research, we learned that the firms encountered many challenges both finding properties and conducting the property measurements. The private sector assessors noted inaccuracies in both the property centroids and their door numbers, leading them to spend considerable time trying to find them, as well as leading them to visit the same streets multiple times (rather than complete all properties on a street at once). However, this may also have been due to the poor planning of the firms, that could have looked up and asked questions about all locations at the start.

As shown in Figure 4, the private firms were more likely to state that they had difficulty finding properties than the government assessors (p-value < 0.001). However, in the end, they were given GPS locations, and they appeared to use technology to overcome this issue: about 75 percent of the private firm assessors used GPS to locate the properties, as compared to 17 percent of the government assessors.

On net, the evidence here suggests that while government workers may have had some advantage finding properties quickly, it is unlikely that a (temporary) advantage of soft information drives the results that we observe since there is no differential targeting based on soft information.

4.2 Incentives and Effort

4.2.1 Do Government Workers Exert More Effort?

In general, government workers do not face strong explicit financial incentives – they have de facto tenure until retirement in their jobs, and promotions are based on seniority, degree qualifications, and qualifying exams rather than performance on assessments. They receive fixed salaries which are not explicitly tied to performance, and they do not have monthly assessment targets.[25] While there can be incentives to perform in the form of softer incentives (better postings within the city, perhaps, as in Khan, Khwaja, and Olken (2019), or informal recognition by leadership), a priori, one does not tend to think of government workers in this context as highly incentivized. Non-monetary incentives such as medals and awards are limited, and typically only awarded to one or two assessors per year. And, yet, one may imagine that people enter government jobs due to a motivation to serve the public good

[25] The GCC imposes monthly revenue targets for tax collectors who are supervised by assessors.

(Dal Bó, Finan, and Rossi, 2013). However, as shown in Figure 5, we observe no difference in the public service motivation of workers in the government versus those at private firms.

However, despite the limited government incentives and the similar levels of public service motivation, government workers clearly put more effort into the process. Specifically, we ask the surveyors the number of minutes they spent on each property. Table 4, column 4 shows that government surveyors spend 35 percent longer per property than private surveyors – 31 vs 23 minutes, respectively (p-value < 0.001). This is again consistent with our administrative data, which showed more "complicated" assessments (with different area values per floor) and more accurate assessments.

Interestingly, the differences in the effort in measuring properties appear to be task-specific effort, rather than reflecting a general propensity for effort. To test generalized effort, in our survey, we included the "A/B" task, from DellaVigna and Pope (2018). This task measures effort by asking the respondent to press the 'A' and 'B' keys on a computer keyboard, in sequence, as many times as possible for 1 minute. Figure 6 shows the results of the A/B task – which shows that actually private assessors do 34.5% more key presses than government assessors (p-value of the difference is 0.02). The effort differences we are detecting here, then, appear to be specific to the task in question rather than a pure capacity for effort or generalized motivation per se. Thus, it is on this particular task that the private assessors were unwilling to exercise effort, and perhaps speaks to how the assessors were motivated and managed on the job.

4.2.2 Would Higher Powered Incentives Have Made a Difference?

Firms may face perverse incentives: if the contract paid firms per property assessed, they could have an interest to go quickly from property to property, which may be a reason for decreasing their accuracy, for avoiding the complex cases, and expending little effort. Indeed, the fastest way to perform reassessment would be to not even go visit the property, but apply a mechanical increase in valuation.

To explore whether stronger output based incentives for private firms would make a difference, we use the fact that, within the set of properties which were being assessed by private firms, the government randomized properties to be assessed either by firms with a fixed fee contract or with a linear incentive contract, wherein firms were paid 4.8 percent of the change in tax demand for all properties re-assessed.[26] The size of these incentives meant

[26] Note that the incentives to firms translated into different incentives faced by employees, but somewhat strangely, they did not bring the revenue-percentage sharing down to the employee level. The fixed firms

that firms were paid about the same on average per property re-assessed, but the fixed-fee firms were paid a fixed rate per property re-assessed, whereas the incentive firms received almost all of their pay from the linear incentives based on tax demand. To see this, note from Table 1 that private firms on average raised tax demand by Rs. 3849. Their average pay per property was therefore Rs. $50 + 0.048 \times 3849 = 235$, very similar to the Rs. 224 per property paid to the fixed fee firms.

Table 5 presents the results on the key outcomes we examine – number of properties re-assessed, tax demand, tax collections, and accuracy – where we separate out assessments by fixed-fee private assessors vs. incentive-based private assessors based on the ward-level randomization.[27] We find no statistically significant differences in performance between fixed fee and incentive-based firms on any of the dimensions we examine. As shown in Columns 2 and 3, while the incentive firm has slightly higher demand and collection than the fixed firm in magnitude, we cannot reject equality between them, and we can reject equality between both private contract types and the government.

While it is possible of course that larger incentives could have generated larger effects, the incentives here were such that the total cost to the government was about the same in both treatments, and at that level of incentive, we find no differences between the two approaches. The bottom line is that the firms' low performance does not come from a basic design flaw in the contract (which is to ignore incentives).

4.3 Procurement and Contracting Challenges

What did explain the differences, then? We begin in this section by exploring challenges in procurement that may have led the highest quality firms not to bid. We then in the next section discuss how the adversarial nature of the task may also play an important role.

4.3.1 Procurement in Practice

Procurement is often treated as a black box, but we had the opportunity to follow the entire process from the tender document to final payment. We qualitatively observed several challenges with the procurement process. The government held two pre-bidding meetings in December 2022 with prospective firms to provide details of the project deliverables and to answer any of the firms' questions. A total of 23 firms attended one or both of these two

paid a flat salary to their surveyors, while 89% of incentive firm surveyors reported that they were being paid a fee for each property they re-assessed (between 50-100 rupees per property).

[27] In Appendix Table 11, we also examine the specific property characteristics.

meetings. The firms asked questions about the pre-qualification criteria, the techniques for measuring properties, the difference between the fixed and incentive-based contracts, and the support that the chosen firms would receive from the government to enter properties.

The meetings revealed that the incentive contract created additional uncertainty for firms. Since firms in this contract would be paid a percentage of increased tax demand, the amount they would be paid in rupee terms would depend on the extent of under-taxation. Firms accordingly asked various questions during the bidding meetings about the expected revenues from this contract; firms were directed to conduct their "own research" to determine the expected gains they could make.

Despite the fact that 23 firms attended at least one of the bidding meetings, in the end, few firms bid on the contract. The stated reasons for this varied. Some firms that we talked to suggested that the prices per property that the government would pay were too low. Past experience was another reason: one firm that we talked with declined to bid because they had outstanding payments from the city government for another project. To further facilitate the bidding process, and allow for more firms to enter, the government extended the deadline from December 20, 2022 to January 2, 2023. In the end, only five firms completed the bidding process.[28] Three firms submitted the documents to bid for the fixed contract only, while two firms bid for both the fixed and the incentive contract. One of the firms that bid for both contracts did not meet the pre-qualification criteria, and was deemed ineligible, leaving four firms remaining.[29]

After submitting the bids, the four firms presented their work plans to the city government in person. One firm (the firm that bid for both contracts) demonstrated a clear understanding of the project requirements; the other three firms presented workplans that exceeded the scope of work, including only hiring engineers (while the task required measuring properties and computing room areas, the government felt that a fully trained engineer was not needed) or conducting surveys to understand taxpayers' "socio-economic behavior." Thus, the GCC panel re-emphasized the project scope of work, and acknowledged a need for a price negotiation so that the prices paid were consistent with the reassessment task.

For the four firms that bid for the fixed contract, the firm that also applied to the incentive contract was the lowest bid at INR 225 (USD 2.70) per property; the other three

[28] Other firms started the process of applying, but then did not submit all the required documents (despite continued communication from city officials to provide assistance). Given the low number of firms that bid, there was a discussion about whether to re-open the tender to allow more firms to bid, but the government ultimately decided against this.

[29] The firm was unable to provide evidence of having completed two property measurement projects of a similar nature within the past ten years.

bids were INR 960, INR 975 and INR 1000.[30] For the incentive contract, the one eligible firm bid 4.8 percent (i.e., it would be paid 4.8 percent of the change in assessed property demand); it is worth noting that this firm, despite bidding for the incentive contract, also stated a preference for the fixed contract.

The government then negotiated the prices with the bidding firms. Essentially, the government used the lowest-priced bid on the fixed price contract as the reference bid for all firms, and demanded that the other firms meet this price. In the end, the one firm that bid for the incentive contract agreed to the contract at 4.8 percent, while two firms agreed to the fixed contract at INR 224 (almost exactly the lowest bid obtained by the government, but only about 25 percent of what these two firms had initially bid themselves). The final firm felt that the proposed price was too low and dropped out.

There were thus several *de facto* challenges with the bidding process that may have deterred firms from bidding and led to lower quality firms being selected. First, as indicated above, the bid submitted was the reference price and the opening to a subsequent negotiation, rather than the actual price that would be paid as in a typical first-price auction. This negotiation is allowed under government regulations if the government perceives the lowest bid is above the prevailing market price (and government workers may have felt pressure to procure at the lowest price possible). This may have created additional uncertainty during the bidding process, as bidding firms could not be sure they would ultimately be hired at the price they were willing to accept.

Second, there was a lack of clarity on some aspects of the contract, which also could have been a deterrent to the degree that firms anticipated this ex-ante. For example, in 2022, the government raised property tax *rates* per square foot. The government did not specify ex-ante that the incentive firm would be remunerated with a percentage of the additional demand generated through the reassessment, calculated based on the tax rates existing *before* a 2022 tax hike. This caused issues down the road, when the firm expected to be compensated on the change in revenue calculated using the higher, post 2022 tax hike rates. Similarly, the government did not provide projections of firms' expected revenues from the incentive contract, despite being asked for this information during the pre-bid information sessions. Uncertainty about potential earnings and lack of familiarity with performance-based contracts likely discouraged firms from bidding on the incentive contract.

Note that these types of issues are not unique to Tamil Nadu, a state known for its

[30] The high bids may have also been because many of the valuation companies had experience dealing with very high-value, commercial properties (e.g. ports, industrial estates), which require highly skilled staff, like engineers, and take quite a bit of time.

high capacity within India. Instead, this type of ex-post ambiguity in the contract reflects precisely the general types of contracting frictions outlined by Hart and Moore (1988). When procuring something more complex than basic supplies, there is a tension between the desire of government to get the lowest possible price and to avoid manipulation by the procuring agents, and the difficulty to hire competent firms and give them the right incentives. In the absence of clarity on the price to be charged in this instance, there are also strong incentives among bureaucrats put in charge of the procurement to extract the lowest price possible from firms, to avoid being accused of colluding with them.

In the end, despite trying to recruit 6 firms, the government was only able to recruit 3. More than that, as discussed earlier, the firms that were skeptical of working with the government perhaps had good reason. Partway through the contract period, in early June 2023, the government abruptly canceled all reassessments, both by government assessors and private firms. Some taxpayers complained to their ward councilors (elected representatives of each ward) that only small properties were being targeted, and that large property owners were being let off the hook. The ward councilors raised these issues to the relevant state government officials, who, in turn, instructed the city to halt the reassessments. General elections were coming up in 2024, and some within the government may have been concerned about political pushback prior to the elections, especially since a previous property tax rate increase (in 2022) had already created some unhappiness.

The firms, left in limbo, lacked clarity on the possibility of resuming the exercise and the duration of the hiatus. They also had not been paid at all at this point. Firm surveyors, who had been actively engaged in fieldwork at the time of the halt, were instructed to cease operations. The firms attempted to discuss this issue with the city government, but the relevant official was on leave. Firms were therefore uncertain about what to do with the staff they had hired for the project – terminating them was undesirable if they were restarting anytime soon, but they also could not absorb them temporarily in other projects given that they had been hired for this project. They were ultimately not paid for 16 months.

4.3.2 Comparison of Bidding and Non-Bidding Firms

Given how the bidding process unfolded, what do bidders versus non-bidders look like? Is it in fact the case that there is differential selection in who actually chose to bid on the project?

Table 6 provides descriptive statistics to compare the firms that ultimately bid on the project with those that did not based on a survey that we conducted. Column 1 provides means for bidders, while Column 2 provides means for non-bidders.

Firm characteristics are presented in Panel A. Both the bidding and non-bidding firms were large, established firms, with an average firm age of 22 years. However, in general, those that chose not to bid were smaller, both in terms of employees (mean of 42 employees for non-bidding firms, compared with 73 employees for the bidding firms) and revenue.

Many of the firms had experiences working with government, but their experiences were mixed (Panel B). Those that bid had more government experience: Seventy-five percent of bidders had previously won a government tender, compared to 43 percent of non-bidders. Interestingly, virtually all firms that had prior government experience – bidding and non-bidding firms alike – reported delays in getting paid by the government. It is plausible that the delays in payment may be more costly for smaller firms (who may be less able to absorb these delays), which may be related to why smaller firms were less likely to ultimately bid on the project.

We next asked firms about what their proposed bids would be for the reassessment contracts, had they bid. Figure 7, Panel A graphs the distribution of bids (i.e. the price per property proposed) for the fixed contract; for comparison, we also graph the actual bids and demark the final contract price in a dotted line. In Panel B, we do the same for the incentive contract where a bid is the percentage of the increase in tax demand given as a bonus payment. Both figures show that the proposed bids would have been much higher than the actual price chosen. In fact, 64% of surveyed firms (including three of the contracted firms) would have bid higher than the actual price in the fixed contract, and almost all of the firms would have done so in the incentive contract. Thus, perhaps few firms bid because they believed it would not have been profitable for them to do so.

4.3.3 The types of workers that were hired

One reason why the firms could have been performing poorly is that they were generally paid much less than they had bid for the job (since the government negotiated with firms to get all of them to accept the lowest bid) and therefore needed to keep costs low given the price being paid. This is one thing that could have prompted them to exercise zero or low effort, and perhaps to hire less experienced or skilled workers (despite these kinds of workers being available in the private market).

In Table 4, we use the assessors survey to explore whether there are observable characteristics that are consistent with the idea that government assessors are higher skill. First, in Column 1, we show that government assessors are much more experienced than the workers hired by private firms to perform the same task. Government assessors have an average of

11 years of work experience surveying properties; private sector assessors, by contrast, have been doing this task for an average of only 1.4 years (p-value < 0.001). While work experience can accumulate over time, and hence this advantage of the government would only be transitory, government workers also have more education. In particular, 85 percent of government assessors are college graduates (with 35 percent having a masters), compared to only 57 and 6 percent for private assessors, respectively (p-values < 0.001). The combination of higher education, and more experience, may be related to why they are more accurate in their assessments. It is possible that the procurement process, by selecting the lowest bidding firms (without adjusting for the quality of the staff hired) led firms to hire staff at a lower salary than that which the government would have paid if it had hired them directly.

4.4 Authority and the Nature of the Task

An important challenge in outsourcing tax assessments is authority – making sure the private workers had the same authority to enter and assess properties as government assessors.

The government took several steps to accomplish this. First, ID cards, jackets and caps with the city government logos were procured and distributed to the assessors, so that they could officially identify themselves – although in practice, not all of them reported receiving all these items, leading to some firms printing some of them at their own expense, and not all private inspectors wore the caps and jackets at all times. Second, all private surveyors were provided a letter on the official GCC letterhead stating that they had been authorized to enter and measure properties as part of the GCC's reassessment drive. Third, the government posted photos of the private inspectors alongside government inspectors on their Instagram page, so that people could verify that the private inspection process was legitimate. All properties identified for the reassessment exercise were also published on the city government website so that taxpayers could check that their property was indeed scheduled to be reassessed.[31]

Nevertheless, the lack of perceived authority still appears to have been a constraint for private sector workers. The private sector assessors expressed frustration at wandering around trying to find properties in the summer heat, sometimes encountering unleashed dogs, only to be not let in and verbally harangued by the citizens when they found the property.

Table 7, Column 1, shows that refusals – whereby a property owner refused access to

[31] The government initially also planned to send text messages to the taxpayers that were identified for reassessment, followed by auto-rickshaw announcements and paper notices. The government did not ultimately do this; instead, the fixed contract firms independently purchased speakers and initiated their own announcements in some areas.

the property – were 27 percent more likely for private assessors compared to government assessors – 19 percent vs. 15 percent of properties (p-value 0.07), even though both had the legal authorization to demand entry to a property for tax assessment purposes.

Perhaps related to this, private surveyors also reported more pushback from property owners about their assessments. Private surveyors reported that owners were 28 percentage points (80 percent; p-value < 0.001) more likely to request a change to the recorded area that the surveyor found, and 12 percentage points (60 percent; p-value < 0.1) more likely to request a change to usage type. While we do not know whether or not assessors actually made changes in response to these requests, it suggests that they were more susceptible to this type of pushback than government assessors.

Finally, as shown in Column 5, the government surveyors appear to try substantially harder if denied entry on the first visit – they revisit 2 times on average then as well, compared to only 0.59 times for private surveyors (p-value < 0.01). They also are less likely to close a file as non-collectible and take an alternative action (column 6) when refused entry. Importantly, whereas both the government and private surveyors revisit if no one is home, the government employees may have felt more authority to go back and revisit if denied entry.

5 Cost effectiveness and implications

Finally, we turn to considering the cost-effectiveness of these two alternative approaches. For private firms, we know the monetary costs the government paid the firm, so determining costs is straightforward. For government inspectors, from our survey data, we can calculate an approximate amount of time that they spend on conducting reassessments, and then multiply it by their salary.[32] We report estimates under different assumptions for how long the increases in revenue last.

For each alternative, we calculate the net benefits to the government, i.e. $NetBenefits = E[\Delta Revenue] - E[Costs]$. As long as $NetBenefits > 0$, the reassessments are expected to be profitable for the government. If the government is not budget constrained in how much

[32] Specifically, we first take the number of hours they work in a day and multiply by the fraction of days that they work in a week on reassessments to have the total time in a day conducting reassessments. We divide that by the average number of reassessments that they complete in the day, and multiply it by their wage. We had many refusals for time spent per day working, so we have fewer observations to calculate the cost. Alternatively, we also calculated costs and net benefits using the time spent on the assessment, and we find similar costs (about INR 199). Given that this measure omits the time spent on finding the property, despite having less data, we use the full calculation above.

it can allocate to this task – which presumably it is not, as long as it produces positive expected revenues – it should choose whichever approach has the highest *NetBenefits*.

Table 8 presents the results. Under all assumptions for the duration of expected benefits, the expected net benefits are substantially larger using government assessors than private firms. For example, if results last 5 years, the net benefits from using government enumerators are 52% larger than from using private firms.[33]

More generally, it is worth noting that both approaches –government or outsourcing– have extremely positive returns. For example, if results last 5 years, the net benefits to the government are INR 29,930 for the government and INR 19,663 for the firms. If the higher assessments last for 10 years – plausible given that properties in Chennai had not been systematically reassessed in years – net benefits are INR 60,146 and 39,556 respectively.[34] Costs per property are about INR 230 (private) and 285 (government), so in both cases, these are astronomical rates of return – between 8,550% - 10,500% if benefits last only five years, and between 17,200% - 21,100% if benefits last a full 10 years. Thus, while the returns are substantially larger from using government surveyors, both approaches are staggeringly cost effective.

Given that they are much better at the task, one can argue that the government should perform it. Moreover, it is not clear that they would have to hire more workers to do so. When they are not helping to collect on tax bills twice a year, the government workers often have downtime in which they engage in other government activities, e.g. flood control. Thus, the assessments could be done during these non-collection periods. On the other hand, given that both the government and the private firms have very high net benefits, one could argue that if the government is constrained by manpower, it should devote as many people to the task as possible, and then hire a private firm to do the rest, potentially planning to conduct a reassessment faster for the properties that have been allocated to private firms. Another option would be for the private firms to be assigned "easier" properties (if they could be identified), or smaller properties where mistakes are less costly in terms of the future revenue stream, and to let the government employees handle the larger or harder properties. For example, as discussed earlier, there are returns to assigning government workers the highest baseline value properties (as they are good at complex properties) rather than a

[33] Appendix Table 12 reports the analogous cost-benefit calculations using tax collection.

[34] Given that this was the first large-scale reassessment drive implemented by the GCC in history, it is certainly plausible that these higher valuations would last for 10 years or more. Previous exercises, conducted sporadically between 2007 and 2022, have focused exclusively on commercial properties and were limited to approximately 10,000 properties.

random assignment. Table 8 shows that in this case, optimally assigning government workers to larger properties where they have larger treatment effects increases the net benefits by about 7 percent compared to randomly assigning some properties to government and some to private firms.

6 Conclusion

Economic theories of outsourcing often highlight the tension between a more efficient private sector that can "get things done" much better than the government, but that is stymied by contracting failures that end up providing the wrong incentives to firms. While these are important issues, this paper highlights several other obstacles to effective outsourcing of some government activities.

In this paper, we experimentally compare in-house property tax assessments conducted by the government to assessments that had been outsourced to private firms. Tax assessments, especially in "campaign mode" to bring the assessments up to date, are fairly often subcontracted to private firms. Yet, in this context, we find that the government was more effective than the private firms: not only did the government employees assess more properties over the same time period and with fewer man days – despite having other responsibilities as well – but they also raised the valuation of reassessed property more. The government employees were more accurate with each measurement, including complex ones, while private firms seem to err on the side of lower reassessment at almost every step.

Prima facie, the private firms should have had incentives to be as quick as possible (possibly at the cost of cutting corners and thus being less accurate); the government's solution to this concern was therefore to experiment with incentivizing the firms based on the amount of additional revenue they generated. However, interestingly, the specific contract incentives did not matter much: the fixed fee contracts for private firms led to similar revenue gains as an incentive contract which paid out based on the increased revenue from the reassessments. The fact that moving from flat fee payments to revenue-based incentives did not change much suggests that these types of incentive issues may not have been at the heart of the outsourcing challenge.

Rather, we find two other classes of potential explanations for the differences between the in-house and private assessments. First, some of this difference may have been driven by how the procurement and administrative processes selected firms. Despite the initial interest, very few firms actually ended up bidding. Many of the firms were frustrated by

frequent payment delays for past government projects. In fact, those that did not bid also believed that the prices that the government was willing to pay were not worth the effort. In retrospect, this was perhaps practical foresight, as the government unexpectedly halted the project early on in the contract due to perceived political pressures, and the firms that were contracted by the government ended up facing long delays on being paid for the work they did complete. As such, given that the firms that ended up bidding agreed to work for very low pay (and likely uncertain prospect of getting their payment on time) it may not have been surprising that the government employees were more educated, had more experience, and exerted more effort in the tasks than the temporary workers that the firms recruited to complete the contract.

Second, government officials also had more authority with citizens to enter properties and interact with citizens for the inspections, despite the government trying to confer authority onto the private firms' workers.

Taken together, this suggests that the limits to government outsourcing may go beyond multitasking issues, and instead that there may be certain tasks – such as tax assessments as we study here – where government provision is actually more efficient than the private sector. Developing a clear understanding of which tasks fit this pattern, and which do not, is an important direction for future research. These results also should lead us to question the frequent conclusion drawn by academics, policy makers, or citizens, when they observe that the government appears to be very inefficient at some of the tasks it does, compared to private firms. In effect, this is an apples to oranges comparison: many of the tasks the government undertakes are particularly difficult to carry out (because after all, no one wants to pay taxes, or to follow pollution or safety regulations), which might explain why governments do them slowly and seemingly inefficiently. And we see that at least in this context, private firms are worse than the government when trying to carry out the same tasks.

In fact, the main difficulty in collecting taxes in this setting was perhaps not the technical ability to collect the relevant information, but rather the political willingness for the government to enforce its own tax code. As noted, several months into the reassessment exercise, after some taxpayers complained about the reassessment and in advance of an election, the entire exercise was stopped, even though it was very clear by then that this would be a very significant source of revenues. While most political economy analyses focus on changes to the tax code, this suggests that understanding the political economy of tax enforcement itself may be an important direction for future work. Indeed, in this context, the recent tax

code change may have made enforcement politically more difficult: some local politicians complained that it was too much to do a reassessment soon after a change in the tax rate. In retrospect, if there is a tradeoff between the two, focusing on enforcement would likely have brought in considerably more money, given the generalized under-valuation of the properties.

References

Balán, P., Bergeron, A., Tourek, G., and Weigel, J. L. (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):762–797.

Banerjee, A., Hanna, R., Kyle, J., Olken, B. A., and Sumarto, S. (2019). Private outsourcing and competition: Subsidized food distribution in indonesia. *Journal of Political Economy*, 127(1):101–137.

Banerjee, A., Hanna, R., and Mullainathan, S. (2012). Corruption. In Gibbons, R. and Roberts, J., editors, *The Handbook of Organizational Economics*, pages 1109–1147. Princeton University Press.

Behaghel, L., Crépon, B., and Gurgand, M. (2014). Private and public provision of counseling to job seekers: Evidence from a large controlled experiment. *American Economic Journal: Applied Economics*, 6(4):142–174.

Bergeron, A., Bessone, P., Kabeya, J. K., Tourek, G. Z., and Weigel, J. L. (2022). Optimal assignment of bureaucrats: Evidence from randomly assigned tax collectors in the drc. Technical report, National Bureau of Economic Research.

Bergeron, A., Tourek, G., and Weigel, J. L. (2024). The state capacity ceiling on tax rates: Evidence from randomized tax abatements in the drc. *Econometrica*, 92(4):1163–1193.

Besley, T., Burgess, R., Khan, A., and Xu, G. (2022). Bureaucracy and development. *Annual Review of Economics*, 14(1):397–424.

Brockmeyer, A., Estefan, A., Ramírez Arras, K., and Suárez Serrato, J. C. (2021). Taxing property in developing countries: Theory and evidence from mexico. Technical report, National Bureau of Economic Research.

Cheeseman, N. and de Gramont, D. (2017). Managing a mega-city: learning the lessons from lagos. *Oxford Review of Economic Policy*, 33(3):457–477.

Coube, M., Fontes, L. F., and Rocha, R. (2026). Public services under private management. *The Quarterly Journal of Economics*.

Dal Bó, E., Finan, F., and Rossi, M. A. (2013). Strengthening state capabilities: The role of financial incentives in the call to public service*. *The Quarterly Journal of Economics*, 128(3):1169–1218.

Decarolis, F., Fisman, R., Pinotti, P., and Vannutelli, S. (2025). Rules, discretion, and corruption in procurement: Evidence from italian government contracting. *Journal of Political Economy Microeconomics*, 3(2):213–254.

DellaVigna, S. and Pope, D. (2018). What motivates effort? evidence and expert forecasts. *The Review of Economic Studies*, 85(2):1029–1069.

Duflo, E., Greenstone, M., Pande, R., and Ryan, N. (2013). Truth telling by third-party auditors and the response of polluting firms. *The Quarterly Journal of Economics*, 128(4):1499–1545.

Fjeldstad, O.-H., Katera, L., and Ngalewa, E. (2008). Outsourcing revenue collection: Experiences from local government authorities in tanzania. Technical report, Research on Poverty Alleviation.

Galiani, S., Gertler, P., and Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, 113(1):83–120.

Haas, A. and Manwaring, P. (2017). Enhancing local tax revenues: private or public collection? Technical report, IGC Cities that Work Policy brief.

Hart, O. and Moore, J. (1988). Incomplete contracts and renegotiation. *Econometrica*, 56(4):755–785.

Hart, O., Shleifer, A., and Vishny, R. W. (1997). The proper scope of government: theory and an application to prisons. *The Quarterly Journal of Economics*, 112(4):1127–1161.

Iversen, V., Fjeldstad, O.-H., Bahiigwa, G., Ellis, F., and James, R. (2006). Private tax collection—remnant of the past or a way forward? evidence from rural uganda. *Public Administration and Development*, 26(4):317–328.

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

Khan, A. Q., Khwaja, A. I., and Olken, B. A. (2019). Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings. *American Economic Review*, 109(1):237–270.

Knebelmann, J., Pouliquen, V., and Sarr, B. (2026). Discretion versus algorithms: Bureaucrats, tax equity and acceptability. IFS Working Paper 26/34, Institute for Fiscal Studies.

Levin, J. and Tadelis, S. (2010). Contracting for government services: Theory and evidence from us cities. *The Journal of Industrial Economics*, 58(3):507–541.

Perry, J. L. (1996). Measuring public service motivation: An assessment of construct reliability and validity. *Journal of public administration research and theory*, 6(1):5–22.

Romero, M., Sandefur, J., and Sandholtz, W. A. (2020). Outsourcing education: Experimental evidence from liberia. *American Economic Review*, 110(2):364–400.

Yang, D. (2008). Integrity for hire: an analysis of a widespread customs reform. *The Journal of Law and Economics*, 51(1):25–57.

Table 1: Difference between Properties Re-Assessed by Government Assessors and Firms on Tax Demand and Collection

	Reassessed (Aug. 2023)	Reassessed (Apr. 2024)	Tax Demand (Rs.)	Tax Collection (Rs.)
	(1)	(2)	(3)	(4)
Panel A: Reduced Form Estimates				
Assigned to government	0.473***	0.471***	2835.14***	1849.01***
	(0.043)	(0.042)	(360.65)	(269.28)
Assigned to private	0.338***	0.377***	1450.73***	792.64***
	(0.024)	(0.022)	(188.45)	(99.15)
Government = Private	0.006	0.050	0.00	0.00
Panel B: Instrumental Variables Estimates				
Reassessed by government			6013.83***	3922.35***
			(527.41)	(424.10)
Reassessed by private			3849.03***	2104.53***
			(372.92)	(226.11)
Government = Private			0.00	0.00
Control Mean	0.000	0.001	7120.43	5654.41
Observations	206255	206255	206255	206255

Note: This table experimentally estimates the difference in the likelihood of re-assessment, tax demand, and tax collection for properties assigned to government tax assessors and private firms. Reassessed (Aug 2023) is an indicator variable capturing whether the property was re-assessed by August 2023. Reassessed (Apr 2024) is an indicator variable capturing whether the property was re-assessed by April 2024. Tax Demand (Rs.) is the property tax liability in the 2023-24 tax year. Tax Collection (Rs.) is the property tax paid in the 2023-24 tax year. Panel A reports reduced-form effects of random assignment to government vs. private assessors on each outcome. Panel B reports 2SLS estimates of the effect of reassessment on tax outcomes, instrumenting reassessment with treatment assignment. All specifications include ward fixed effects and columns (3) and (4) include controls for the outcome variable at baseline. Tax demand data come from the DCB dataset and Tax Collection data come from the Receipts dataset. We impute tax collection as zero if there is no receipt associated with that property in 2023-24 in the Receipts Data. The sample is limited to the first list of properties the agents were instructed to re-assess and all control properties. Standard errors are clustered at the ward level and displayed in parentheses. We provide p-values for the test of equality between the government and private firms. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Accuracy of Reassessed Properties (OLS)

	Tax Gap (1)	Inaccuracy (2)
Assigned to private	-0.107^{**}	0.138^{***}
	(0.051)	(0.038)
Government = Private	0.04	0.00
Government Mean	-0.09	0.28
Observations	1,499	1,499

Note: This table estimates differences in non-revenue outcomes for properties re-assessed by government tax assessors versus private firms. Tax Gap is the difference between the official Tax Demand for the 2023–24 tax year and our estimated Tax Demand, divided by the sum of the two and multiplied by 2. Inaccuracy is the absolute value of Tax Gap. Official Tax Demand data come from the DCB dataset. We estimate OLS regressions where the omitted category is properties re-assessed by the government. Standard errors are clustered at the ward level. The row *Government = Private* reports p-values from tests of equality between government and private firms. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Impacts of Reassessment on Property Characteristics (IV Estimates)

		Area (sq. ft.)			Property Usage		
	Total	Residential	Non-Residential	Floor Count	Usage Upgraded	Usage Downgraded	Irregular-shaped Property
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reassessed by government	572.087***	505.917***	41.310***	0.572***	0.042***	0.009***	0.241***
	(44.32)	(39.55)	(9.00)	(0.04)	(0.01)	(0.00)	(0.02)
Reassessed by private	476.212***	438.222***	11.016	0.494***	0.025***	0.017***	0.188***
	(30.32)	(26.94)	(7.15)	(0.03)	(0.01)	(0.00)	(0.02)
Government = Private	0.074	0.157	0.008	0.149	0.04	0.021	0.057
Control Mean	1336.860	1107.833	204.099	1.869	0.015	0.004	0.445
Observations	206255	203581	203581	203581	206255	206255	203570

Note: This table decomposes the experimental estimate of the difference in tax demand for properties re-assessed by government tax assessors and private firms to determine whether the increase in demand was driven by increases in built-up area or property usage types. Property Usage Upgraded is an indicator recording whether the property usage was upgraded relative to baseline as a result of the re-assessment, including switches from residential to mixed-use and from mixed-use to non-residential. Property Usage Downgraded is an indicator recording whether the property usage was downgraded relative to baseline, including switches from non-residential to mixed-use and from mixed-use to residential. Irregular-shaped Property identifies multifloor properties where the floors have different areas. All specifications include ward fixed effects and controls for the respective outcome variable at baseline. Area and Property Usage data come from the ERP Master Table. Area data at baseline comes from the Floor Table. The sample is limited to the first list of properties the agents were instructed to re-assess and all control properties. Standard errors clustered by ward. We provide p-values for the test of equality between the government and private firms. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Differences in Skill and Effort between Government and Private Firms

	Skill			Effort	
	Total years experience surveying properties	Graduate degree or above	Master's degree or above	Average time spent on measuring one property	Number of revisits when no-one was present at the property
	(1)	(2)	(3)	(4)	(5)
Assigned to private	-9.60***	-0.28***	-0.29***	-8.01***	-0.42
	(1.09)	(0.07)	(0.06)	(1.97)	(0.33)
	[-11.82, -5.49]	[-0.43, -0.24]	[-0.40, -0.08]	[-13.23, -1.86]	[-1.04, 0.70]
Government = Private	< 0.001	< 0.001	< 0.001	< 0.001	0.206
Government Mean	11.03	0.85	0.35	30.72	1.96
Observations	147	147	147	139	147

Note: This table presents differences in employee skill and effort between government assessors and surveyors hired by private firms. We estimate OLS regressions of the outcome variable on an indicator for assigned to private firms (compared to government). Robust standard errors are reported in parentheses. Lee bounds are reported in square brackets. We provide p-values for the test of equality between government assessors and surveyors hired by private firms. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Instrumental Variables Estimates of Reassessment by Contract Type

	Reassessed (Apr 2024)	Tax Demand (Rs.)	Tax Collection (Rs.)	Tax Gap	Inaccuracy
	(1)	(2)	(3)	(4)	(5)
Reassessed by government	0.471***	6013.830***	3922.351***		
	(0.04)	(527.41)	(424.10)		
Reassessed by fixed-contract firms	0.369***	3456.806***	2029.115***	-0.098*	0.137***
	(0.02)	(457.97)	(289.41)	(0.06)	(0.04)
Reassessed by incentive-contract firms	0.392***	4541.495***	2237.686***	-0.114	0.140***
	(0.05)	(608.18)	(359.03)	(0.07)	(0.05)
Government = Fixed	0.035	<0.01	<0.01	0.098	0.001
Government = Incentive	0.211	0.067	0.002	0.109	0.008
Fixed = Incentive	0.667	0.154	0.651	0.845	0.955
Control Mean	0.00	7150.26	5735.05		
Government Mean				-0.09	0.28
Observations	206255	206255	206255	1499	1499

Notes: This table shows experimental effects of property reassessment by government assessors and private firms. Column 1 presents the first stage estimates. Columns 2 and 3 show treatment effects using 2SLS estimation with treatment assignment as instruments. Columns 4-5 show OLS estimates comparing private firms to government assessors (reference category). Tax Demand and Collection are for 2023-24. Tax Gap is the difference in the official Tax Demand for the 2023-24 tax year minus our estimated Tax Demand, divided by the sum of the two, and multiplied by 2. Estimated Tax Demand was calculated based on property characteristics recorded during the Property Survey and the government's property tax formula. Inaccuracy is the absolute value of Tax Gap. All specifications include ward fixed effects. Columns (1)-(3) include baseline controls. Standard errors clustered by ward. We provide p-values for the tests of equality between government and fixed-contract firms, government and incentive-contract firms, and fixed-contract and incentive-contract firms. Control Mean shows means for untreated properties (columns 1-3). Government Mean shows means for government-assessed properties (columns 4-5). *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Descriptive Statistics: Bidders vs Non-bidders

	Bidder	Non-bidder
	(1)	(2)
Panel A: Firm Characteristics		
Firm Age	22.75	22.33
	(4.92)	(26.82)
Permanent Staff Count	72.50	42.18
	(85.39)	(64.05)
Temporary Staff Count	55.00	24.30
	(63.51)	(44.51)
Turnover: INR 10,00,000 to INR 50,00,000	0.00	0.33
	(0.00)	(0.49)
Turnover: INR 50,00,000 to INR 1,00,00,000	0.00	0.33
	(0.00)	(0.49)
Turnover: INR 1,00,00,000 to INR 5,00,00,000	0.50	0.00
	(0.58)	(0.00)
Turnover: More than INR 5,00,00,000	0.25	0.25
	(0.50)	(0.45)
Firm Online Rating (0–5)	4.38	4.50
	(0.26)	(0.60)
Panel B: Past Experience with Government		
Past Tender Bids	1.00	0.64
	(0.00)	(0.50)
Past Tender Win	0.75	0.43
	(0.50)	(0.53)
Government Temporarily Halted Past Projects	0.25	0.00
	(0.50)	(0.00)
Government Delayed Payments in the Past	1.00	0.83
	(0.00)	(0.41)
Panel C: Contract Preferences		
Contract Preferred: Incentive	0.50	0.25
	(0.58)	(0.45)
Contract Preferred: Fixed	0.50	0.58
	(0.58)	(0.51)
Contract Preferred: Indifferent	0.00	0.17
	(0.00)	(0.39)
Observations	4	12

Note: This table reports descriptive statistics comparing bidder and non-bidder firms. The data come from a survey of firms that we conducted.

Table 7: Differences in Authority between Government and Private Firms

	Authority				Revisits/Action	
		Owner requested change to...				
	Share of refusals	Property area	Property usage	Number of floors	Number of revisits when owner denied permission for re-assessment	Action taken after refusal/ owner absence
	(1)	(2)	(3)	(4)	(5)	(6)
Assigned to private	0.04*	0.28***	0.12*	0.06	-1.42***	-0.39***
	(0.02)	(0.08)	(0.07)	(0.07)	(0.30)	(0.08)
	[-0.00, 0.12]	[0.17, 0.48]	[0.06, 0.32]	[0.01, 0.22]	[-2.05, -0.26]	[-0.62, -0.29]
Government = Private	0.069	< 0.001	0.100	0.340	< 0.001	< 0.001
Government Mean	0.15	0.35	0.20	0.16	2.01	0.70
Observations	137	147	147	147	147	144

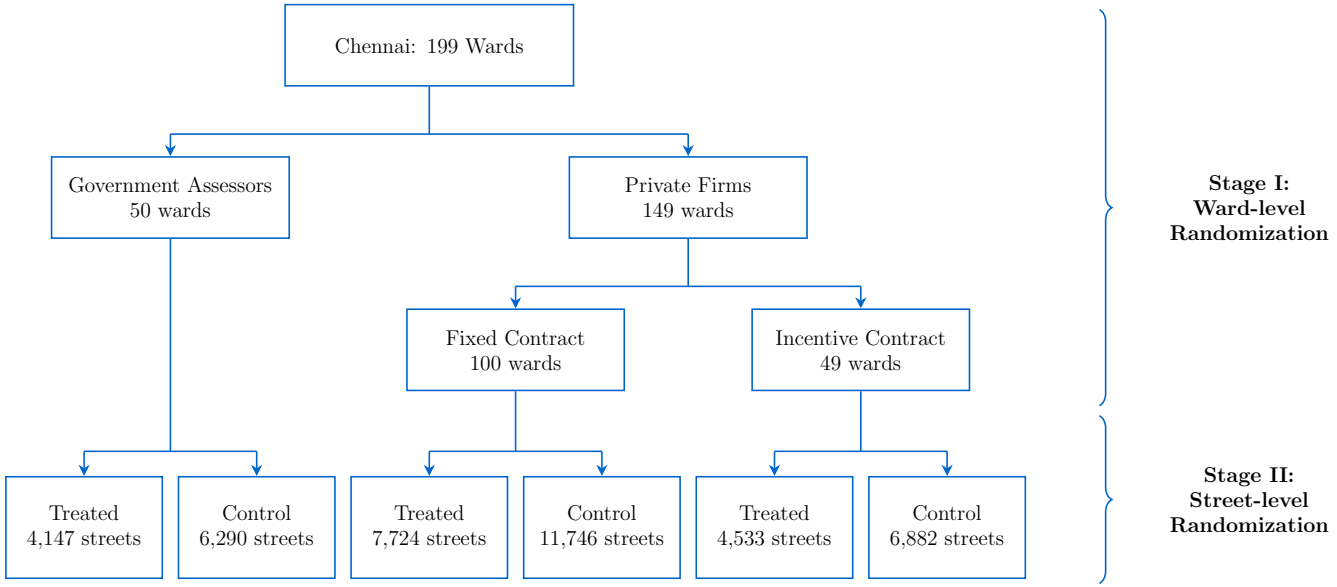
Note: This table presents differences in employee authority between government assessors and surveyors hired by private firms. We estimate OLS regressions of the outcome variable on an indicator for assigned to private firms (compared to government) with robust standard errors between parentheses. Lee bounds are reported in square brackets. We provide p-values for the test of equality between the government and private firms. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Cost-Benefit Analysis of Reassessment Assignment Rules (Demand-Based Estimates)

	Duration (in years)		
	1	5	10
Expected Benefit Per Property			
Government	6043.10	30215.49	60430.98
Private Firms	3978.65	19893.27	39786.55
Optimal Assignment Mix	5350.47	26752.34	53504.68
Random Assignment Mix	5010.88	25054.38	50108.76
Costs Per Property			
Government	285.02	285.02	285.02
Private Firms	230.09	230.09	230.09
Optimal Assignment Mix	257.56	257.56	257.56
Random Assignment Mix	257.56	257.56	257.56
Net Benefit Per Property			
Government	5758.08	29930.47	60145.96
Private Firms	3748.56	19663.18	39556.46
Optimal Assignment Mix	5092.91	26494.78	53247.12
Random Assignment Mix	4753.32	24796.83	49851.21

Note: Net benefit equals the average treatment effect on tax demand minus the average cost of reassessment. Government and private-firm rows report weighted average treatment effects obtained by averaging the estimated treatment-effect curves over the baseline-value distribution of properties. The random assignment mix reports the expected benefit and cost for a scenario in which, at each point in the baseline-value distribution, properties are assigned 50% to the government arm and 50% to the private-firm arm. The optimal assignment mix reports the expected benefit and cost for a scenario in which the largest 50% of properties, based on baseline value, are assigned to the government arm and the smallest 50% are assigned to the private-firm arm. The mixed cost equals the average of the government and private-firm per-property costs because both assignment scenarios allocate half of properties to each reassessment mode. The government per-property cost is estimated using employee survey data as total salary per employee, defined as monthly salary multiplied by project duration, divided by the number of properties handled per employee and adjusted by the fraction of total hours spent on reassessment. Private-firm costs are based on procurement records. We assume constant annual benefits over the specified duration periods.

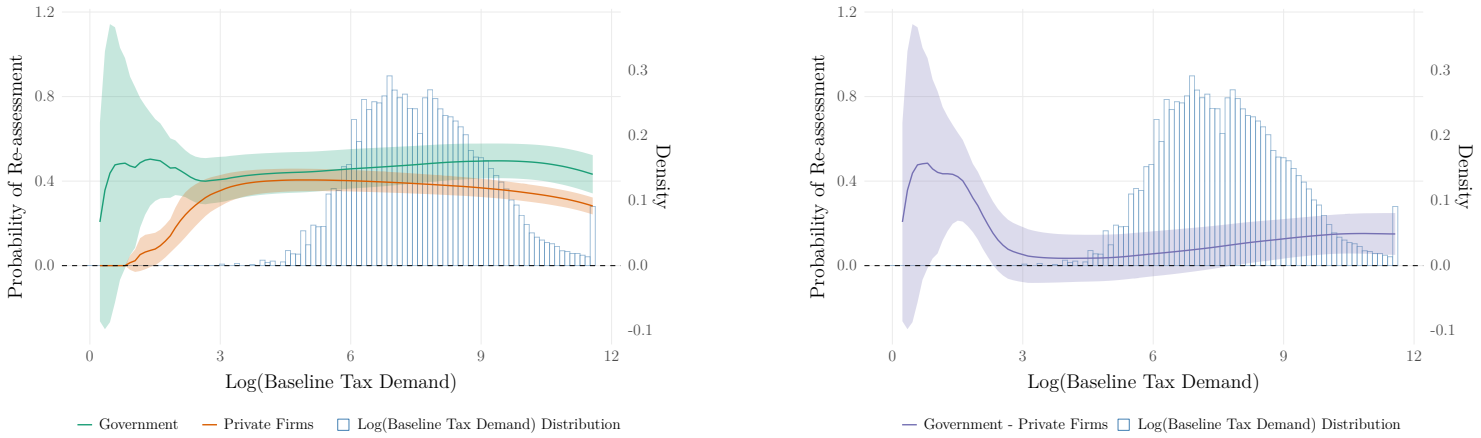
Figure 1: Experimental Design



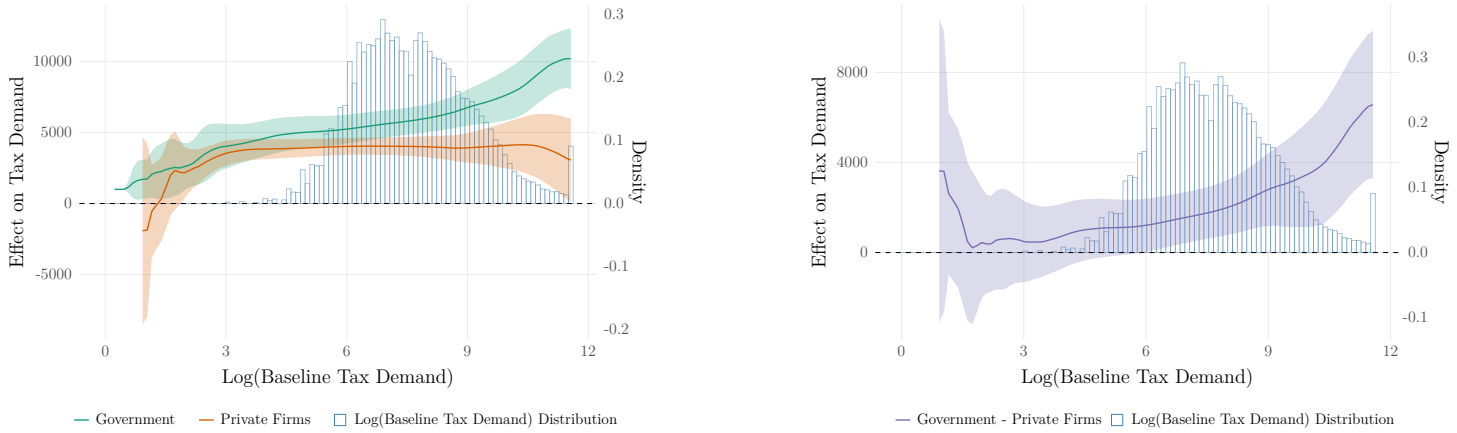
Note: The figure summarizes the two-stage experimental design. In the first stage, Chennai's 199 wards were randomly assigned to government assessors (50 wards) or private firms (149 wards). Wards assigned to private firms were further allocated between fixed-price contracts (100 wards) and incentive contracts (49 wards). In the second stage, streets within each assignment arm were randomly allocated to treatment or control. The numbers in the bottom row report the number of streets assigned to each group.

Figure 2: Heterogeneity in Reassessments and Demand, by Log(Baseline Tax Demand)

Panel A: First-stage Impacts on Probability of Reassessment



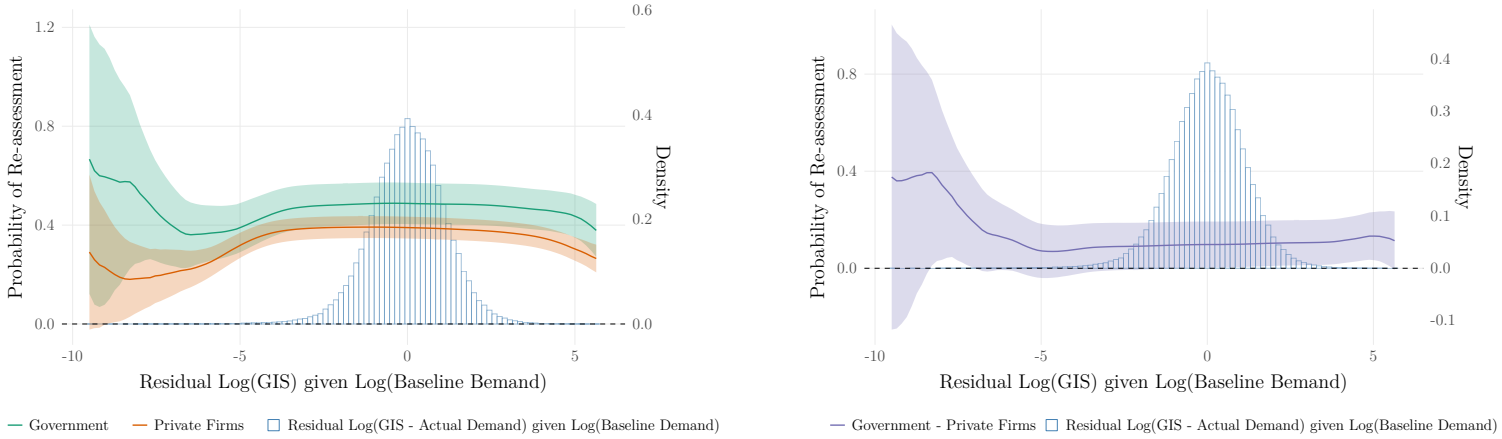
Panel B: IV Estimates of Heterogeneous Treatment Effects on Tax Demand



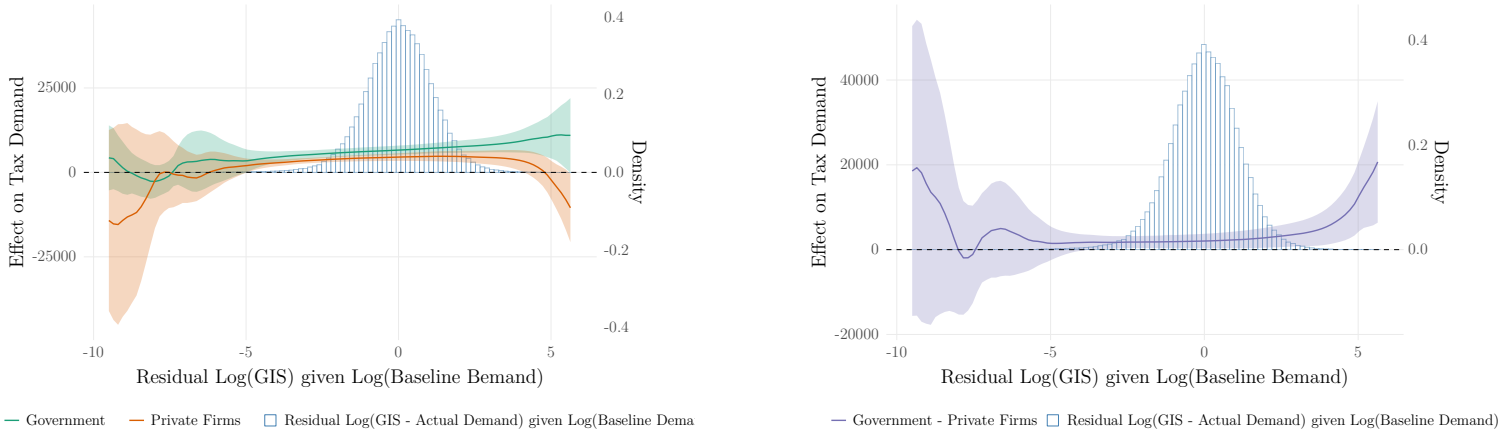
Note: This figure reports locally weighted estimates by log baseline tax demand, evaluated at 100 equally spaced points using a triangular kernel with a bandwidth equal to one-quarter of the horizontal-axis range. Panel A reports effects of assignment to government and private assessors on reassessment. Panel B reports the corresponding IV effects of reassessment on tax demand. Left figures show government and private estimates relative to the control group; right figures show government-minus-private differences. Histograms show the distribution of log baseline tax demand. All specifications include ward fixed effects, with standard errors clustered at the ward level.

Figure 3: Targeting Undervalued Properties

Panel A: First-stage Impacts on Probability of Reassessment

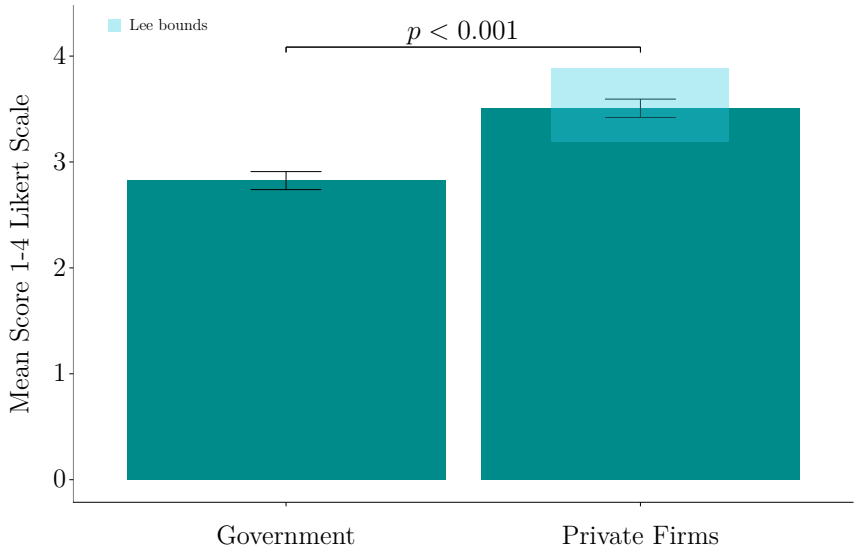


Panel B: IV Estimates of Heterogeneous Treatment Effects on Tax Demand



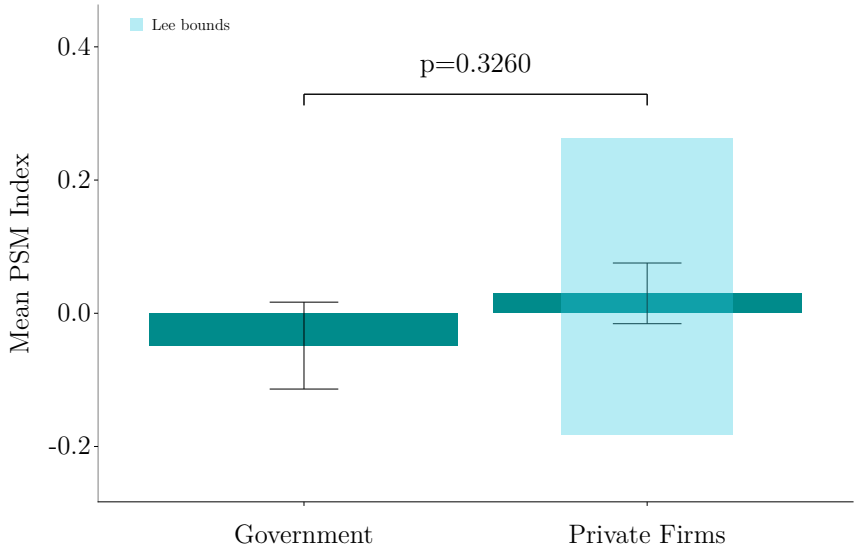
Note: This figure reports locally weighted estimates by residualized demand from the GIS drone survey, evaluated at 100 equally spaced points using a triangular kernel with a bandwidth equal to one-quarter of the horizontal-axis range. Residualized demand is obtained by regressing GIS-predicted demand on baseline tax demand and captures the component of GIS-predicted demand not explained by baseline tax demand. Panel A reports effects of assignment to government and private assessors on reassessment. Panel B reports the corresponding IV effects of reassessment on tax demand. Left figures show government and private estimates relative to the control group; right figures show government-minus-private differences. Histograms show the distribution of residualized GIS demand. All specifications include ward fixed effects, with standard errors clustered at the ward level.

Figure 4: Difficulty in Locating Properties



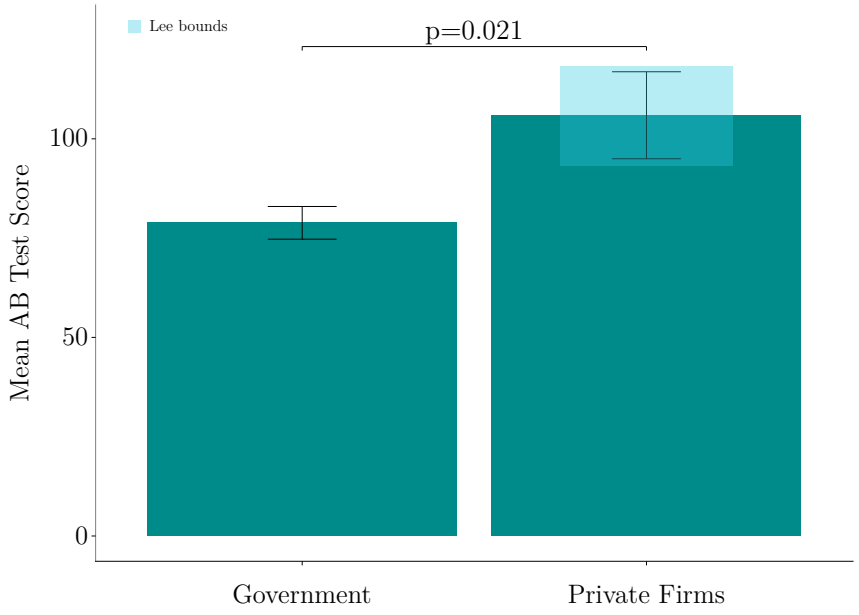
Note: The figure reports government and private assessors' mean self-reported difficulty in locating properties. Difficulty is measured on a four-point scale, where 1 denotes "Big Challenge," 2 denotes "Moderate Challenge," 3 denotes "Easy," and 4 denotes "Very Easy." Error bars represent two robust standard errors from a regression of the difficulty rating on indicators for government and private assessors, estimated without an intercept. The shaded band reports Lee bounds that adjust for differential attrition.

Figure 5: Public Service Motivation Index



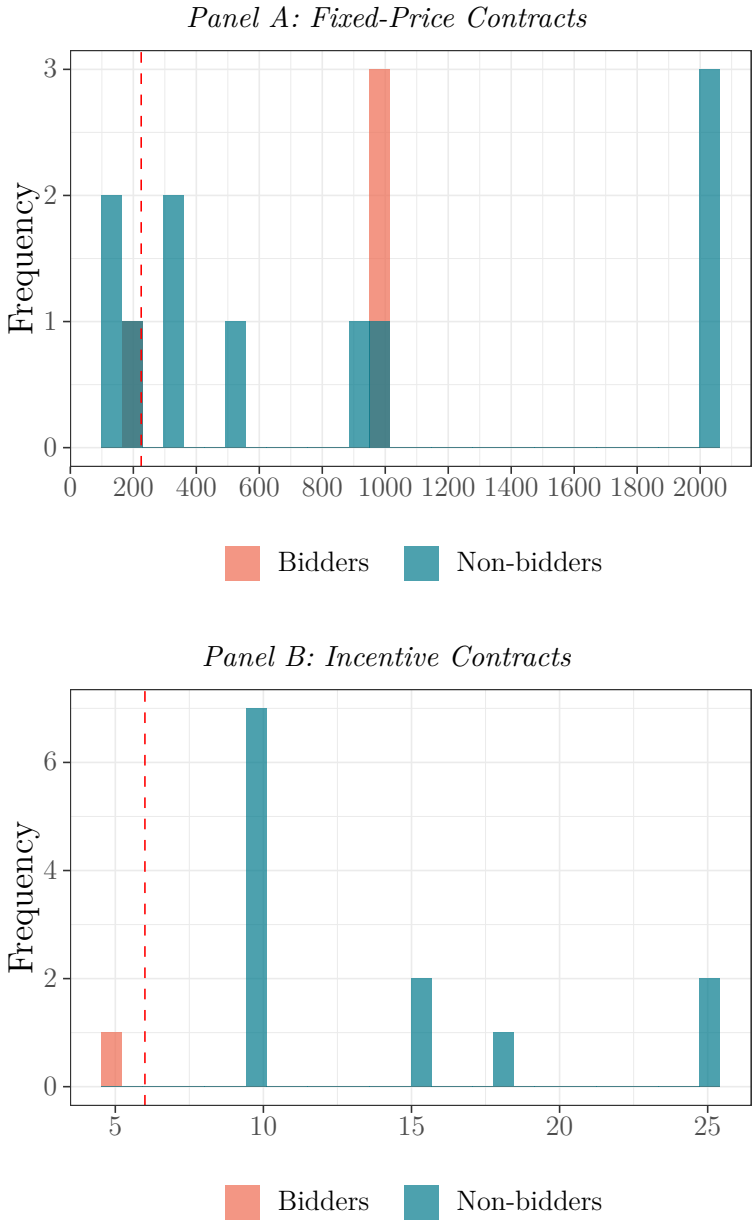
Note: The figure reports the mean public service motivation index for government assessors and surveyors hired by private firms. The index is constructed from 14 items adapted from Perry (1996). The items are standardized, and the index is calculated as the average of four dimensions: commitment to the public interest, social justice, compassion, and self-sacrifice. Error bars represent two robust standard errors. The shaded band reports Lee bounds that adjust for differential attrition.

Figure 6: A–B Test Scores of Government and Private Workers



Note: The figure reports the mean scores of government assessors and surveyors hired by private firms on the A–B effort test. Respondents were instructed to alternate between pressing the "a" and "b" keys as quickly as possible for one minute. Higher scores indicate greater effort. Error bars represent two robust standard errors from a regression of the test score on indicators for government and private workers, estimated without an intercept. The shaded band reports Lee bounds that adjust for differential attrition.

Figure 7: Proposed Bids for Reassessment Contracts



Note: The figure reports the distribution of bids submitted by bidders and the hypothetical bids reported by non-bidders for fixed-price contracts (Panel A) and incentive contracts (Panel B). In Panel A, a bid is the proposed price for reassessing one property. In Panel B, a bid is the proposed incentive payment expressed as a percentage of the increase in tax demand. The dashed line in Panel A indicates the final fixed-contract price of Rs. 224 per property. The dashed line in Panel B indicates the maximum incentive payment specified by the government, equal to 6 percent of the increase in tax demand.

Appendix Tables and Figures

Appendix Table 1: Technical Proposal Evaluation Criteria

Criteria	Marking System	Max Marks
A. Bidder's Prior Experience		
Experience conducting Property Enumeration / Assessment or similar tasks	• Relevant projects > 20,000 survey points: **30 marks** • Relevant projects > 10,000 assessment: **20 marks** • Relevant projects > 5,000 survey points: **10 marks**	30
B. Approach & Methodology		
Approach & Methodology	• Presentation on project understanding and value proposition • Adequacy of work plan responding to Terms of Reference • Project Team & resource-wise KPIs	40
C. Project Key Resource Profiles		
Project Manager	**Qualification:** Bachelors or equivalent in Urban Planning / Engineering / Geography / Management / Public admin / Commerce / Economics: **4 marks** **Experience as Project Manager:** • More than 7 years: **5 marks** • 5-7 years: **4 marks** • 3-5 years: **2 marks** **Experience as Project Manager in Municipal admin/Finance:** • More than 5 projects: **6 marks** • 3-5 projects: **4 marks** • 1-3 projects: **2 marks**	15
Team Lead	**Qualification:** Bachelors or equivalent in Urban Planning / Engineering / Geography / Management / Economics / Public admin / Commerce disciplines: **4 marks** **Experience as Team Lead:** • More than 7 years: **5 marks** • 5-7 years: **4 marks** • 3-5 years: **2 marks** **Experience as Team Lead in Municipal admin/Finance/Property tax:** • More than 5 projects: **6 marks** • 3-5 projects: **4 marks** • 1-3 projects: **2 marks**	15
	Total Maximum Marks	**100**

Appendix Table 2: Property Survey Response Rates, by Treatment Arm

	P(Surveyed)	P(Surveyed)
	(1)	(2)
Assigned to private firms	0.005	
	(0.021)	
Assigned to fixed-contract firms		0.034
		(0.025)
Assigned to incentive-contract firms		−0.019
		(0.024)
Government Mean	0.690	0.690
Observations	2,150	2,150

Note: This table presents results on response rates in the Property Survey, by treatment groups. The outcome is an indicator for whether the property was successfully surveyed. The sample includes treated properties that were randomly selected for the Property Survey. We estimate OLS regressions where the omitted category is properties assigned to the government. Standard errors are clustered at the ward level. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 3: Randomization Check (OLS estimates)

	Tax Demand (2022-23)	Tax Collection (2022-23)	Property Area (square feet)	Share Mixed-Use Properties (Government Records)	Share Non-Residential Properties (Government Records)	Share Mixed-Use Properties (GIS Survey)	Share Non-Residential Properties (GIS Survey)	Share Rental Properties	Share Owner-Occupied Properties	Share Mixed Tenancy Properties
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Assigned to government	276.868	182.303	17.222	0.004	0.000	0.002	0.000	-0.006*	-0.005	0.011**
	(187.99)	(175.80)	(20.43)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)
Assigned to private	229.182	230.007	7.035	-0.002	0.006*	0.001	0.006	0.000	0.002	-0.002
	(181.64)	(154.74)	(14.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Government = Private	0.855	0.839	0.681	0.316	0.261	0.885	0.335	0.138	0.373	0.029
Control Mean	6454.78	5639.72	1295.44	0.10	0.06	0.13	0.08	0.07	0.74	0.18
Observations	206255	206255	206255	206255	206255	206255	206255	203581	203581	203581

Note: This table provides a baseline balance check. Each column presents the coefficients from running an OLS regression of a baseline variable on the treatment arms variables. Regressions include ward fixed effects and standard errors are clustered at the ward level. Data on tax demand come from the DCB Table. Data on tax collection come from the Receipts Table. We provide p-values for the test of equality between the government and private firms. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 4: Difference between Properties Re-Assessed by Government Assessors and Firms on Tax Demand and Collection without baseline controls

	Reassessed (Aug. 2023)	Reassessed (Apr. 2024)	Tax Demand (Rs.)	Tax Collection (Rs.)
	(1)	(2)	(3)	(4)
Panel A: Reduced Form Estimates				
Assigned to government	0.473***	0.471***	3115.84***	1998.63***
	(0.043)	(0.042)	(448.74)	(305.43)
Assigned to private	0.338***	0.377***	1683.08***	981.42***
	(0.024)	(0.022)	(272.65)	(183.56)
Government = Private	0.006	0.050	0.01	0.00
Panel B: Instrumental Variables Estimates				
Reassessed by government			6610.07***	4239.99***
			(724.42)	(500.63)
Reassessed by private			4465.58***	2604.79***
			(632.12)	(465.79)
Government = Private			0.03	0.02
Control Mean	0.000	0.001	7120.43	5654.41
Observations	206255	206255	206255	206255

Note: This table replicates Table 1, but omits baseline control variables in columns (3) and (4). *** p<0.01, ** p<0.05, * p<0.10.

Appendix Table 5: Difference between Properties Re-Assessed by Government Assessors and Firms on Tax Demand and Collection with lists B and C as controls

	Reassessed (Aug. 2023)	Reassessed (Apr. 2024)	Tax Demand (Rs.)	Tax Collection (Rs.)
	(1)	(2)	(3)	(4)
Panel A: Reduced Form Estimates				
Assigned to government	0.474***	0.473***	2999.35***	1859.36***
	(0.043)	(0.043)	(371.47)	(272.13)
Assigned to private	0.338***	0.378***	1633.59***	778.09***
	(0.024)	(0.022)	(190.00)	(96.80)
Government = Private	0.006	0.050	0.00	0.00
Panel B: Instrumental Variables Estimates				
Reassessed by government			6344.06***	3933.07***
			(539.32)	(423.76)
Reassessed by private			4328.53***	2063.15***
			(366.01)	(218.21)
Government = Private			0.00	0.00
Control Mean	0.000	0.001	6951.26	5668.43
Observations	279055	279055	279055	279055

Note: This table replicates Table 1, but includes properties in lists B and C as control properties. *** p<0.01, ** p<0.05, * p<0.10.

Appendix Table 6: Difference between Properties Re-Assessed by Government Assessors and Firms on Actual and Imputed Tax Demand

		Re-calculated demand for Reassessed Properties	
	Actual Tax Demand	Tax Demand (No Rebate for Reassessed)	Tax Demand (Rebate for Reassessed)
	(1)	(2)	(3)
Reassessed by government	6013.83***	6056.23***	4168.99***
	(527.41)	(534.99)	(442.93)
Reassessed by private	3849.03***	4722.60***	2944.79***
	(372.92)	(343.77)	(304.80)
Government = Private	0.00	0.04	0.02
Control Mean	7120.43	7124.81	7122.41
Observations	206255	206246	206246

Note: This table estimates the difference in actual and imputed tax demand for properties re-assessed by government tax assessors and private firms. Actual Tax Demand is the property tax liability in the 2023-24 tax year. Tax Demand (No Rebate for Reassessed) represents the estimated tax demand for reassessed properties by applying the government's tax formula to the property's characteristics without any rebate. For properties that have not been reassessed, the actual tax demand is assigned. Tax Demand (Rebate for Reassessed) estimates the tax demand for reassessed properties by applying the government's tax formula with rebate. As with the previous measure, the actual tax demand is assigned to properties that have not been reassessed. All specifications include ward fixed effects and control for baseline annual demand. Standard errors clustered at the ward level. We provide p-values for the test of equality between the government and private firms. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 7: Accuracy of Reassessed Properties (OLS)

	Tax Gap	Inaccuracy
	(1)	(2)
Assigned to private	0.022	0.176***
	(0.052)	(0.037)
Government = Private	0.67	0.00
Government Mean	-0.05	0.25
Observations	1,489	1,489

Note: This table estimates the difference in non-revenue outcomes for properties re-assessed by government tax assessors versus private firms. Tax Gap is the difference in the official and estimated demand, divided by the estimated demand, while Inaccuracy is the absolute value of this measure. Official Tax Demand data come from the DCB dataset. We estimate OLS regressions where the omitted category is properties re-assessed by the government. Standard errors are clustered at the ward level. We provide p-values for the test of equality between the government and private firms. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table 8: Visit and Unofficial Payment Reports by Assessor Type

	Overall	Government	Private
Visited by assessor (%)	92.6	100	92
	(190)	(14)	(176)
Heard about officials demanding unofficial payments (%)	9.5	0	10.3
	(168)	(12)	(156)
Heard about people making unofficial payments (%)	8.9	0	9.6
	(169)	(12)	(157)

Notes: This table reports descriptive statistics from the property survey pilot data. Assessor type is defined by ward-level treatment assignment. Entries report the share of non-missing respondents answering "Yes" to each question, in percent. Number of observations are reported in parentheses. Visit outcomes refer to whether the respondent's building was recently visited. Unofficial payment outcomes refer to what respondents report hearing about during the most recent building measurement in their neighborhood.

Appendix Table 9: Reduced-Form Effects of Reassessment on Property Characteristics

	Area (sq. ft.)			Property Usage			
	Total	Residential	Non-Residential	Floor Count	Usage Upgraded	Usage Downgraded	Irregular-shaped Property
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Assigned to government	269.686***	241.525***	19.722***	0.273***	0.020***	0.004***	0.115***
	(30.97)	(27.05)	(4.75)	(0.03)	(0.00)	(0.00)	(0.01)
Assigned to private	179.604***	166.190***	4.154	0.187***	0.010***	0.006***	0.071***
	(17.69)	(15.65)	(2.82)	(0.02)	(0.00)	(0.00)	(0.01)
Government = Private	0.012	0.016	0.005	0.02	0.015	0.177	0.009
Control Mean	1336.860	1107.833	204.099	1.869	0.015	0.004	0.445
Observations	206255	203581	203581	203581	206255	206255	203570

Note: This table provides the reduced form estimates for Table 3. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 10: Difference between Properties Re-Assessed by Government Assessors and Firms on Property Characteristics with lists B and C as controls

		Area (sq. ft.)			Property Usage		
	Total	Residential	Non-Residential	Floor Count	Usage Upgraded	Usage Downgraded	Irregular-shaped Property
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reassessed by government	584.923***	513.553***	47.090***	0.586***	0.048***	0.010***	0.247***
	(44.20)	(39.59)	(8.85)	(0.04)	(0.01)	(0.00)	(0.02)
Reassessed by private	497.265***	450.201***	21.613***	0.516***	0.034***	0.019***	0.196***
	(30.03)	(26.90)	(6.71)	(0.03)	(0.01)	(0.00)	(0.01)
Government = Private	0.101	0.185	0.022	0.188	0.063	0.008	0.064
Control Mean	1330.254	1103.373	201.393	1.862	0.012	0.003	0.442
Observations	279055	275406	275406	275406	279055	279055	275395

Note: This table provides the reduced form estimates for Table 3, but includes properties in lists B and C as control properties. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 11: Impacts of Reassessment on Property Characteristics (Fixed vs Incentive Contracts)

		Area (sq. ft.)			Property Usage		
	Total	Residential	Non-Residential	Floor Count	Usage Upgraded	Usage Downgraded	Irregular-shaped Property
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Reassessed by government	572.085***	505.916***	41.310***	0.572***	0.042***	0.009***	0.241***
	(44.32)	(39.55)	(9.00)	(0.04)	(0.01)	(0.00)	(0.02)
Reassessed by fixed-contract firms	431.600***	399.656***	1.658	0.480***	0.026***	0.016***	0.175***
	(35.48)	(32.51)	(8.32)	(0.04)	(0.01)	(0.00)	(0.02)
Reassessed by incentive-contract firms	554.972***	506.025***	27.468**	0.519***	0.025***	0.017***	0.211***
	(50.59)	(42.91)	(12.47)	(0.05)	(0.01)	(0.00)	(0.02)
Government = Fixed	0.013	0.038	0.001	0.128	0.087	0.055	0.031
Government = Incentive	0.799	0.999	0.368	0.413	0.094	0.085	0.342
Fixed = Incentive	0.046	0.048	0.085	0.539	0.976	0.845	0.226
Control Mean	1336.860	1107.833	204.099	1.869	0.015	0.004	0.445
Observations	206255	203581	203581	203581	206255	206255	203570

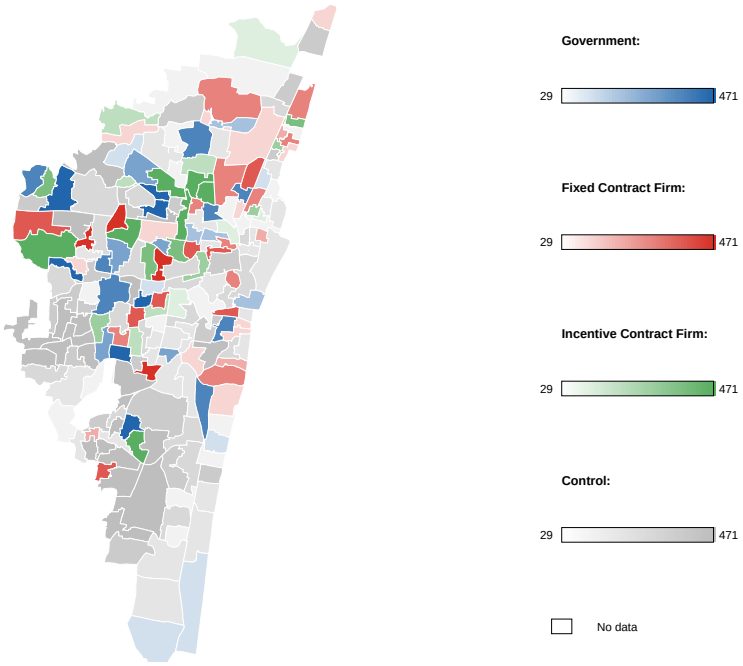
Note: This table replicates Table 3, decomposing the private category into properties reassessed by fixed-contract and incentive-contract firms. *** p<0.01, ** p<0.05, * p<0.1.

Appendix Table 12: Cost-Benefit Analysis of Reassessment Assignment Rules (Collection-Based Estimates)

	Duration (in years)		
	1	5	10
Expected Benefit Per Property			
Government	3938.25	19691.24	39382.49
Private Firms	2178.32	10891.61	21783.22
50/50 Baseline Split	3100.09	15500.45	31000.90
Random Assignment Mix	3058.29	15291.43	30582.85
Costs Per Property			
Government	285.02	285.02	285.02
Private Firms	230.09	230.09	230.09
50/50 Baseline Split	257.56	257.56	257.56
Random Assignment Mix	257.56	257.56	257.56
Net Benefit Per Property			
Government	3653.23	19406.22	39097.47
Private Firms	1948.23	10661.52	21553.13
50/50 Baseline Split	2842.54	15242.90	30743.35
Random Assignment Mix	2800.73	15033.87	30325.30

Note: Net benefit equals the average treatment effect on tax collection minus the average cost of reassessment. Government and private-firm rows report weighted average treatment effects obtained by averaging the estimated treatment-effect curves over the baseline tax collection distribution of properties. The random assignment mix reports the expected benefit and cost for a scenario in which, at each point in the baseline tax collection distribution, properties are assigned 50% to the government arm and 50% to the private-firm arm. The 50/50 baseline split reports the expected benefit and cost for a scenario in which the largest 50% of properties, based on baseline tax collection, are assigned to the government arm and the smallest 50% are assigned to the private-firm arm. The mixed cost equals the average of the government and private-firm per-property costs because both assignment scenarios allocate half of properties to each reassessment mode. The government per-property cost is estimated using employee survey data as total salary per employee, defined as monthly salary multiplied by project duration, divided by the number of properties handled per employee and adjusted by the fraction of total hours spent on reassessment. Private-firm costs are based on procurement records. We assume constant annual benefits over the specified duration periods.

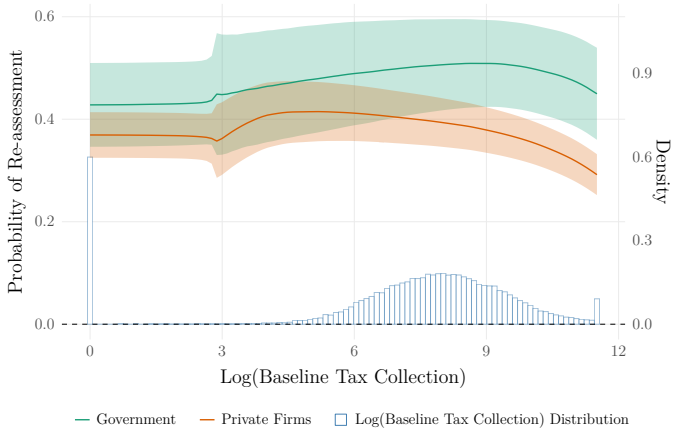
Appendix Figure 1: Properties Assigned for Re-Assessment by Ward and by Treatment

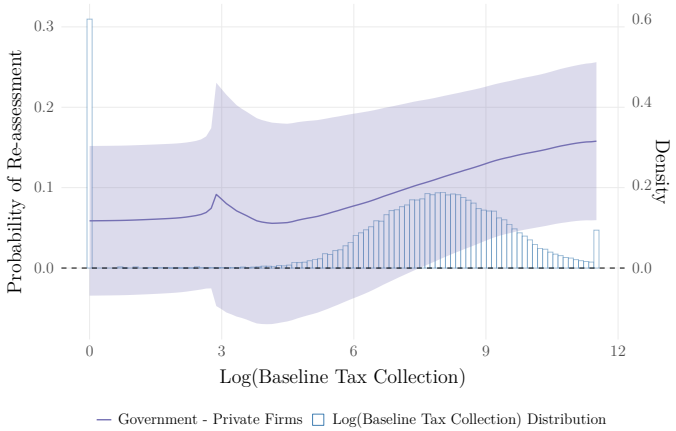


Note: This figure depicts the geographic distribution of properties selected for re-assessment. The sample is limited to the first list of properties the agents were instructed to re-assess. The color indicates the treatment arm and the color intensity the number of properties in each ward.

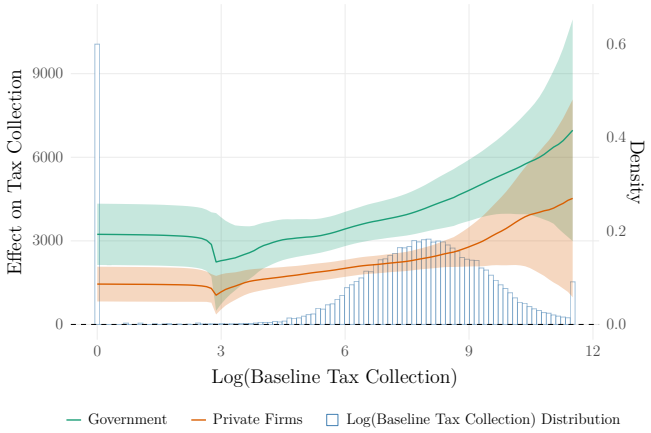
Appendix Figure 2: Heterogeneity in Reassessments and Tax Collection, by Log(Baseline Tax Collection)

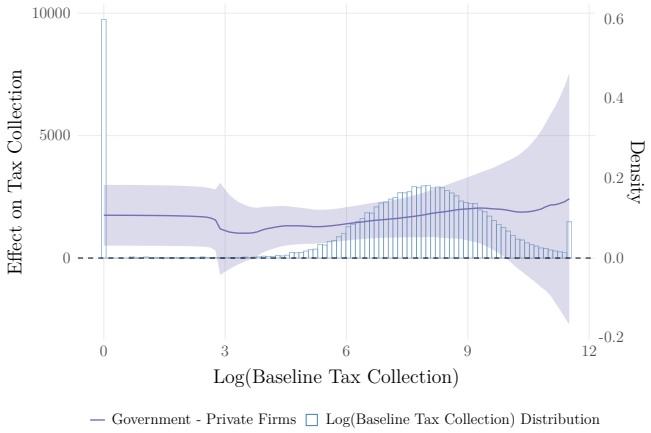
Panel A: First-Stage Effects on the Probability of Reassessment





Panel B: IV Estimates of Heterogeneous Treatment Effects on Tax Collection





Note: This figure reports locally weighted estimates by log baseline tax collection, evaluated at 100 equally spaced points using a triangular kernel with a bandwidth equal to one-quarter of the horizontal-axis range. Panel A reports effects of assignment to government and private assessors on reassessment. Panel B reports the corresponding IV effects of reassessment on tax collection. Left figures show government and private estimates relative to the control group; right figures show government-minus-private differences. Histograms show the distribution of log baseline tax collection. All specifications include ward fixed effects, with standard errors clustered at the ward level.