Online Appendix for "Tax Farming Redux"

Adnan Q. Khan, Asim I. Khwaja, and Benjamin A. Olken

Abstract

The Appendix contains additional information about the model, data, treatment implementation, and a series of robustness checks to our main tables. Tables are organized by category of robustness check. For additional results, we continue table numbering from the main text. For those tables that are variations of tables from the main text, we maintain the same table number, followed by a letter indicating the category of the robustness check.

Part I
Appendices

A Generalizations of Model

In the model presented in Section 3, we assumed for simplicity that the costs of reducing tax liabilities were linear in the amount reduced (i.e. the cost for taxpayers is $\alpha_i(\tau_i^* - \tau_i)$ and for inspectors is $\beta_i(\tau_i^* - \tau_i)$). Here we show two generalizations of this setup. In Section A.1 we generalize this cost function. In Section A.2 we derive a related model where the costs are in terms of bribes paid, rather than reductions in tax liability. The general conclusions are unchanged from the simple version in the text.

A.1 Generalizing the Cost Function

Instead of having the costs of evasion (α_i and β_i) be linear, which gives corner solutions of either full or no evasion, an alternative model would consider convex costs of evasion, yielding interior solutions for evasion. Specifically, suppose instead that the taxpayer's utility cost of accepting a reduced tax liability is $\alpha_i g(\tau_i^*, \tau_i)$, and the tax inspector's cost is $\beta_i h(\tau_i^*, \tau_i)$ where the marginal costs of collusion are weakly positively increasing in $\tau_i^* - \tau_i$. To guarantee an interior solution, as in Chetty (2009), we assume that $\frac{\partial g}{\partial \tau}|_{(\tau_i^*, \tau_i^*)} = 0$ and $\frac{\partial g}{\partial \tau}|_{(\tau_i^*, 0)} = \infty$, with analogous conditions for h. The taxpayer's utility function is $-\tau_i - \alpha_i g(\tau_i^*, \tau_i) - b_i$, where b_i is the bribe that the taxpayer needs to pay to reduce tax liability to τ_i from τ_i^*. The tax inspector's utility is $r\tau_i - \beta_i h(\tau_i^*, \tau_i) + b_i$.

As before, the tax inspector and taxpayer choose τ_i to maximize their net (of outside options) joint surplus, i.e. solve

$$\max_{\tau_i} -\tau_i - \alpha_i g(\tau_i^*, \tau_i) + r\tau_i - \beta_i h(\tau_i^*, \tau_i) + \tau_i^* - r\tau_i^* \tag{A.1}$$

The bribe paid is a fixed share $1 - \gamma$ of the joint surplus from evasion, plus the tax inspector's outside option if no evasion takes place, $r\tau_i^*$. Denote the solution to the maximization problem in (A.1) as $\hat{\tau}_i$.

$$b(\tau_i^*, \tau_i, r) = (1 - \gamma)[(1 - r)(\tau_i^* - \hat{\tau}_i) - \alpha_i g(\tau_i^*, \hat{\tau}_i) - \beta_i h(\tau_i^*, \hat{\tau}_i)] + r(\tau_i^* - \hat{\tau}_i) + \beta_i h(\tau_i^*, \hat{\tau}_i)$$

Note that since $\hat{\tau}_i$ was chosen to maximize (A.1), we can use the envelope theorem (the term in square brackets is the value function) to calculate

$$\frac{db}{dr} = -(1 - \gamma)(\tau_i^* - \hat{\tau}_i) + (\tau_i^* - \hat{\tau}_i) + \beta_i \frac{\partial h}{\partial \tau}(\tau_i^*, \hat{\tau}_i)\frac{d\tau}{dr} \tag{A.2}$$

If $\beta_i = 0$, then equation (A.2) shows that bribes increase in response to an increase in r. If $\beta_i > 0$, the change in bribes is theoretically ambiguous. If we assume that costs are convex (i.e. that $\frac{\partial^2 g}{\partial \tau^2}(\tau_i^*, \tau) > 0$ and $\frac{\partial^2 h}{\partial \tau^2}(\tau_i^*, \tau) > 0$), then taxes paid also increase with r under this model (i.e. $\frac{d\tau}{dr} > 0$).[33] The change in bribes b in response in this case depends on whether the increase in τ,

---

[33]To see this, note that that first-order condition from (A.1) is $(r - 1) = \alpha_i \frac{\partial g}{\partial \tau} + \beta_i \frac{\partial h}{\partial \tau}$. An increase in r increases the left hand side of this equation, so τ must increase.

and corresponding decline in h, offset the first-order increase in bribes needed to compensate for foregone incentive payments.

To calculate the change in social welfare from a change in r, we can rewrite the social welfare equation (equation 3.6) as

$$W = \int_i \left[\underbrace{-\hat{\tau}_i - \alpha_i g(\tau_i^*, \hat{\tau}_i) - \hat{b}_i}_{taxpayer} + \underbrace{r\hat{\tau}_i - \beta_i h(\tau_i^*, \hat{\tau}_i) + \hat{b}_i}_{taxinspector} + \underbrace{\hat{\tau}_i - r\hat{\tau}_i}_{government} \right] f(\alpha, \beta, \tau^*) \qquad (A.3)$$

For a given taxpayer-tax inspector combination i, the change in social welfare from a marginal change in r can be derived as follows. Note that the taxpayer and tax inspector have already chosen $\hat{\tau}_i$ to maximize their joint surplus. The envelope theorem says that the change in welfare from the first two terms together from a marginal change in r is just the mechanical effect, τ_i, since changes in the joint surplus from the induced change in τ_i are second-order. The derivative of the government's term, however, is $(1 - r)\frac{d\tau}{dr} - \tau_i$, since it also includes the effects from the change in τ_i. Adding up we have

$$\frac{dW}{dr} = \int_i \left[\underbrace{\hat{\tau}_i}_{taxpayer+taxinspector} + \underbrace{(1 - r)\frac{d\tau}{dr} - \hat{\tau}_i}_{government} \right] f(\alpha, \beta, \tau^*)$$

$$= \int_i \left[(1 - r)\frac{d\tau}{dr} \right] f(\alpha, \beta, \tau^*)$$

Again, the change in social utility is given by the change in tax revenue net of incentive payments.

While this model generates similar qualitative predictions to the model in the text (taxes increase with an increase in r, though the effect on bribes is ambiguous, and the change in social welfare can be calculated by examining the change in government budget (i.e. change in tax receipts net of change in incentive payments)), one important difference is that everyone continues to collude; that is, since we have set up the cost functions to guarantee an interior solution to evasion, there are no taxpayers induced to switch to the non-collusive equilibrium. However, this is simply due to the conditions imposed on the $g(.)$ and $h(.)$ functions that force an interior solution. Relaxing these assumptions would allow for situations where a change in r will have individuals that switch from positive evasion to no evasion and higher taxes.

A.2   Specifying Costs Based on Bribes Rather than Evasion

In the model in the text, the cost to taxpayers was in terms of tax evasion, i.e. in terms of $(\tau^* - \tau)$. An alternative specification of the model would be to specify the costs in terms of the bribes paid. i.e. in terms of b. This is similar to the model in the text, but slightly more cumbersome to work with, since now when solving the Nash bargaining one needs to take into account the fact that the transfers between the parties, b, are differentially costly for both sides. This means that the model can no longer be solved by computing the surplus and assigning a share γ of it to the tax inspector plus his outside option, but must be solved directly by maximizing the

product of the surpluses.

Specifically, in this model, we suppose that the utility of the taxpayer is $-\tau - b - \alpha b$ and the utility of the tax inspector is $r\tau + b - \beta b$, where $0 \leq \alpha, \beta < 1$. Note the difference between this setup and the utility specified in Section 3 – now there is no utility penalty from tax evasion per se, but instead in this model bribes are less valuable than money, reflecting perhaps the possibility of being caught or the need to launder the bribes to evade detection. To simplify the algebra, we assume equal bargaining weights (i.e. $\gamma = \frac{1}{2}$). Since this model is linear, as in the model in Section 3, conditional on collusion one will always set $\tau = 0$ and fully evade.

To solve the problem, the taxpayer and the tax inspector maximize the product of their excess utilities, i.e. solve

$$\max_b \left\{ r\tau + (1 - \beta) b - r\tau^* \right\} \left\{ -\tau - (1 + \alpha) b + \tau^* \right\} \tag{A.4}$$

subject to the constraint that $0 \leq b$ and $0 \leq \tau \leq \tau^*$. Conditional on collusion taking place, $\tau = 0$ and $b = \left(\frac{1}{2(1+\alpha)} + \frac{r}{2(1-\beta)} \right) \tau^*$. Collusion will only take place if the surplus is positive, which will be true if $r < \frac{1 - \frac{\alpha+\beta}{2(1+\alpha)}}{1 + \frac{\alpha+\beta}{2(1-\beta)}}$.

Note that the same qualitative results from the model in Section 3 apply here as well: the range of parameters at which collusion takes places is decreasing in r, α, and β, and conditional on collusion taking place, bribes are increasing in r.

This model easily accommodates an effort margin as well. Suppose that, prior to bargaining with the taxpayer, if the tax inspector exerts 0 effort, he learns that the property is valued at an amount τ_1. If he exerts effort e, with probability e he discovers that the true property valuation is $\tau_2 > \tau_1$. His cost of effort is $\frac{1}{2} c e^2$. Once the valuation τ_1 or τ_2 is revealed, this is common knowledge between taxpayer and tax inspector, and they jointly solve the bargaining game in (A.4). In this context, effort is also increasing in the incentive rate r, regardless of whether the equilibrium is collusion, full payment of taxes, or a combination.

B   Data

We use two main sources of data for analysis. Section B.1 describes the administrative data we use as our main measures of tax performance, and Section B.2 describes the data from surveys we conducted to obtain measures of tax assessment accuracy, customer satisfaction, and corruption.

B.1   Administrative Data

Our primary data are quarterly administrative data on tax collections. Each quarter, as part of their normal reporting requirements, circle inspectors report their revenue collected during the fiscal year cumulatively through the end of the quarter, which they compile from tax paid receipts retrieved from the national bank. In addition, they report their total assessed tax base before exemptions are granted (known as "gross demand") and after exemptions have been granted (known as "net demand"). These records are compiled separately for current year taxes and arrears. We digitized these quarterly reports for each of the approximately 500 tax circles in Punjab for a total of 6 years (4 years prior to the project beginning and the 2 years the project was in place).

Given the importance of this data in determining payments and measuring impact, it is important to validate its accuracy. Since reported tax receipts have to match actual money deposited in

the bank and ultimately received by the provincial Treasury department, the margin for discrepancies is low.[34] However, to ensure that the department's administrative data is correct circle-by-circle (since department's internal cross-checks are usually run at a higher level of aggregation), we instituted an additional re-verification program where we cross-checked the department's administrative records against the bank records. This entailed selecting a subset of circles (done identically in treatment and control areas), obtaining the individual records of payment received from the bank for each property, and manually tallying the sums from the thousands of properties in each circle to ensure that it matched the department total. Each circle had about an 18 percent chance of being verified by our cross-checks at some point during the two year experiment. We found virtually no systematic discrepancies between the administrative data we had received from the department and what we found in these independent verifications; the average difference between our independent verifications and what the circle had reported revealed under-reporting of -0.28%, or about zero.[35]

The administrative data also contain information on the identity of the inspector, which allows us to track if inspectors are relocated. We supplemented this by conducting a survey, each quarter, of the locations of all inspectors, constables, and clerks in the tax department.

B.2   Taxpayer/Property Survey Data

The second major data source is an independent property survey we conducted. This survey had three main purposes. First, it allowed us to obtain data on people's interactions with the tax department, both in terms of their overall perceptions of the quality of this interaction and on corruption.[36] Second, we obtained an independent assessment of the property's characteristics (e.g. land area, covered area, location), which we could use to construct an independent assessment of the property's valuation and compare to the department's official assessments.[37] Third, we obtained information about the owners and property characteristics that allows us to understand whether any observed impact of the schemes varied by the types of properties and owners.[38]

---

[34]When a taxpayer pays his tax due at the local bank, in addition to a receipt that he retains as proof of payment, two additional receipts are generated and collected by the bank. One of these is returned to the tax department and the other is given to the Treasury. The latter's totals (at the district level) are then matched to both the department's aggregates and also to the actual amount transferred by the bank to the Treasury.

[35]One complication is that circle boundaries are modified over time, as circles are merged and split to better reflect realities on the ground. In our data, out of the 482 circles present at the time of randomization, a total of 117 were affected by merges or splits throughout the 6 year period covered by our administrative data. To maintain consistency, we reconstruct the data at the level of the 482 circles present at the time of randomization. For those circles that merged prior to randomization, or split post randomization, one can simply add the two split circles together to obtain correct values for circles with randomization-era borders. For circles that split prior to randomization, or merged after randomization, we use the ratio of current year tax base net of exemptions (called "net demand" by the department) among the new and old circles reported in the quarters immediately before and after the split/merge to apportion the new circles to the randomization-era circles. The main results are qualitatively similar if we instead simply restrict analysis to the 365 circles that were unaffected (see Appendix Table 3-G2).

[36]The quality of interaction questions are solicited on a traditional Likert scale. For corruption, given that respondents are often not comfortable revealing their own bribe payments, we ask about the incidence of corruption and amount of bribe that would need to be paid for a property similar to theirs.

[37]Assessing areas is relatively straightforward since most properties in Punjab use standard lot sizes. To calibrate this, surveyors practiced assessing the size of various sizes of properties in training so that they could reliably estimate property sizes.

[38]In cases when the property was rented, we were not always able to obtain information about the owner. Although renters form 25% of our sample, Appendix Tables 4-M and 6-M show that the results are qualitatively similar if we

To do so, we surveyed approximately 16,000 properties. Properties were sampled in one of two ways. First, to obtain a general population sample of all properties (including those not necessarily on the tax rolls), we created GIS maps of the circle boundaries for all 482 circles, and used GIS software to randomly select five points within each circle. We then surveyed the property nearest that point, and selected seven more properties nearby (chosen by walking left from the point and choosing every other property) of which an additional four were surveyed based on a randomization table. Once this was completed, we matched these properties to the property-level administrative data to obtain the corresponding administrative records for these properties. On average, 85% of properties we randomly sampled in the field could be matched to corresponding administrative records, suggesting quite high coverage of properties throughout the province. These properties, which we refer to as our "general population sample," represent a random sample of 25 properties per circle, or 12,000 total.[39]

Second, since we were particularly interested in the properties whose tax valuation had changed, we also sampled properties directly off the separate tax lists that are maintained for newly assessed or reassessed properties to ensure we had sufficient representation of these properties in our sample. Specifically, we randomly selected 10 properties in each circle from those that had been reassessed during FY11-12 and FY12-13, and then located these properties in the field and surveyed them. We denote this sample of over 4,300 properties as the "reassessed" sample.

The survey was conducted at the end of the experiment. For logistical reasons it was split into two phases. The first phase was conducted during June and July 2013 (with a few properties finished in August and September), and covered approximately half the circles (randomly selected). The second phase was from October 2013 to January 2014, and covered the remaining half of the circles. For subjective measures (e.g. bribes, customer satisfaction), we focus on the results from the first wave of the survey, when respondents would surely be answering for the correct time period when the treatments were in effect. For objective measures (e.g. accuracy of assessment, property characteristics), we use both survey waves.

While the other measures are fairly self-explanatory, we should explain the assessment (in)accuracy measure. This is calculated by comparing the inspector's official assessment with the assessment we compute using our independent survey, normalized by the sum of the two measures, according to the following formula:

$$Inaccuracy = \frac{|GARV_{Inspector} - GARV_{Survey}|}{(GARV_{Inspector} + GARV_{Survey})} \tag{B.1}$$

This variable ranges from 0 to 1, with 1 indicating the greatest difference between the two metrics.[40] We also examine the tax gap, which is the same metric without the absolute value and which

---

drop rented properties.

[39]For budgetary reasons, one-fifth of the surveys in a circle were conducted using a shortened version of the questionnaire. The choice of which properties received short versus long survey was pre-determined as part of the sampling protocol and hence effectively randomized. When analyzing the survey data, we control for the format of the survey with a short survey dummy.

[40]Note that this measure is normalized by the sum of $(GARV_{Inspector} + GARV_{Survey})$. An alternative would just be to treat our measure, $GARV_{Survey}$, as the truth and normalize by that. However, if there is iid measurement error in each estimate, the average of the two will help smooth out the measurement error.

measures average amounts of over/under taxation:

$$TaxGap = \frac{GARV_{Inspector} - GARV_{Survey}}{(GARV_{Inspector} + GARV_{Survey})} \quad\quad\quad (B.2)$$

This measure ranges from 1 (complete over-taxation: inspector assesses the property positively whereas independent survey reveals 0 liability) to -1 (complete under-taxation: inspector assesses the property at 0 whereas independent survey reveals positive tax), with 0 indicating agreement.

C Knowledge and Credibility of Schemes

To ensure that collectors understood the specifics of the scheme they were in, we carried out detailed trainings for each scheme at the start of the year, post-training quizzes, and refresher trainings. By seven months after treatments started, quiz results revealed that virtually all inspectors were able to understand the scheme and accurately calculate the payments to which they would be entitled.

A total of five training sessions for circle staff were conducted over the treatment period: three in the first year (August, February, May 2011-12) and two in the second (August, October, 2012-13). Trainings were conducted after each randomization lottery, and retraining sessions were conducted after an interval of 2 - 4 months, covering all treatment staff. The training for each subtreatment was conducted separately to avoid any confusion between them. During the training session, treatment staff was given a detailed description of each subtreatment, the variables on which their performance would be judged, and the formula for calculating individual payments. Circle staff worked through examples calculating their earnings under different scenarios and the sessions concluded with a question and answer session.

To judge circle staff's understanding of the treatments, they were quizzed six times at regular intervals over two years. The quizzes tested staff on knowledge of the subtreatment they were selected in, the criteria they would be judged on, the formula for calculating their payments, and their ability to calculate their own payments under different scenarios. The initial understanding of the treatments was low, with around 37% inspectors failing the quiz conducted after the first training, but this improved substantially after multiple trainings, and within 7 months after randomization, failure rates dropped to under 1% and remained under 10% for the remainder of the treatment period. The quiz results show that shortly into the project, staff had a good understanding of their respective subtreatments, how their payments would be calculated, and what they needed to do to increase their payments.

Finally, to check that inspectors in fact knew what schemes they were in, towards the end of the second year (in July), all inspectors (including those not in treatments) were called in a brief phone survey. As part of this, they were asked whether they were participating in one of the circle-level incentives schemes, and if so, which one. Appendix Table 18 reports the result of a dummy for being in each of the 4 inspector-level schemes (Revenue, Revenue Plus, Flexible Bonus, or Information) on actual participation (instrumented with the results of the randomization). As is evident from the table, inspectors clearly knew both which scheme they were in and were able to accurately differentiate between the schemes, with the only mistake being that a small number of inspectors from the Revenue scheme mistakenly reported being in the Revenue Plus scheme.

Since performance-based pay had never been introduced in the tax (or, for that matter, any other) Punjab department before, another important aspect was ensuring credibility of the schemes

– i.e. that inspectors believed they would be paid as promised. This was partly helped by the successful completion of a pilot project with 11 tax circles selected for incentive schemes in the year preceding the project. We also ensured that at each stage of the process, formal approvals were received from all relevant government departments, including the Chief Minister (the equivalent of the provincial governor), the Tax department, the Finance department, and the Planning and Development department, and that these approvals were communicated to all parties concerned. Finally, the payment process was designed to further establish credibility. Although the final payments each year were determined only based on end-of-year totals, staff were paid each quarter based on their cumulative earnings under the scheme through that quarter, with corresponding quarterly benchmarks computed in the same way as the final annual benchmark.[41] Payments were carried out separately in each division, with checks handed out to every staff member along with a detailed calculation of the amount paid so that the staff could verify that the amount was indeed correct.

D   Decomposing the Revenue Impact

To better understand the source of the revenue changes observed, we can decompose them further using the administrative data. There are three margins that could be affected: the officially assessed tax-base (before exemptions are granted), the amount of exemptions granted before issuing tax bills, and the amount of tax revenue collected conditional on the (post-exemption) tax base. These are related as follows:

$$Revenue = TaxBase * NonExemptionRate * RecoveryRate \qquad (D.1)$$

where $NonExemptionRate = \frac{TaxBaseAfterExemptions}{TaxBaseBeforeExemptions}$ and $RecoveryRate = \frac{Revenue}{TaxBaseAfterExemptions}$. Taking logs of equation (D.1), we obtain an expression that additively decomposes the source of tax revenue.

Appendix Table 14 reports the results of this decomposition exercise for both treatment years. Columns 1 and 5 begin by reproducing the same regressions of $\ln Revenue$ from Columns 1 and 4 of Table 3, with the addition of baseline $\ln NonExemptionRate$ and $\ln TaxBase$ as controls.[42] The results show that virtually all the impact is coming from changes in the tax base, particularly for current year revenue. This implies that reassessments – which, recall, can be either finding new properties or changing the assessment on existing properties – are the main margin through which tax inspectors raise revenue.[43]

Appendix Table 14-H shows this decomposition separately for the three schemes. The only

---

[41]To do this, we computed benchmarks for each quarter in the same manner as we computed annual benchmarks, and made payments based on the cumulative amount of revenue collected through that quarter compared to the analogously computed benchmark. Only half of cumulative earnings were paid out to mitigate the possibility that staff were overpaid in the event that the pace of collections slowed over the year. This process was clearly explained in advance to inspectors and formed part of the training.

[42]The addition of these controls both reduces the sample size slightly (due to incomplete baseline values) and also slightly changes the point estimates. The coefficients in Columns 2 to 4 (6 to 8) should add up to the coefficient in Column 1 (5)

[43]It is also interesting to note that the tax base for arrears increases in year 1 but not in year 2. The arrears tax base can only increase if there are collections not made the year before, or if past uncollected amounts are now added in due to a valuation adjustment that is retroactively applied. Given the performance pay incentives in year 1, therefore, by year 2 there is likely not much room left to improve the arrears tax base.

notable difference among the sub-treatments is that the Revenue treatment also shows a statistically significant impact on recovery rate in years 1 and 2, driven by the recovery rate in arrears. Thus, it appears that in the Revenue treatment, inspectors worked both on tax base and recovery rate (particularly for arrears and in the second year), whereas in the other treatments that had multidimensional incentives, inspectors focused more on the tax base, at least as we can measure it here.

E   Mechanisms Beyond Price Effects

E.1   Information-only Scheme

To examine the role of information and perceived monitoring more directly, starting in Year 2 we introduced an "information-only" scheme. In this scheme inspectors received the same type of training and quarterly information sheets (with benchmarks) as inspectors in the Revenue scheme, but without any additional financial compensation. This scheme thus nets out the effects from the actual incentive payments from any other effect that may be in the other schemes.

While in our baseline specifications we included these circles as part of the control group, in Panel A of Appendix Table 20 we separate out the information scheme and compare it to the control group. The results in Appendix Table 20 show generally positive point estimates associated with the information scheme (particularly in the arrears treatment), but for total and current year revenue they are not distinguishable from zero. For current-year revenue, the point estimate is that the information scheme is associated with 7.1 log points higher revenue, compared with 16.8 log points for the Revenue scheme (a test for equality has a p-value of 0.093). More important, when one digs into the details, one does not see the same pattern of response to information as to any of other schemes. In particular, all three of our three treatment schemes seem to have obtained the bulk of their effect on current revenues primarily through a change in the tax base (see Appendix Table 14-H). The point estimates on the information scheme, however, show no change in either the number of properties added to the tax rolls (Panel A of Appendix Table 8-I) or in tax base for current revenue (See Appendix Table 14-I). This suggests that the (statistically insignificant) information results are unlikely to be a robust finding. Even if one assumes this informational/perceived monitoring effect is more systematic, the large (and likely more sustained) fraction of the effect of the Revenue scheme is attributable to the financial aspect of the performance incentives.

E.2   Testing for Income Effects

To test for the presence of income effects, we take advantage of the fact that benchmarks in the Revenue and Revenue Plus schemes were determined based on the 2nd, 3rd, and 4th lags of revenue, but not the 1st lag.[44] Since revenue can be closely approximated by an AR(1) process,

---

[44]The reason for this was both due to design and logistical considerations. In terms of design, not having the previous year's performance be part of the benchmark helped lessen ratchet effects in subsequent years (i.e. doing well in year 1 did not mean benchmarks for year 2 were higher). Logistically, benchmarks for 2012-2013 needed to be announced by the second week of July 2012, but the 2011-2012 revenue collection data would not be fully compiled, data-entered, and cleaned until August 2012. This meant that 2011-2012 revenue collection data could not be used in the computation of benchmarks for the 2012-2013 fiscal year.

this suggests a way of identifying income effects.[45] Specifically, we regress

$$LnRevenue_t = \gamma_1 LnRevenue_{t-1} + \gamma_2 LnBenchmark_{t-2} + \epsilon \tag{E.1}$$

to form the prediction $Ln\hat{Revenue}_t$, and then exponentiate to get $\hat{Revenue}_t$ and $\hat{Benchmark}_{t-2}$. This is the amount of revenue that would be collected under business-as-usual. An inspector in the Revenue or Revenue Plus group would therefore expect to earn

$$IncomeShock_t = \alpha \left(\hat{Revenue}_t - \hat{Benchmark}_{t-2} \right) \tag{E.2}$$

simply from business-as-usual. Since there is heterogeneity across inspectors in $IncomeShock_t$ (due to idiosyncratic variation in $Revenue_{t-1}$ conditional on $Benchmark_{t-2}$), this identifies the pure income effect that an inspector randomized into Revenue or Revenue Plus would receive compared to an inspector in the control group.[46] Since $IncomeShock_t$ is defined in both treatment and control areas, we can interact treatment status with the income shock that they would receive if they received a treatment to identify the treatment effect. Specifically, we estimate:

$$
\begin{aligned}
LnRevenue_{ct} = {} & \beta_1 Treatment_c + \beta_2 LnIncomeShock_{ct} \times Treatment_c + \beta_3 LnIncomeShock_{ct} \\
& + \beta_4 LnRevenue_{ct-1} + \beta_5 LnBenchmark_{ct-2} + \epsilon
\end{aligned} \tag{E.3}
$$

The key coefficient of interest is β_2.[47]

Panel B of Appendix Table 20 presents the results. There is no difference in performance based on the infra-marginal component of the revenue treatments, i.e. there is no evidence of any income effect. Again, this suggests that the key component is the price effect, not the income effect.[48]

E.3 Supervisory incentives

To the extent that circle staff were also aided by their supervisors, one could examine this effect by directly examining the impact of performance-pay for supervisory tiers. Starting in Year 2 of

---

[45]Specifically, if one regresses log revenue on its first 4 lags, the first lag has a coefficient close to 1 with an F-statistic of 314.4; the remaining 3 lags together have a joint F-statistic of only 5.2. We should note that this does not mean that benchmarks based on the 2nd through 4th lags are meaningless, just that the first lag is close to a sufficient statistic. For example, if one regresses current revenue on the 2nd, 3rd, and 4th lags only (omitting the first lag), one obtains an F-statistic of over 1,000; it is only once one also includes the 1st lag that the remaining lags have little explanatory power.

[46]In practice, this calculation is slightly more complicated, since in the first year of the schemes there were separate benchmarks for current-year revenue and arrears revenue (in year 2 they were combined). In year 1 we therefore estimate separate income shocks using equations (E.1) and (E.2) for current and arrears and then add to get the total income shock. Note also that this only works for the first year of the program (2012); in the second year, the first lag of revenue is the revenue realized in the first treatment year, which is endogenous.

[47]Note that $IncomeShock_{ct}$ is not quite a linear combination of $Revenue_{ct-1}$ and $Benchmark_{ct-2}$, given the exponentiation and subtraction, so we include the main effect of this as well.

[48]An alternate approach would be to directly examine price effects using the fact that different circles received different incentive rates α. In principle, since the reward rate changes discontinuously at the 50th and 75th percentile of baseline circle size, one can apply RD techniques to estimate the impact of a higher reward rate. The challenge is power, as the number of circles close to the discontinuity is very small. When we apply this approach, we find positive but noisy estimates of being in the 30% or 40% reward rate compared to the 20% reward rate in both years, though the results in 2013 are somewhat sensitive to the functional form used for the running variable (results available on request). However, the standard errors on these estimates are quite large (around 0.12 log-points for current and 0.35 log-points for arrears).

the experiment, such supervisory performance-pay was also introduced. These rewards were very similar to the Revenue incentive treatment, but they were paid based on the average performance above benchmarks for all of the taxable units (and hence staff) under their supervision.

There are two levels of supervisors – Excise and Taxation Officers (ETOs) and Assistant ETOs (AETOs). We randomized at the level of ETO and treated all AETOs who worked underneath them. Supervisors are not only responsible for monitoring the performance of the field staff and ensuring that collection targets are being met, but they can also directly aid in the collection process, especially in terms of supporting and imposing stronger sanctions on non-taxpayers and in handling appeals. All ETOs and AETOs had a mix of treatment and control circles working beneath them. Note that since we randomized at the level of 51 ETOs, we report randomized-inference based p-values, which are accurate in small samples and account for the clustering of the randomization at the ETO level.[49] With only 51 ETOs randomized in this treatment, compared to almost 500 circles in the main experiment, the level of statistical power is much lower here, but the Tax department wanted to include this scheme nonetheless, especially given the success of the circle staff schemes in the first year.

Panel C of Appendix Table 20 reports the results of the supervisory scheme. The unit of observation remains a circle. We find no effect – the point estimates are in fact negative for total, current, and arrears revenue, though they are never statistically distinguishable from zero. We have further investigated whether there are interactions between supervisors and the staff under them – i.e. if there is a particular synergy from having both supervisor and staff incentivized, or if the effects are orthogonal to one other. The results suggest that, if anything, paying supervisors only may in fact be detrimental to overall collections, though this is only marginally statistically significant and only for current-year revenue (See Appendix Table 21). We should caution that, given the imprecise estimates on the supervisory treatments due to lower sample size, it may be prudent not to make too much of this effect.

We also examine whether inspectors in treatment schemes believed they were being pressured more extensively by their supervisors to work harder. The results are shown in columns 1 and 2 of Appendix Table 22, and show that there is, on average, no difference in perceived pressure from supervisors between treatment and control areas. On net, the results presented here suggest that increased pressure from supervisors does not appear to be an important part of the channel driving the effects, both because rewarding supervisors directly has little effect and because we do not find any reported increase in supervisory pressure/support when tax circle staff are incentivized.

---

[49]We found through Monte-Carlo simulations that conventional cluster-robust standard errors appear too small in this context and over-reject the null.

Part II

Appendix Tables

F   Additional Results

This section presents additional results of interest that are referred to elsewhere.

Table 10: Spillovers

	(1) Total	(2) Current	(3) Arrears
Treatment	0.103***	0.103***	0.117
	(0.035)	(0.037)	(0.097)
Spillover Control	0.017	0.024	0.014
	(0.026)	(0.026)	(0.077)
N	480	480	476

Notes: This table examines possible geographic spillover effects. We present instrumental variable regressions, with (own-circle) randomization results instrumenting for (own-circle) implementation. The unit of observations is a circle, as defined in FY 2013 Q4. The dependent variable is log recovery; columns seperate margins of collection. A spillover control circle is a control circle for which more than half of circles within 1km are treatment circles. Treatment is own circle treatment status. Specifications control for baseline log recovery, and include stratum fixed effects. Information treatment is included in the controls. Standard errors are clustered by robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 11: Balance

	Main Treatment					Information Treatment		Supervisory Treatment	
	Control	Treatment	Revenue	Revenue Plus	Flexible Bonus	Control	Treatment	Control	Treatment
Log Revenue	15.47	-0.037	0.024	-0.053	-0.055	15.46	0.050	15.65	-0.167
		(0.042)	(0.059)	(0.057)	(0.058)		(0.065)		(0.089)
		[0.366]	[0.732]	[0.373]	[0.372]		[0.449]		[0.155]
Log Recovery Rate	-0.333	-0.015	0.007	-0.002	-0.039	-0.330	-0.026	-0.366	0.006
		(0.024)	(0.035)	(0.034)	(0.040)		(0.037)		(0.041)
		[0.575]	[0.846]	[0.966]	[0.297]		[0.530]		[0.909]
Log Non-exemption Rate	-0.251	-0.024	-0.001	0.009	-0.059**	-0.245	-0.035	-0.267	0.009
		(0.019)	(0.021)	(0.023)	(0.038)		(0.035)		(0.019)
		[0.210]	[0.967]	[0.742]	[0.0380]		[0.228]		[0.797]
Number of staff posted	2.564	0.055	0.038	0.056	0.088	2.576	-0.077	2.549	0.054
		(0.053)	(0.070)	(0.076)	(0.077)		(0.080)		(0.062)
		[0.326]	[0.638]	[0.500]	[0.274]		[0.359]		[0.625]
All positions filled	0.519	0.059	0.043	0.094	0.054	0.531	-0.072	0.538	-0.011
		(0.044)	(0.065)	(0.064)	(0.066)		(0.065)		(0.056)
		[0.215]	[0.488]	[0.157]	[0.436]		[0.286]		[0.908]
Log benchmark	15.44	-0.017	0.036	0.014	-0.073	15.44	-0.010	15.59	-0.114
		(0.044)	(0.062)	(0.059)	(0.064)		(0.074)		(0.095)
		[0.690]	[0.621]	[0.823]	[0.258]		[0.905]		[0.369]
FY 10-11 log growth rate	0.0280	0.003	0.017	-0.005	-0.005	0.0233	0.026	0.0259	-0.008
		(0.013)	(0.014)	(0.016)	(0.023)		(0.022)		(0.019)
		[0.806]	[0.399]	[0.780]	[0.811]		[0.206]		[0.777]
P-val, joint sig.		0.412	0.793	0.006	0.455		0.261		0.003
P-val, from RI		0.409	0.820	0.011	0.592		0.395		0.150

Notes: The table presents balance tests for the randomization into the different treatments. Columns labelled Control reflect control group means. Values in the treatment columns are the coefficients of a regression of the baseline value of the variable indicated in the row on a treatment dummy (or the set of subtreatments dummies), controlling for the relevant randomization strata. In the Main Treatment tests, the Information treatment is included in the controls. In the Information Treatment tests, the Information treatment group is compared against pure controls. The Supervisory Treatment test compares against its own control means (which are different from column 1 means). Robust standard errors in parentheses. Randomization inference (RI) based p-values in brackets. RI statistics are based on 1000 re-randomization iterations. * p<0.10, ** p<0.05, *** p<0.01. Stars reflect randomization inference based p-values.

Table 12: Correlation of Satisfaction and Corruption Variables

	(1) Satisfaction	(2) Quality	(3) Bribes (normalized)	(4) Perception of Corruption
Panel A: Correlations with own response				
Satisfaction		.677*** (.0196)	-.0015 (.1124)	-.0463 (.0305)
Quality	.7459*** (.02)		-.2584** (.1119)	-.0046 (.0283)
Bribes (normalized)	-4.1e-05 (.003)	-.0063*** (.0018)		-.007 (.0052)
Perception of Corruption	-.0215 (.0144)	-.0019 (.0119)	-.1211 (.0773)	
N	5541	5541	5541	5541
Panel B: Correlations with others in circle				
Satisfaction	.8953*** (.0152)	.1053*** (.0152)	-.2431 (.224)	.0014 (.0045)
Quality	.0844*** (.0157)	.8589*** (.0172)	.1351 (.1514)	-.002 (.0049)
Bribes (normalized)	-.0023 (.0019)	5.8e-04 (.0012)	.731*** (.2428)	-.0011 (.0011)
Perception of Corruption	5.8e-04 (.002)	-8.8e-05 (.0024)	-.0424 (.0523)	.972*** (.0016)
N	6350	6350	5541	6350
Mean of dependent variable	0.568	0.548	0.305	0.645

Notes: This table presents OLS regressions where we examine the relationship of our subjective non-monetary outcomes with each other (Panel A) and other respondents (Panel B). The unit of observation is a property. Panel A reports regressions of dependent variable on own response of independent variables. Panel B reports regressions of dependent variable on circle-level estimators of independent variables (where the own measure is excluded). Bribes are normalized by gross annual rental value (GARV) to be comparable within circle. Satisfaction, Quality, and Perception of Corruption are measured on 5-point Likert scaled, normalized to a [0,1] interval. Robust standard errors in parentheses. Standard errors are clustered by robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 13: Impacts on Satisfaction with the Government

	(1) Quality of Electricity Dept.	(2) Quality of Water Dept.	(3) Satisfaction with Electricity Dept.	(4) Satisfaction with Water Dept.	(5) Likelihood of Picking up Note	(6) Indicated Preference for Incumbent Party
Panel A: Main Treatment						
Treatment	-0.006	-0.011	-0.003	0.003	-0.020	0.004
	(0.028)	(0.022)	(0.030)	(0.022)	(0.029)	(0.037)
Panel B: Subtreatments						
Revenue	-0.023	-0.031	-0.015	-0.039	-0.011	0.048
	(0.042)	(0.030)	(0.047)	(0.031)	(0.044)	(0.049)
Revenue Plus	0.057	0.039	0.050	0.074**	-0.005	-0.026
	(0.043)	(0.031)	(0.044)	(0.030)	(0.041)	(0.050)
Flexible Bonus	-0.051	-0.041	-0.042	-0.028	-0.041	-0.003
	(0.037)	(0.029)	(0.040)	(0.032)	(0.043)	(0.056)
N	4840	4840	4840	4840	4840	4840
Sample	Phase 1	Phase 1	Phase 1	Phase 1	Phase 1	Phase 1
Mean of control group	0.416	0.507	0.431	0.520	0.351	0.605
Rev. vs. Multitasking p.	0.575	0.345	0.708	0.071	0.800	0.227
Objective vs. Subjective p.	0.102	0.156	0.182	0.190	0.479	0.794
Equality of Schemes	0.091	0.069	0.206	0.004	0.770	0.434
Joint significance	0.175	0.123	0.361	0.011	0.821	0.630

Notes: This table examines the impact of the overall treatments (Panel A) and subtreatments (Panel B) on a wider range of non-monetary outcomes compared to those in Table 4. We present estimates from instrumental variables regressions, where treatment status is instrumented with randomization results. The unit of observation is a property. Quality and Satisfaction were measured on a 5 point Likert scale and re-scaled to a [0,1] interval. Likelihood of Picking up Note is the respondent's assessment (on a 5 point scale, also re-scaled to a [0,1] interval) of how likely a stranger would return them a Rs. 1000 they had accidentally dropped. Indicated Preference for Incumbent Party is a binary variable = 1 if the respondent indicates that any member of the household voted for the incumbent party at either the provincial or national level in the most recent elections. Sample restricted to circles surveyed in the first phase of the survey (see text for details). Specification includes strata fixed effects and controls for whether the property was surveyed using the short version of the survey. Information treatment included in the control group. Robust standard errors in parentheses. Standard errors are clustered by robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 14: Impacts on Tax Base and Recovery Rates, All Treatments

	Year 1				Year 2			
	(1)	(2)	(3) Non-Exemption Rate	(4) Recovery Rate	(5)	(6)	(7) Non-Exemption Rate	(8) Recovery Rate
	Revenue	Tax Base			Revenue	Tax Base		
Total								
Any Treatment	0.076***	0.090***	-0.025	0.011	0.090***	0.054	0.006	0.029
	(0.027)	(0.029)	(0.018)	(0.023)	(0.030)	(0.033)	(0.020)	(0.019)
Curent								
Any Treatment	0.073***	0.084***	0.000	-0.012	0.096***	0.067**	0.018	0.011
	(0.028)	(0.028)	(0.014)	(0.022)	(0.032)	(0.030)	(0.016)	(0.016)
Arrears								
Any Treatment	0.111*	0.133*	-0.053	0.032	0.075	-0.006	0.053	0.028
	(0.065)	(0.069)	(0.036)	(0.036)	(0.081)	(0.090)	(0.046)	(0.035)
N (Total)	473	470	470	473	474	474	474	474
Mean of control group (Total)	15.680	16.114	-0.201	-0.225	15.756	16.149	-0.229	-0.165

Notes: This table decomposes the treatment effect on revenue collection (Columns 1, 5) into the effect on three components: changes in the Tax Base (Columns 2, 6); changes in the Non-Exemption Rate (Columns 3, 7); and changes in the Recovery Rate (Columns 4, 8). We use instrumental variables regressions, where treatment status is instrumented with randomization results. The unit of observation is a circle, as defined at the time of randomization. Rows indicate the relevant margins of collection (total revenue, current year revenue, and collections against past arrears). Control variables include baseline tax base, non-exemption rate, and recovery rate. Outcome variables and controls in logs. Specification includes stratum fixed effects. The Information treatment is included in the control group. Number of observations and means of control group are reported for total collections (current and arrears are similar). Robust standard errors in parentheses. Standard errors are clustered by robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 15: Change in Bribe Payments on the Intensive and Extensive Margin

	(1) Bribe Payment	(2) Bribe Payment	(3) Bribe Payment = 0
Treatment	594.1*	744.5*	-.035
	(341.7)	(450)	(.0436)
N	5993	4207	5993
Mean of control group	1874.542	2734.269	0.314
Restricted to positive bribe payments	No	Yes	N/A

Notes: This table separates the impact of performance pay schemes on bribe payments into 1) the amount of bribe paid and 2) the likelihood of paying any bribe. Column 1 replicates Column 2 of Table 6 of the main text. Column 2 repeats the same specification, restricting to taxpayers that report positive bribe payments. Column 3 uses as a dependent variable a dummy equal to 1 if the taxpayer reports 0 bribe payments. The Information treatment is included in the controls. Robust standard errors in parentheses. Standard errors are clustered by robust partition, the partition of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 16: Impacts on Inspector Effort

	(1) Team Effort	(2) Hours/Day spent in field	(3) Hours/Day spent in office	(4) Total hours worked in typical day
Panel A: Main Treatment				
Any treatment	-.5	-.286**	.156	-.13
	(2.01)	(.141)	(.147)	(.0792)
Panel B: Subtreatments				
Revenue	-2.5	-.0782	.0171	-.0611
	(3.77)	(.213)	(.216)	(.128)
Revenue Plus	2.45	-.289	.103	-.185**
	(2.69)	(.182)	(.191)	(.0765)
Flexible Bonus	-1.68	-.445**	.313	-.132
	(2.83)	(.215)	(.204)	(.1)
N	353	353	353	353
Mean of control group	90.6	5.46	2.7	8.16

Notes: This table examines the impact of performance pay on self-reported inspector effort. We use instrumental variables regressions, where treatment status is instrumented with randomization results. The unit of observation is a circle, as defined at the end of Year 2. Team Effort is assessed on a 100 point scale. The Information treatment is included in the controls. Robust standard errors in parentheses. Standard errors are clustered by a robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 17: Treatment Effect on Probability of Transfer by Position

	(1) Inspector	(2) Constable	(3) Clerk	(4) All
Panel A: Main Treatment				
Any treatment	-0.042	-0.045	-0.091	-0.051
	(0.052)	(0.046)	(0.064)	(0.033)
Panel B: Subtreatments				
Revenue	-0.052	-0.015	-0.073	-0.043
	(0.080)	(0.073)	(0.096)	(0.051)
Revenue Plus	-0.049	-0.071	-0.035	-0.053
	(0.080)	(0.060)	(0.096)	(0.053)
Flexible Bonus	-0.025	-0.050	-0.161	-0.057
	(0.077)	(0.071)	(0.102)	(0.048)
N	426	428	240	1094
Mean of control group	0.502	0.285	0.430	0.401

Notes: OLS regressions of transfer probability on randomization status. Columns separate effects by staff type. Unit of observation is a staff members. Sample restricted to circles that remained in treatment or control for two years. Specification includes strata fixed effects. Column 4 also includes staff type fixed effects. The Information treatment is included in the controls. Robust standard errors in parentheses. Standard errors are clustered by a robust partition of circles, the group of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 18: Inspectors' Knowledge of Treatments

Perceived Treatment	Revenue	Revenue Plus	Flexible Bonus	Information	Control	N
Actual Treatment						
Revenue	0.82	0.16	0.02	0.00	0.00	51
Revenue Plus	0.07	0.93	0.00	0.00	0.00	54
Flexible Bonus	0.04	0.05	0.86	0.00	0.05	56
Information	0.00	0.00	0.00	0.80	0.20	51
Control	0.03	0.02	0.00	0.01	0.94	159

Notes: This table provides a tabulation of inspector's understanding of their own treatment status. Rows list inspector's actual treatment status. Columns list inspector's perceived treatment status. Cells list the fraction of inspectors in the treatment given by the row who believe themselves to be in the treatment specified in the column (i.e. columns 1 through 5 should total to 1.00).

Table 19: Inspector Beliefs

	(1) Chance of being rewarded after 2012 ballot	(2) Chance of being selected in 2013 ballot	(3) Chance of being selected in 2013 ballot relative to incentive circles	(4) Chance of being selected in 2013 ballot relative to control circles
Information	-.00205 (.185)	.0091 (.183)	.283 (.211)	.422 (.272)
Revenue		.163 (.183)		
Revenue Plus		.172 (.176)		.112 (.28)
Flexible Bonus		.266 (.195)		.222 (.253)
N	178	340	180	207
Omitted group	Control	Control	Control	Revenue
Mean of omitted group	4.00	3.98	2.12	2.00

Notes: This table examines how treatment assignment may have affected inspectors' subjective assessments of their chances of being selected for future schemes. We report ordered probit regressions, where treatment is measured with randomization results. The unit of observation is a circle, as defined at the time of randomization. Responses in columns 1 and 2 were coded on a 5 point probablity scale, with 1 indicating 'Not likely at all' and 5 indicating 'Definitely.' Responses in columns 3 and 4 were coded on a 3 point probability scale, with 1 indicating less chance, 2 indicating same chance, and 3 indicating higher chance. Robust standard errors in parentheses. Standard errors are clustered by a robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 20: Mechanisms Beyond Price Effects

Panel A

	(1) Total	(2) Current	(3) Arrears
Information	0.064 (0.051)	0.071 (0.050)	0.226* (0.136)
N	482	482	479
Mean of control group	15.709	15.486	13.864

Panel B

	(1) Total	(2) Current	(3) Arrears
Revenue * Income Shock	0.0229 (0.104)	0.0499 (0.111)	-0.0481 (0.218)
Revenue Plus * Income Shock	0.0459 (0.167)	0.0317 (0.143)	-0.250 (0.362)
Income Shock	-0.0797 (0.0678)	0.00846 (0.0461)	0.119 (0.0809)
N	478	478	478

Panel C

	(1) Total	(2) Current	(3) Arrears
Supervisory treatment	-0.065 [0.319]	-0.106 [0.220]	-0.035 [0.852]
N	482	482	479
Mean of control group	15.924	15.679	14.176

Notes: This table presents results on mechanisms other than price effects that may be contributing to the observed impact of the performance pay schemes. In all panels, we use instrumental variables regressions, where treatment status is instrumented with randomization results. The unit of observations is a circle, as defined at the time of randomization. The outcome variable is log recovery as of the end of the second year of the study (FY 2012-2013). Columns separate recovery by total recovery (Column 1), current year recovery (Column 2), and collections against past arrears (Column 3). Panel A: This table re-estimates main revenue outcomes by subtreatment, with the Information treatment separated from the control group. Coefficients for the Revenue, Revenue Plus, and Flexible Bonus treatments not shown. Panel B: This table examines whether income effects contribute to the observed outcomes. The income Shock is calculated as the amount circle staff team would have earned under the scheme due to business-as-usual, plus their combined base salary, and is measured in logs (see Section for details). Specification follows Equation 6.4 of the main text. The coefficient for Flexible Bonus is not shown. Panel C: This table examines the impact of the supervisory treatments. Standard errors: For Panel A, we present robust standard errors in parentheses. Standard errors are clustered by robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. For Panel B, standard errors, in parentheses, are bootstrapped with 1000 iterations over the two-step estimation procedure (i.e. first estimating the income shock, and then estimating the model). The bootstrap sampling procedure is clustered by robust partition. For Panel C, randomization-inference p-values are shown in brackets, calculated over 1000 iterations. * p<0.10, ** p<0.05, *** p<0.01

Table 21: Impact of Interactions between Supervisory and Inspector Treatments on Revenue Outcomes

	Total	Current	Arrears
Supervisory treatment	-0.089	-0.131*	-0.111
	[0.127]	[0.058]	[0.543]
Inspector treatment	0.063	0.068	0.034
	[0.159]	[0.158]	[0.778]
Inspector Treatment * Supervisory Treatment	0.103	0.076	0.271
	[0.299]	[0.457]	[0.328]
N	482	482	479
Mean of control group	15.901	15.662	14.157

Notes: This table checks for potential interaction effects between performance pay incentives for inspectors and supervisors. We use instrumental variables regressions, where both supervisory and circle treatments are instrumented with randomization results. The unit of observation is a circle, as defined at the time of randomization. Columns separate recovery by total recovery (Column 1), current year recovery (Column 2), and collections against past arrears (Column 3). Specification include division fixed effects and baseline log revenue collection. Mean of control group reflects mean for pure controls, i.e. circles that did not fall under either the inspector or supervisory treatments. The Information treatment is included in the controls. Randomization inference based p-values in brackets. * p<0.10, ** p<0.05, *** p<0.01

Table 22: Impact of Treatment on Inspector Monitoring

	(1) Pressure from supervisors	(2) Level of monitoring
Panel A: Main Treatment		
Any treatment	.0186	-.0251
	(.0302)	(.0438)
Panel B: Subtreatments		
Revenue	.0169	-.111
	(.0479)	(.0706)
Revenue Plus	-.0561*	.0362
	(.0296)	(.0614)
Flexible Bonus	.0893*	-.0153
	(.0519)	(.0592)
N	353	352
Mean of control group	.104	.775

Notes: This table examines the impact of performance pay on inspectors' perception of monitoring by supervisors. We use instrumental variables regressions, where treatment status is instrumented with randomization results. The unit of observation is a circle, as defined at the end of Year 2. Pressure from supervisors was assessed on a 5-point Likert scale. Level of monitoring was assessed on a 4 point scale. The Information treatment is included in the controls. Robust standard errors in parentheses. Standard errors are clustered by a robust partition of circles, i.e. the group of circles such that all circles that merged or split with each other are included within the same partition. * p<0.10, ** p<0.05, *** p<0.01

Table 23: Newly Constructed Properties vs. Reassessed New Properties

Panel A: Easy to observe characteristics

	(1) Percent Area Commercial	(2) Total Covered Area	(3) Number of Floors
Newly Constructed Properties in gen. pop. sample	0.018	-90.537	0.072**
	(0.023)	(200.370)	(0.035)
Newly Assessed	0.217***	262.325	0.049**
	(0.017)	(418.175)	(0.021)
N	14285	14393	14393
Mean of non-new properties in gen. pop. sample	0.347	2577.180	1.579
Newly Constructed = Reassessed New	0.000	0.398	0.540

Panel B: Harder to observe characteristics

	(1) Percent Area Rented	(2) Percent Area Rented	(3) Number of Years Occupying Property	(4) Number of Years Occupying Property
Newly Constructed Properties in gen. pop. sample	0.030	0.022	-12.894***	-12.827***
	(0.022)	(0.020)	(0.694)	(0.676)
Newly Assessed	0.189***	0.103***	-11.325***	-10.624***
	(0.014)	(0.012)	(0.535)	(0.542)
N	14285	14284	11231	11152
Mean of non-new properties in gen. pop. sample	0.236	0.236	20.271	20.271
Newly Constructed = Reassessed New	0.000	0.000	0.032	0.003
Controls for Panel A characteristics	No	Yes	No	Yes

Notes: This table compares properties in the general population sample constructed before 2010 with 1) properties built in or after 2010; and 2) properties marked as newly assessed in the Section 9 register. 'Newly constructed properties' refers to the former and 'Newly Assessed' refers to the latter. Specification includes circle fixed effects. Panel A uses easily observable characteristics as dependent variables. Panel B uses harder to observe charateristics as dependent variables, also including variables from Panel A as controls. Standard errors are clustered by robust partition of circles, the group of circles such that all circles that merged or split with each other are included within the same partition. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Figure F.1: Heterogeneity in Reassessment Probability by Tax Density, Treatment vs. Control
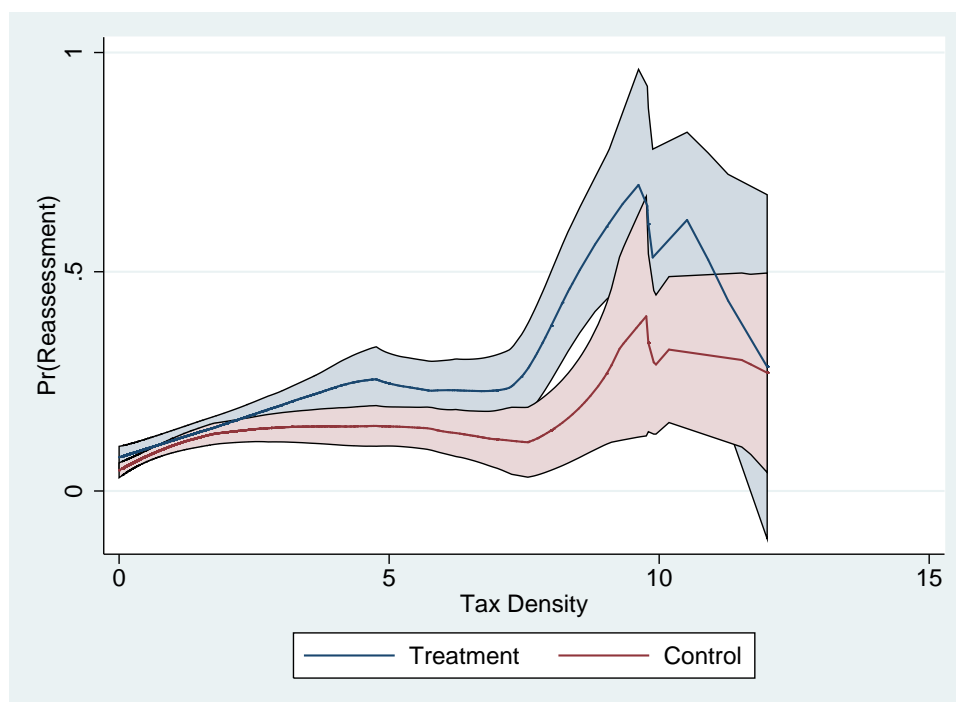
Notes: This figure plots a local linear regression of the likelihood of appearing in the reassessed sample against tax density (average tax value per square foot of area). Observations are weighted by their respective sampling probabilities. Confidence intervals at the 95% level (given by the shaded regions) are based on standard errors that are clustered by robust partition, and bootstrapped (point-wise) over 1000 iterations.

G   Extensions to Main Tables and Robustness Checks

This section runs a series of data and specification robustness checks.

THIS PAGE IS INTENTIONALLY LEFT BLANK

Table 2-G: Summary Statistics, Extended Version

	Mean	SD	N
Panel A: Administrative Data			
Log Revenue (Total) FY 2012	15.67	0.75	481
Log Revenue (Current) FY 2012	15.37	0.72	481
Log Revenue (Arrears) FY 2012	14.07	1.21	481
Log Tax Base (Total) FY 2012	16.12	0.82	477
Log Tax Base (Current) FY 2012	15.77	0.71	477
Log Tax Base (Arrears) FY 2012	14.56	1.43	477
Log Non-Exemption Rate (Total) FY 2012	-0.22	0.18	477
Log Non-Exemption Rate (Current) FY 2012	-0.19	0.12	477
Log Non-Exemption Rate (Arrears) FY 2012	-0.22	0.31	477
Log Recovery Rate (Total) FY 2012	-0.23	0.22	481
Log Recovery Rate (Current) FY 2012	-0.20	0.19	481
Log Recovery Rate (Arrears) FY 2012	-0.26	0.32	481
Log Revenue (Total) FY 2013	15.75	0.74	482
Log Revenue (Current) FY 2013	15.52	0.73	482
Log Revenue (Arrears) FY 2013	13.91	1.17	479
Log Tax Base (Total) FY 2013	16.14	0.80	482
Log Tax Base (Current) FY 2013	15.86	0.73	482
Log Tax Base (Arrears) FY 2013	14.40	1.37	479
Log Non-Exemption Rate (Total) FY 2013	-0.23	0.20	482
Log Non-Exemption Rate (Current) FY 2013	-0.19	0.13	482
Log Non-Exemption Rate (Arrears) FY 2013	-0.30	0.41	479
Log Recovery Rate (Total) FY 2013	-0.16	0.18	482
Log Recovery Rate (Current) FY 2013	-0.14	0.14	482
Log Recovery Rate (Arrears) FY 2013	-0.19	0.29	479
Panel B: Survey Data			
Could the property be located on the official tax rolls?	0.84	0.37	11,971
Quality of Tax Department [0-1]	0.53	0.22	6,050
Satisfaction with Tax Department [0-1]	0.55	0.23	6,050
Inaccuracy	0.34	0.27	9,870
Tax Gap	-0.099	0.42	9,870
(sum) GARV	32,302	252,426	10,787
Self-reported tax payment in FY 2013	3,562	18,604	12,000
Self-reported tax payment in FY 2011	3,011	18,007	11,586
Degree of Corruption in Tax Department [0-1]	0.64	0.22	6,050
Bribe Payment	2,073	3,932	5,993
Frequency of Bribe Payment	0.76	0.88	4,802
Number of floors	1.60	0.66	12,000
Last Renovation was $\leq$ 2 years ago	0.017	0.13	12,000
Land Area (sq. ft.)	280	742	12,000
Total Covered Area (sq. ft.)	2,570	18,515	12,000
On Main Road	0.46	0.50	12,000
Taxation category (1-7)	3.70	1.60	10,787
Fraction of covered area - Commercial	0.35	0.42	11,912
Fraction of covered area - Commercial and Rented	0.16	0.33	11,913
Age of owner	51	11	9,222
Owner's level of education	9.30	5.30	11,934
Per-capita wages	16,566	16,662	9,459
Predicted values of expenditure given assets owned	6,215	2,964	9,600
Connected to Politician	0.05	0.22	12,000
Connected to Politician/Government/Police	0.35	0.48	12,000
Panel C: Inspector Survey Data			
How would you score your circle team effort in your duties?	90	14	371
In a typical work day how many hours did you spend in the field	5.40	0.92	371
In a typical work day how many hours did you spend in the office	2.80	0.97	371
Total hours worked in a typical day (field + office)	8.10	0.54	371
Pressure from supervisors	0.10	0.20	371
Level of monitoring	0.77	0.29	370

Notes: Panel A statistics from administrative data, shown for both the first year (FY 2012) and the second year (FY 2013) of the study. Each observation is one of the 482 circles as defined at the time of randomization. Panel B statistics from the property survey are for randomly sampled properties only. Subjective variables - i.e., Quality, Satisfaction, Degree of Corruption, Bribe Payment, and Frequency of Bribe Payment - are reported for circles from the first phase of the survey only (see text for more details). Panel C statistics from the inspector survey.

Table 3-G1: Impacts on Revenue Collected, Reduced Form Estimates

	Year 1			Year 2		
	(1) Total	(2) Current	(3) Arrears	(4) Total	(5) Current	(6) Arrears
Panel A: Main Treatment						
Any treatment	0.057***	0.046***	0.096**	0.060***	0.058***	0.072
	(0.018)	(0.017)	(0.044)	(0.020)	(0.021)	(0.053)
Panel B: Subtreatments						
Revenue	0.078***	0.073***	0.087	0.085***	0.100***	0.001
	(0.024)	(0.023)	(0.067)	(0.029)	(0.030)	(0.088)
Revenue Plus	0.049	0.053	0.041	0.060**	0.052	0.114
	(0.034)	(0.033)	(0.070)	(0.030)	(0.032)	(0.074)
Flexible Bonus	0.045*	0.014	0.159**	0.036	0.022	0.099
	(0.025)	(0.024)	(0.065)	(0.028)	(0.029)	(0.072)
N	481	481	481	482	482	479
Mean of control group	15.671	15.379	14.030	15.745	15.518	13.915
Rev. vs. Multitasking p.	0.286	0.161	0.860	0.241	0.055	0.261
Objective vs. Subjective p.	0.538	0.091	0.207	0.244	0.097	0.618
Equality of Schemes	0.532	0.137	0.408	0.387	0.101	0.527
Joint significance	0.005	0.010	0.076	0.014	0.007	0.304

Notes: This table re-estimates Table 3 from the main text, showing the reduced form estimates instead. See Notes to Table 3 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 3-G2: Impacts on Revenue Collected, Dropping Circles with Boundary Changes

	Year 1			Year 2		
	(1) Total	(2) Current	(3) Arrears	(4) Total	(5) Current	(6) Arrears
Panel A: Main Treatment						
Any treatment	0.057**	0.033	0.128*	0.083***	0.071**	0.101
	(0.024)	(0.023)	(0.076)	(0.031)	(0.030)	(0.096)
Panel B: Subtreatments						
Revenue	0.092***	0.079**	0.118	0.124***	0.141***	-0.037
	(0.035)	(0.034)	(0.112)	(0.043)	(0.044)	(0.149)
Revenue Plus	0.032	0.036	0.019	0.068	0.047	0.161
	(0.035)	(0.033)	(0.116)	(0.044)	(0.037)	(0.139)
Flexible Bonus	0.043	-0.023	0.259**	0.057	0.024	0.179
	(0.037)	(0.036)	(0.112)	(0.044)	(0.044)	(0.124)
N	364	364	364	365	365	362
Mean of control group	15.665	15.389	13.977	15.744	15.528	13.911
Rev. vs. Multitasking p.	0.171	0.057	0.868	0.186	0.021	0.192
Objective vs. Subjective p.	0.647	0.041	0.128	0.405	0.120	0.395
Equality of Schemes	0.373	0.076	0.251	0.409	0.069	0.416
Joint significance	0.061	0.059	0.113	0.026	0.013	0.366

Notes: This table re-estimates Table 3 from the main text, dropping circles that have experienced any change in circle boundaries during the course of treatment. See Notes to Table 3 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-G1: Impacts on Non-Revenue Outcomes, Reduced Form Estimates

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Panel A: Main Treatment				
Any treatment	-0.005	-0.008	0.003	0.005
	(0.016)	(0.017)	(0.009)	(0.016)
Panel B: Subtreatments				
Revenue	0.005	-0.004	0.001	-0.017
	(0.027)	(0.029)	(0.013)	(0.022)
Revenue Plus	0.031	0.022	0.020*	0.011
	(0.019)	(0.020)	(0.012)	(0.023)
Flexible Bonus	-0.044*	-0.039*	-0.012	0.021
	(0.023)	(0.023)	(0.013)	(0.023)
N	6050	6050	9870	9870
Sample	Phase 1	Phase 1	Full	Full
Mean of control group	0.538	0.555	0.339	-0.103
Rev. vs. Multitasking p.	0.699	0.895	0.817	0.163
Objective vs. Subjective p.	0.015	0.064	0.103	0.312
Equality of Schemes	0.013	0.056	0.094	0.351
Joint significance	0.033	0.122	0.165	0.535

Notes: This table re-estimates Table 4 from the main text, showing reduced form estimates instead. See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-G2: Impacts on Non-Revenue Outcomes, Controlling for Objective Property Characteristics

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Panel A: Main Treatment				
Any treatment	-0.005	-0.010	0.004	0.006
	(0.022)	(0.022)	(0.012)	(0.019)
Panel B: Subtreatments				
Revenue	0.005	-0.006	0.005	-0.013
	(0.035)	(0.037)	(0.016)	(0.025)
Revenue Plus	0.043*	0.031	0.024	-0.010
	(0.025)	(0.026)	(0.016)	(0.029)
Flexible Bonus	-0.058*	-0.051	-0.016	0.041
	(0.031)	(0.031)	(0.017)	(0.028)
N	5998	5998	9869	9869
Sample	Phase 1	Phase 1	Full	Full
Mean of control group	0.538	0.555	0.339	-0.103
Rev. vs. Multitasking p.	0.732	0.924	0.975	0.317
Objective vs. Subjective p.	0.014	0.062	0.084	0.085
Equality of Schemes	0.011	0.051	0.131	0.217
Joint significance	0.028	0.113	0.234	0.374

Notes: Re-estimation of Table 4 from the main text, including controls for property characteristics. Property controls include: land area, total covered area, a dummy for whether the property was located on a main road, number of floors, a dummy for whether the property had a renovation in the past 2 years, and the percent of covered area dedicated to each of the four main usage categories (residential vs. commercial, owner occupied vs. rented). See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-G3: Impacts on Non-Revenue Outcomes, Controlling for Perceptions of Electricity Bureau

	(1) Quality	(2) Satisfaction
Panel A: Main Treatment		
Any treatment	0.001	-0.003
	(0.019)	(0.019)
Panel B: Subtreatments		
Revenue	0.008	-0.004
	(0.031)	(0.030)
Revenue Plus	0.025	0.017
	(0.020)	(0.020)
Flexible Bonus	-0.028	-0.020
	(0.029)	(0.029)
N	4840	4840
Sample	Phase 1	Phase 1
Mean of control group	0.529	0.543
Rev. vs. Multitasking p.	0.778	0.945
Objective vs. Subjective p.	0.154	0.375
Equality of Schemes	0.188	0.386
Joint significance	0.283	0.570

Notes: This table re-estimates Table 4 from the main text, controlling for perceptions of the electricity department. Column (1) controls for perceptions of electricity department quality, and Column (2) controls for electricity department satisfaction. See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-G4: Impacts on Non-Revenue Outcomes, Ordered Probit Specifications

	(1) Quality	(2) Satisfaction
Panel A: Main Treatment		
Any treatment	-0.042	-0.064
	(0.090)	(0.091)
Panel B: Subtreatments		
Revenue	0.059	0.008
	(0.151)	(0.153)
Revenue Plus	0.093	0.050
	(0.106)	(0.110)
Flexible Bonus	-0.242*	-0.224*
	(0.133)	(0.133)
N	6050	6050
Sample	Phase 1	Phase 1
Mean of control group	3.153	3.220
Rev. vs. Multitasking p.	0.399	0.554
Objective vs. Subjective p.	0.029	0.084
Equality of Schemes	0.068	0.181
Joint significance	0.145	0.304

Notes: This table re-estimates columns 1 and 2 of Table 4 from the main text, using ordered probit regressions instead. See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

H Separating Effects by Subtreatment

This section presents some of the secondary results in the paper separately for each subtreatment.

Table 6-H: Impacts on Tax Payments and Corruption, by Subtreatment

	Self-reported Tax Payment	Bribe Payment	Frequency of Bribe Payment	Perception of Corruption
Panel A: General Population Sample Only				
Revenue	490	738	.203	-.0431
	(553)	(611)	(.157)	(.0341)
Revenue Plus	-721***	101	.118	.016
	(264)	(422)	(.126)	(.0331)
Flexible Bonus	16.7	938**	.28*	.0507
	(250)	(428)	(.149)	(.032)
N	11586	5993	4802	6050
Mean of control group	4069.4	1874.5	0.7	0.6
Rev. vs. Multitasking p.	0.119	0.714	0.983	0.023
Objective vs. Subjective p.	0.697	0.279	0.471	0.037
Equality of Schemes	0.016	0.270	0.651	0.043
Joint significance	0.018	0.152	0.185	0.091
Panel B: Re-assessed and General Population Sample				
Re-assessed * Revenue	371	-457	-.0996	-.0155
	(1567)	(598)	(.173)	(.0333)
Re-assessed * Revenue Plus	2195*	-363	-.208	.0115
	(1270)	(379)	(.128)	(.0298)
Re-assessed * Flexible Bonus	3024*	-802	-.159	-.00682
	(1594)	(594)	(.13)	(.0334)
Re-assessed	2762***	-65.1	.0134	-.019*
	(573)	(178)	(.0403)	(.0107)
N	16353	8207	6993	8268
Sample	Full	Phase 1	Phase 1	Phase 1
Mean of control group in gen. pop. sample	3928.3	1874.5	0.7	0.6
Rev. vs. Multitasking p.	0.162	0.832	0.643	0.620
Objective vs. Subjective p.	0.280	0.504	0.971	0.894
Equality of Schemes	0.364	0.772	0.862	0.782
Joint significance	0.150	0.515	0.316	0.918

Notes: This table re-estimates Table 6 from the main text, estimating impacts separately by subtreatment. See Notes to Table 6 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 7-H: Impacts on Satisfaction and Accuracy by Reassessed Status, by Subtreatment

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Re-assessed * Revenue	-0.011	-0.009	0.005	-0.014
	(0.033)	(0.035)	(0.026)	(0.040)
Re-assessed * Revenue Plus	0.009	0.003	-0.026	0.011
	(0.033)	(0.031)	(0.022)	(0.041)
Re-assessed * Flexible Bonus	0.023	0.017	0.023	-0.011
	(0.034)	(0.033)	(0.025)	(0.039)
Re-assessed	0.049***	0.044***	-0.061***	0.122***
	(0.013)	(0.013)	(0.009)	(0.015)
N	8268	8268	14173	14173
Sample	Phase 1	Phase 1	Full	Full
Mean of control group in gen. pop. sample	0.538	0.555	0.339	-0.103

Notes: This table re-estimates Table 7 from the main text, estimating impacts separately by sub-treatment. See Notes to Table 7 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 14-H: Impacts on Tax Base and Recovery Rates, by Subtreatment

	Year 1				Year 2			
	(1)	(2)	(3) Non-Exemption Rate	(4) Recovery Rate	(5)	(6)	(7) Non-Exemption Rate	(8) Recovery Rate
	Revenue	Tax Base			Revenue	Tax Base		
Panel A: Total								
Revenue	0.122***	0.090**	-0.019	0.050*	0.131***	0.060	0.020	0.050**
	(0.033)	(0.039)	(0.028)	(0.028)	(0.043)	(0.050)	(0.030)	(0.024)
Revenue Plus	0.068	0.108**	-0.009	-0.031	0.093**	0.056	0.015	0.022
	(0.051)	(0.050)	(0.027)	(0.039)	(0.043)	(0.047)	(0.024)	(0.030)
Flexible Bonus	0.033	0.070	-0.046	0.010	0.041	0.043	-0.016	0.014
	(0.035)	(0.046)	(0.035)	(0.036)	(0.040)	(0.045)	(0.032)	(0.023)
Panel B: Current								
Revenue	0.113***	0.086**	0.013	0.014	0.162***	0.108**	0.026	0.029
	(0.035)	(0.036)	(0.016)	(0.029)	(0.043)	(0.043)	(0.027)	(0.022)
Revenue Plus	0.081	0.105*	0.011	-0.035	0.079	0.064	0.009	0.005
	(0.053)	(0.058)	(0.019)	(0.041)	(0.049)	(0.050)	(0.023)	(0.026)
Flexible Bonus	0.021	0.059*	-0.023	-0.016	0.042	0.026	0.017	-0.001
	(0.036)	(0.032)	(0.026)	(0.034)	(0.041)	(0.036)	(0.021)	(0.018)
Panel C: Arrears								
Revenue	0.161*	0.152*	-0.063	0.069	0.005	-0.127	0.047	0.084**
	(0.088)	(0.088)	(0.056)	(0.045)	(0.128)	(0.136)	(0.064)	(0.040)
Revenue Plus	0.024	0.079	-0.004	-0.055	0.137	0.008	0.142**	-0.012
	(0.107)	(0.117)	(0.053)	(0.051)	(0.109)	(0.117)	(0.057)	(0.056)
Flexible Bonus	0.142	0.164	-0.089	0.078	0.077	0.097	-0.031	0.012
	(0.096)	(0.110)	(0.057)	(0.067)	(0.103)	(0.124)	(0.066)	(0.047)
N (Total)	473	470	470	473	474	474	474	474
Mean of control group (Total)	15.670	16.108	-0.204	-0.228	15.742	16.140	-0.230	-0.168

Notes: This table re-estimates Appendix Table 14, estimating impacts seperately by subtreatment. See Notes to Appendix Table 14 for additional information. * p<0.10, ** p<0.05, *** p<0.01

I Information Treatment Separated from the Control Group

This section replicates the main tables, excluding the information treatment from the control group. These tables show no qualitative differences in the results by doing so.

Table 3-I: Impacts on Revenue Collected, Separating Information Treatment

	Year 1			Year 2		
	(1) Total	(2) Current	(3) Arrears	(4) Total	(5) Current	(6) Arrears
Panel A: Main Treatment						
Any treatment	0.091***	0.073***	0.152**	0.110***	0.108***	0.166*
	(0.028)	(0.027)	(0.069)	(0.034)	(0.035)	(0.093)
Information				0.068	0.075	0.232*
				(0.052)	(0.051)	(0.138)
Panel B: Subtreatments						
Revenue	0.118***	0.109***	0.134	0.144***	0.168***	0.056
	(0.035)	(0.034)	(0.099)	(0.045)	(0.046)	(0.139)
Revenue Plus	0.080	0.086*	0.072	0.107**	0.097*	0.226*
	(0.053)	(0.052)	(0.110)	(0.046)	(0.050)	(0.119)
Flexible Bonus	0.071*	0.024	0.243**	0.071*	0.051	0.198*
	(0.038)	(0.035)	(0.098)	(0.042)	(0.044)	(0.114)
Information				0.064	0.071	0.226*
				(0.051)	(0.050)	(0.136)
N	481	481	481	482	482	479
Mean of control group	15.671	15.379	14.030	15.727	15.507	13.860
Rev. vs. Multitasking p.	0.323	0.193	0.830	0.232	0.049	0.263
Objective vs. Subjective p.	0.530	0.090	0.212	0.217	0.082	0.636
Equality of Schemes	0.562	0.143	0.433	0.356	0.084	0.528
Joint significance	0.004	0.010	0.073	0.006	0.002	0.163

Notes: This table re-estimates Table 3 from the main text, separating the Information treatment from the control group. See Notes to Table 3 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-I: Impacts on Non-Revenue Outcomes, Separating Information Treatment

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Panel A: Main Treatment				
Any treatment	-0.008	-0.013	0.006	0.009
	(0.025)	(0.026)	(0.013)	(0.024)
Information	-0.006	-0.005	0.005	0.009
	(0.031)	(0.033)	(0.022)	(0.037)
Panel B: Subtreatments				
Revenue	0.004	-0.007	0.003	-0.020
	(0.037)	(0.039)	(0.018)	(0.031)
Revenue Plus	0.038	0.027	0.029*	0.017
	(0.029)	(0.029)	(0.017)	(0.033)
Flexible Bonus	-0.062*	-0.054	-0.015	0.032
	(0.034)	(0.035)	(0.019)	(0.034)
Information	-0.007	-0.005	0.005	0.009
	(0.031)	(0.033)	(0.022)	(0.037)
N	6050	6050	9870	9870
Sample	Phase 1	Phase 1	Full	Full
Mean of control group	0.541	0.559	0.339	-0.105
Rev. vs. Multitasking p.	0.681	0.874	0.810	0.158
Objective vs. Subjective p.	0.015	0.064	0.100	0.314
Equality of Schemes	0.014	0.060	0.091	0.342
Joint significance	0.036	0.129	0.158	0.527

Notes: This table re-estimates Table 4 from the main text, separating the Information treatment from the control group. See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 5-I: Impacts on Number of Reassessed Properties, Separating Information Treatment

	Total Number of Section 9 Properties Added to Tax Rolls in Treatment Period	Number of New Properties Added to Tax Rolls in Treatment Period	Number of Reassessed Properties Added to Tax Rolls in Treatment Period
Treatment	85.28*	77.23**	8.05
	(47.42)	(34.00)	(25.07)
Information	10.58	15.17	-4.58
	(68.06)	(34.19)	(46.21)
N	234	234	234
Mean of control group	98.7	36.1	62.6

Notes: This table re-estimates Table 5 from the main text, separating the Information treatment from the control group. See Notes to Table 5 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 6-I: Impacts on Tax Payments and Corruption by Reassessed Status, Separating Information Treatment

	(1) Self-reported Tax Payment	(2) Bribe Payment	(3) Frequency of Bribe Payment	(4) Perception of Corruption
Panel A: General Population Sample Only				
Treatment	-85.88	720.1**	.2212**	.024
	(293.1)	(356.7)	(.1018)	(.0303)
Information	-97.41	470.7	.071	.0476
	(318.2)	(546.5)	(.1325)	(.0435)
N	11586	5993	4802	6050
Mean of control group	4249.973	1806.560	0.668	0.637
Panel B: Re-assessed and General Population Sample				
Re-assessed * Treatment	1970	-345.3	-.149	-.0024
	(1222)	(377.3)	(.0979)	(.0246)
Re-assessed * Information	363.4	804.6	.0381	.0026
	(1913)	(715.8)	(.1494)	(.0334)
Re-assessed	2700***	-224.7	.0061	-.0196
	(726)	(170.7)	(.0446)	(.0134)
N	16353	8207	6993	8268
Sample	Full	Phase 1	Phase 1	Phase 1
Mean of control group in gen. pop. sample	4119.372	1806.560	0.668	0.637

Notes: This table re-estimates Table 6 from the main text, separating the Information treatment from the control group. See Notes to Table 6 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 7-I: Impacts on Satisfaction and Accuracy by Reassessed Status, Separating Information Treatment

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Re-assessed * Treatment	0.031	0.020	0.008	0.003
	(0.027)	(0.026)	(0.019)	(0.031)
Re-assessed * Information	0.086**	0.057	0.029	0.034
	(0.043)	(0.043)	(0.033)	(0.049)
Re-assessed	0.032**	0.033**	-0.066***	0.117***
	(0.016)	(0.015)	(0.011)	(0.018)
N	8268	8268	14173	14173
Sample	Phase 1	Phase 1	Full	Full
Mean of control group in gen. pop. sample	0.541	0.559	0.339	-0.105

Notes: This table re-estimates Table 7 from the main text, separating the Information treatment from the control group. See Notes to Table 7 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 8-I: Selection Effects in Reassessment, Separating Information Treatment

Panel A

	GARV	Number of floors	Last renovation was ≤ 2 years ago	Components of GARV						Tax Liability
				Land area (sq. feet)	Total covered area, all uses	Main Road	Category	Percent of property commercial	Percent of property commercial and rented	
Re-assess * Treatment	14226.650	-0.018	-0.022	-63.561	456.299	-0.018	-0.209**	0.023	0.072**	2725.417
	(17371.383)	(0.056)	(0.023)	(109.336)	(697.532)	(0.053)	(0.095)	(0.040)	(0.031)	(3815.833)
Re-assess * Information	-25525.546	-0.083	-0.073**	-131.343	-1678.964	-0.065	0.073	0.023	-0.014	-5063.368
	(28195.863)	(0.083)	(0.033)	(132.493)	(1780.996)	(0.079)	(0.159)	(0.067)	(0.053)	(6158.812)
Re-assess dummy	29053.615***	0.092***	0.107***	60.258	135.634	0.075***	0.199***	0.213***	0.178***	6370.336***
	(8904.727)	(0.032)	(0.013)	(78.042)	(239.663)	(0.029)	(0.051)	(0.022)	(0.017)	(1975.343)
N	15090	16352	16354	16352	16352	16352	15090	16226	16227	15090
Mean of control group in gen. pop. sample	35039.319	1.563	0.020	313.886	2512.189	0.463	3.789	0.360	0.168	6257.659

Panel B

	Approximate age of owner	Owner's level of education	Per-capita wages	Predicted expenditure given assets	Connected to Politician	Connected to Politician/ Government/ Police
Re-assess * Treatment	-0.210	-0.385	-982.247	189.045	0.012	0.007
	(0.859)	(0.336)	(1193.329)	(242.458)	(0.013)	(0.028)
Re-assess * Information	0.571	0.585	-676.670	331.684	-0.038*	0.007
	(1.446)	(0.579)	(1627.107)	(401.868)	(0.020)	(0.055)
Re-assess dummy	-0.758	0.201	131.347	-152.279	-0.006	0.004
	(0.461)	(0.178)	(632.992)	(149.176)	(0.007)	(0.015)
N	13406	16254	13765	13954	16354	16354
Mean of control group in gen. pop. sample	50.696	9.231	16343.153	6273.619	0.051	0.373

Notes: This table re-estimates Table 8 from the main text, separating the Information treatment from the control group. See Notes to Table 8 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 14-I: Impacts on Tax Base and Recovery Rates, Separating Information Treatment

| | Year 2 | | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
	Revenue	Tax Base	Non-Exemption Rate	Recovery Rate	Revenue	Tax Base	Non-Exemption Rate	Recovery Rate
Total								
Any Treatment	0.076***	0.090***	-0.025	0.011	0.100***	0.045	0.021	0.034
	(0.027)	(0.029)	(0.018)	(0.023)	(0.032)	(0.035)	(0.020)	(0.021)
Information					0.044	-0.037	0.061*	0.019
					(0.050)	(0.061)	(0.036)	(0.037)
Curent								
Any Treatment	0.073***	0.084***	0.000	-0.012	0.107***	0.068**	0.022	0.017
	(0.028)	(0.028)	(0.014)	(0.022)	(0.034)	(0.032)	(0.017)	(0.018)
Information					0.050	0.005	0.018	0.027
					(0.048)	(0.043)	(0.024)	(0.030)
Arrears								
Any Treatment	0.111*	0.133*	-0.053	0.032	0.113	0.023	0.061	0.029
	(0.065)	(0.069)	(0.036)	(0.036)	(0.091)	(0.099)	(0.047)	(0.037)
Information					0.162	0.123	0.034	0.005
					(0.135)	(0.160)	(0.114)	(0.069)
N (total)	473	470	470	473	474	474	474	474
Mean of control group	15.680	16.114	-0.201	-0.225	15.739	16.129	-0.227	-0.164

Notes: This table re-estimates Appendix Table 14, separating the Information treatment from the control group. See Notes to Appendix Table 14 for additional information. * p<0.10, ** p<0.05, *** p<0.01

J Dropping Revenue Plus Circles

While the Revenue Plus treatment was not unbalanced on any individual outcome variables, the balance tests in Table 11 suggested that one could reject balance for the joint test for the Revenue Plus treatment. Therefore these tables replicate our main results by excluding this treatment, to check whether average treatment effects are robust. We find that these results continue to hold.

Table 3-J: Impacts on Revenue Collected, Dropping Revenue Plus Circles

	Year 1			Year 2		
	(1) Total	(2) Current	(3) Arrears	(4) Total	(5) Current	(6) Arrears
Panel A: Main Treatment						
Any treatment	0.093***	0.066**	0.188**	0.092***	0.092**	0.082
	(0.029)	(0.029)	(0.077)	(0.035)	(0.036)	(0.096)
Panel B: Subtreatments						
Revenue	0.118***	0.109***	0.135	0.130***	0.153***	0.002
	(0.036)	(0.035)	(0.100)	(0.043)	(0.044)	(0.132)
Flexible Bonus	0.074**	0.027	0.248**	0.059	0.038	0.150
	(0.037)	(0.035)	(0.096)	(0.040)	(0.041)	(0.107)
N	430	430	430	411	411	408
Mean of control group	15.671	15.379	14.030	15.745	15.518	13.915
Equality of Schemes	0.002	0.008	0.025	0.009	0.003	0.354
Joint significance	0.002	0.008	0.025	0.009	0.003	0.354

Notes: This table re-estimates Table 3 from the main text, dropping Revenue Plus circles. See Notes to Table 3 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-J: Impacts on Non-Revenue Outcomes, Dropping Revenue Plus Circles

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Panel A: Main Treatment				
Any treatment	-0.029	-0.031	-0.008	0.007
	(0.026)	(0.027)	(0.014)	(0.024)
Panel B: Subtreatments				
Revenue	0.007	-0.005	0.003	-0.025
	(0.036)	(0.037)	(0.017)	(0.029)
Flexible Bonus	-0.061*	-0.054*	-0.016	0.029
	(0.031)	(0.032)	(0.018)	(0.031)
N	5150	5150	8398	8398
Sample	Phase 1	Phase 1	Full	Full
Mean of control group	0.538	0.555	0.339	-0.103
Equality of Schemes	0.135	0.234	0.605	0.345
Joint significance	0.135	0.234	0.605	0.345

Notes: This table re-estimates Table 4 from the main text, dropping Revenue Plus circles. See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 5-J: Impacts on Number of Reassessed Properties, Dropping Revenue Plus Circles

	(1) Total Number of Section 9 Properties Added to Tax Rolls in Treatment Period	(2) Number of New Properties Added to Tax Rolls in Treatment Period	(3) Number of Reassessed Properties Added to Tax Rolls in Treatment Period
Treatment	102.4* (56.90)	98.6** (45.73)	3.8 (24.37)
N	202	202	202
Mean of control group	96.7	36.7	60.0

Notes: This table re-estimates Table 5 from the main text, separating the Information treatment from the control group. See Notes to Table 5 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 6-J: Impacts on Tax Payments and Corruption by Reassessed Status, Dropping Revenue Plus Circles

	(1) Self-reported Tax Payment	(2) Bribe Payment	(3) Frequency of Bribe Payment	(4) Perception of Corruption
Panel A: General Population Sample Only				
Treatment	247.2 (331.4)	854.8** (412.5)	.2504** (.1153)	.0094 (.0277)
N	9928	5096	4086	5150
Mean of control group	4069.425	1874.542	0.683	0.644
Panel B: Re-assessed and General Population Sample				
Re-assessed * Treatment	1675 (1259)	-662.1 (482.2)	-.1353 (.1134)	-.0105 (.026)
Re-assessed	2781*** (572.5)	-72.11 (178)	.0134 (.0404)	-.0178* (.0107)
N	14020	6998	5970	7056
Sample	Full	Phase 1	Phase 1	Phase 1
Mean of control group in gen. pop. sample	3928.252	1874.542	0.683	0.644

Notes: This table re-estimates Table 6 from the main text, dropping Revenue Plus circles. See Notes to Table 6 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 7-J: Impacts on Satisfaction and Accuracy by Reassessed Status, Dropping Revenue Plus Circles

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Re-assessed * Treatment	0.006 (0.028)	0.003 (0.029)	0.015 (0.020)	-0.015 (0.031)
Re-assessed	0.050*** (0.013)	0.045*** (0.013)	-0.061*** (0.009)	0.123*** (0.015)
N	7056	7056	12097	12097
Sample	Phase 1	Phase 1	Full	Full
Mean of control group in gen. pop. sample	0.538	0.555	0.339	-0.103

Notes: This table re-estimates Table 7 from the main text, dropping Revenue Plus circles. See Notes to Table 7 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 8-J: Selection Effects in Reassessment, Dropping Revenue Plus Circles

Panel A

				Components of GARV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	GARV	Number of floors	Last renovation was ≤ 2 years ago	Land area (sq. feet)	Total covered area (sq. feet)	Main Road	Tax Category	Percent of property commercial	Percent of property commercial and rented	Tax Liability
Re-assess * Treatment	23335.773	-0.002	-0.008	-41.662	1047.209	-0.045	-0.263***	0.009	0.055*	4433.625
	(20632.890)	(0.054)	(0.022)	(84.510)	(1005.980)	(0.054)	(0.101)	(0.042)	(0.033)	(4492.322)
Re-assess	24776.627***	0.079***	0.094***	37.811	-159.900	0.065***	0.211***	0.217***	0.176***	5519.381***
	(7995.795)	(0.026)	(0.011)	(57.205)	(385.660)	(0.024)	(0.044)	(0.019)	(0.015)	(1765.226)
N	12897	14019	14021	14019	14019	14019	12897	13913	13914	12897
Mean of control group in gen. pop. sample	36808.77	1.57	0.02	301.13	2779.82	0.46	3.78	0.35	0.17	6642.00

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
	Approximate age of owner	Owner's level of education	Per-capita wages	Predicted expenditure given assets	Connected to Politician	Connected to Politician/ Government/ Police
Re-assess * Treatment	-0.429	-0.458	-1819.317	73.750	0.027*	-0.002
	(0.927)	(0.352)	(1311.592)	(241.537)	(0.014)	(0.031)
Re-assess	-0.656	0.308*	20.685	-94.246	-0.013**	0.005
	(0.398)	(0.158)	(509.982)	(122.393)	(0.006)	(0.014)
N	11488	13948	11812	11966	14021	14021
Mean of control group in gen. pop. sample	50.70	9.19	16281.55	6292.58	0.05	0.36

Notes: This table re-estimates Table 8 from the main text, dropping Revenue Plus circles. See Notes to Table 8 for additional information.
* p<0.10, ** p<0.05, *** p<0.01

Table 14-J: Impacts on Tax Base and Recovery Rates, Dropping Revenue Plus Circles

	Year 1				Year 2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Revenue	Tax Base	Non-Exemption Rate	Recovery Rate	Revenue	Tax Base	Non-Exemption Rate	Recovery Rate
Total								
Any Treatment	0.074***	0.077**	-0.034	0.031	0.084**	0.049	0.002	0.033
	(0.027)	(0.032)	(0.023)	(0.025)	(0.034)	(0.038)	(0.024)	(0.020)
Curent								
Any Treatment	0.067**	0.074***	-0.004	-0.004	0.102***	0.067**	0.021	0.013
	(0.029)	(0.027)	(0.017)	(0.024)	(0.035)	(0.032)	(0.019)	(0.017)
Arrears								
Any Treatment	0.147**	0.153**	-0.080*	0.077*	0.044	-0.018	0.016	0.047
	(0.072)	(0.076)	(0.043)	(0.043)	(0.093)	(0.104)	(0.053)	(0.037)
N (Total)	423	420	420	423	404	404	404	404
Mean of control group (Total)	15.680	16.114	-0.201	-0.225	15.756	16.149	-0.229	-0.165

Notes: This table re-estimates Appendix Table 14, dropping Revenue Plus circles. See Notes to Appendix Table 14 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 20-J: Mechanisms Beyond Price Effects, Dropping Revenue Plus Circles

Panel A

	(1) Total	(2) Current	(3) Arrears
Information	0.065	0.072	0.212
	(0.051)	(0.050)	(0.135)
N	411	411	408
Mean of control group	15.709	15.486	13.864

Panel B

	(1) Total	(2) Current	(3) Arrears
Revenue * Income Shock	0.0109	0.0379	-0.0588
	(0.115)	(0.117)	(0.242)
Flexible Bonus * Income Shock	-0.0181	0.0257	-0.124
	(0.0945)	(0.0671)	(0.312)
Income Shock	-0.0353	0.0228	0.150*
	(0.0554)	(0.0443)	(0.0847)
N	427	427	427

Panel C

	(1) Total	(2) Current	(3) Arrears
Supervisory treatment	-0.054	-0.104	0.022
	[0.413]	[0.227]	[0.911]
N	411	411	408
Mean of control group	15.926	15.686	14.160

Notes: This table re-estimates Appendix Table 20, dropping Revenue Plus circles. See Notes to Appendix Table 20 for additional information. * p<0.10, ** p<0.05, *** p<0.01

K Controlling for Variables from Balance Check

While the Revenue Plus treatment was not unbalanced on any individual outcome variables, the balance tests in Table 11 suggested that one could reject balance for the joint test for the Revenue Plus treatment. Therefore these tables replicate our main results by controlling for variables from the balance check, to check whether the estimates on the Revenue Plus treatment are robust. We find that these results continue to hold.

Table 3-K: Impacts on Revenue Outcomes, Controlling for Balance Check Variables

	Year 1			Year 2		
	(1) Total	(2) Current	(3) Arrears	(4) Total	(5) Current	(6) Arrears
Panel A: Main Treatment						
Any treatment	0.061**	0.052**	0.118*	0.088***	0.085***	0.066
	(0.024)	(0.023)	(0.065)	(0.028)	(0.029)	(0.080)
Panel B: Subtreatments						
Revenue	0.113***	0.101***	0.150*	0.149***	0.171***	-0.003
	(0.033)	(0.034)	(0.087)	(0.040)	(0.040)	(0.126)
Revenue Plus	0.032	0.047	0.046	0.070*	0.056	0.126
	(0.036)	(0.032)	(0.108)	(0.038)	(0.035)	(0.112)
Flexible Bonus	0.031	0.004	0.149	0.039	0.024	0.072
	(0.035)	(0.033)	(0.092)	(0.039)	(0.041)	(0.103)
N	470	470	470	471	471	468
Mean of control group	15.683	15.390	14.044	15.756	15.527	13.934
Rev. vs. Multitasking p.	0.032	0.048	0.603	0.021	0.001	0.444
Objective vs. Subjective p.	0.279	0.055	0.621	0.088	0.029	0.927
Equality of Schemes	0.099	0.087	0.666	0.058	0.006	0.693
Joint significance	0.009	0.021	0.197	0.002	0.000	0.671

Notes: This table re-estimates Table 3 from the main text, controlling for variables included in the balance checks (see Appendix Table 1 for details). See Notes to Table 3 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 4-K: Impacts on Non-Revenue Outcomes, Controlling for Balance Check Variables

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Panel A: Main Treatment				
Any treatment	0.001	-0.004	-0.000	0.011
	(0.023)	(0.023)	(0.012)	(0.022)
Panel B: Subtreatments				
Revenue	0.011	-0.001	-0.003	-0.017
	(0.036)	(0.038)	(0.017)	(0.031)
Revenue Plus	0.053**	0.042	0.027*	0.016
	(0.026)	(0.027)	(0.016)	(0.032)
Flexible Bonus	-0.055*	-0.048	-0.026	0.034
	(0.032)	(0.033)	(0.018)	(0.032)
N	6000	6000	9606	9606
Sample	Phase 1	Phase 1	Full	Full
Mean of control group	0.538	0.555	0.339	-0.103
Rev. vs. Multitasking p.	0.742	0.969	0.859	0.215
Objective vs. Subjective p.	0.009	0.052	0.043	0.301
Equality of Schemes	0.005	0.033	0.034	0.413
Joint significance	0.013	0.073	0.078	0.582

Notes: This table re-estimates Table 4 from the main text, controlling for variables included in the balance checks (see Appendix Table 11 for details). See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

L Dropping GPS Sampled Properties

Since there is a possible concern that the first point surveyed over-samples larger properties, in this section we replicate our main survey-result based tables excluding this first point. We see that the results are qualitatively unchanged.

Table 6-L: Impacts on Tax Payments and Corruption by Reassessed Status, Dropping GPS Sampled Properties

	(1) Self-reported Tax Payment	(2) Bribe Payment	(3) Frequency of Bribe Payment	(4) Perception of Corruption
Panel A: General Population Sample Only				
Treatment	-86.51	488.2	.2089**	.011
	(307.2)	(344.4)	(.0948)	(.0257)
N	9383	4819	4095	4856
Mean of control group	3883.010	1895.657	0.682	0.645
Panel B: Re-assessed and General Population Sample				
Re-assessed * Treatment	1828*	-462.9	-.1645*	-.0037
	(1103)	(381)	(.0938)	(.0224)
Re-assessed	2834***	-120.9	.0143	-.0193*
	(597.8)	(180.5)	(.0396)	(.0109)
N	14003	7033	6286	7074
Sample	Full	Phase 1	Phase 1	Phase 1
Mean of control group in gen. pop. sample	3774.129	1895.657	0.682	0.645

Notes: This table re-estimates Table 6 from the main text, dropping GPS-sampled properties (see Appendix B for details). See Notes to Table 6 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 8-L: Selection Effects in Reassessment, Dropping GPS Sampled Properties

Panel A

	(1)	(2)	(3)	Components of GARV						(10)
				(4)	(5)	(6)	(7)	(8)	(9)	
	GARV	Number of floors	Last renovation was ≤ 2 years ago	Land area (sq. feet)	Total covered area (sq. feet)	Main Road	Tax Category	Percent of property commercial	Percent of property commercial and rented	Tax Liability
Re-assess * Treatment	19543.940 (16628.064)	0.011 (0.050)	-0.007 (0.020)	-37.031 (86.931)	927.685 (844.245)	-0.004 (0.048)	-0.229*** (0.088)	0.021 (0.037)	0.081*** (0.030)	3713.816 (3633.759)
Re-assess	25745.796*** (8452.681)	0.079*** (0.026)	0.095*** (0.011)	54.480 (60.824)	-157.838 (444.591)	0.064*** (0.024)	0.209*** (0.044)	0.215*** (0.019)	0.174*** (0.015)	5717.251*** (1858.254)
N	12979	14002	14004	14002	14002	14002	12979	13896	13897	12979
Mean of control group in gen. pop. sample	35308.88	1.57	0.02	285.43	2759.23	0.46	3.78	0.35	0.17	6356.80

Panel B

	(1)	(2)	(3)	(4)	(5)	(6)
	Approximate age of owner	Owner's level of education	Per-capita wages	Predicted expenditure given assets	Connected to Politician	Connected to Politician/ Government/ Police
Re-assess * Treatment	-0.455 (0.799)	-0.543 (0.330)	-942.500 (1052.622)	127.269 (211.188)	0.021* (0.013)	0.005 (0.028)
Re-assess	-0.544 (0.398)	0.350** (0.166)	147.272 (521.266)	-83.122 (123.121)	-0.012* (0.006)	0.010 (0.014)
N	12047	13924	12380	12544	14004	14004
Mean of control group in gen. pop. sample	50.59	9.14	16163.52	6283.49	0.05	0.35

Notes: This table re-estimates Table 8 from the main text, dropping GPS-sampled properties (see Appendix B for details). See Notes to Table 8 for additional information. * p<0.10, ** p<0.05, *** p<0.01

M Dropping Renters

There is a potential concern that rented properties may not have as much information about interactions with the tax department since ultimately the owner bears the responsibility for the tax bill. This section therefore shows our results on the subjective non-revenue outcomes are similar if we just include properties occupied by owners.

Table 4-M: Impacts on Non-Revenue Outcomes, Dropping Renters

	(1) Quality	(2) Satisfaction
Panel A: Main Treatment		
Any treatment	-0.004	-0.008
	(0.022)	(0.023)
Panel B: Subtreatments		
Revenue	0.005	-0.005
	(0.037)	(0.039)
Revenue Plus	0.037	0.028
	(0.026)	(0.026)
Flexible Bonus	-0.052*	-0.046
	(0.030)	(0.032)
N	5125	5125
Sample	Phase 1	Phase 1
Mean of control group	0.540	0.556
Rev. vs. Multitasking p.	0.739	0.925
Objective vs. Subjective p.	0.027	0.095
Equality of Schemes	0.028	0.088
Joint significance	0.068	0.182

Notes: This table re-estimates Table 4 from the main text, dropping rental properties. See Notes to Table 4 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 6-M: Impacts on Tax Payments and Corruption by Reassessed Status, Dropping Renters

	(1) Self-reported Tax Payment	(2) Bribe Payment	(3) Frequency of Bribe Payment	(4) Perception of Corruption
Panel A: General Population Sample Only				
Treatment	34.4	608*	.2043**	.01
	(287.9)	(340.3)	(.1)	(.0256)
N	9769	5071	3893	5125
Mean of control group	3639.438	1847.679	0.676	0.640
Panel B: Re-assessed and General Population Sample				
Re-assessed * Treatment	826	-365.2	-.1495	-.0072
	(1017)	(403.4)	(.1027)	(.0247)
Re-assessed	2244***	-103.5	.026	-.0091
	(534.7)	(184.1)	(.0486)	(.0123)
N	12829	6419	5230	6476
Sample	Full	Phase 1	Phase 1	Phase 1
Mean of control group in gen. pop. sample	3489.919	1847.679	0.676	0.640

Notes: This table re-estimates Table 6 from the main text, dropping rental properties. See Notes to Table 6 for additional information. * p<0.10, ** p<0.05, *** p<0.01

N Including Both Phases of the Survey

As explained in the paper, given that the second phase of the survey occurred after the third (post-treatment) scheme had begun, there is a concern that the Phase 2 responses may not reflect treatment year behavior. Therefore in the main paper we only include data from Phase 1 for subjective survey outcomes. In this section we report results for both phases pooled and show this does not qualitatively affect our results.

Table 4-N: Impacts on Non-Revenue Outcomes, Both Phases

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Panel A: Main Treatment				
Any treatment	-0.003	-0.008	0.004	0.007
	(0.016)	(0.017)	(0.012)	(0.022)
Panel B: Subtreatments				
Revenue	0.004	-0.009	0.002	-0.022
	(0.023)	(0.025)	(0.017)	(0.029)
Revenue Plus	0.040**	0.039*	0.028*	0.015
	(0.020)	(0.021)	(0.016)	(0.032)
Flexible Bonus	-0.054**	-0.054**	-0.016	0.029
	(0.025)	(0.026)	(0.018)	(0.031)
N	12000	12000	9870	9870
Sample	Full	Full	Full	Full
Mean of control group	0.548	0.568	0.339	-0.103
Rev. vs. Multitasking p.	0.672	0.945	0.813	0.159
Objective vs. Subjective p.	0.004	0.012	0.099	0.315
Equality of Schemes	0.003	0.004	0.090	0.344
Joint significance	0.009	0.012	0.160	0.533

Notes: This table re-estimates Table 4 from the main text, including properties from both phases of the survey. See Notes to Table 4 for additional information. Specification includes a dummy that controls for survey phase (see text for details). * p<0.10, ** p<0.05, *** p<0.01

Table 6-N1: Impacts on Tax Payments and Corruption by Reassessed Status, Both Phases

	(1) Self-reported Tax Payment	(2) Bribe Payment	(3) Frequency of Bribe Payment	(4) Perception of Corruption
Panel A: General Population Sample Only				
Treatment	-62.81	389.6*	.0664	.0023
	(264.7)	(231)	(.0705)	(.0182)
N	11586	11448	9562	12000
Mean of control group	4069.425	1690.712	0.704	0.644
Panel B: Re-assessed and General Population Sample				
Re-assessed * Treatment	1884*	-161.3	-.028	.0063
	(1083)	(278.1)	(.0607)	(.0163)
Re-assessed	2763***	-84.62	.0013	-.0031
	(572.9)	(122.4)	(.0282)	(.0078)
N	16353	15793	13889	16354
Sample	Full	Full	Full	Full
Mean of control group in gen. pop. sample	3928.252	1690.712	0.704	0.644

Notes: This table re-estimates Table 6 from the main text, including properties from both phases of the survey. See Notes to Table 6 for additional information. Specification includes a dummy that controls for survey phase (see text for details). * p<0.10, ** p<0.05, *** p<0.01

Table 6-N2: Impacts on Tax Payments and Corruption by Reassessed Status and Subtreatment, Both Phases

	Self-reported Tax Payment	Bribe Payment	Frequency of Bribe Payment	Perception of Corruption
Panel A: General Population Sample Only				
Revenue	490	755*	.134	-.0262
	(553)	(416)	(.105)	(.0269)
Revenue Plus	-721***	-84.5	-.0545	.00998
	(264)	(273)	(.0977)	(.0242)
Flexible Bonus	16.7	485	.117	.0243
	(250)	(295)	(.104)	(.0245)
N	11586	11448	9562	12000
Mean of control group	4069.4	1690.7	0.7	0.6
Rev. vs. Multitasking p.	0.119	0.184	0.371	0.120
Objective vs. Subjective p.	0.697	0.659	0.488	0.206
Equality of Schemes	0.016	0.112	0.251	0.267
Joint significance	0.018	0.123	0.326	0.439
Panel B: Re-assessed and General Population Sample				
Re-assessed * Revenue	371	49	.0598	.0155
	(1567)	(552)	(.0961)	(.0258)
Re-assessed * Revenue Plus	2195*	-206	-.0648	.0169
	(1270)	(249)	(.087)	(.0221)
Re-assessed * Flexible Bonus	3024*	-320	-.0758	-.0123
	(1594)	(378)	(.0846)	(.0238)
Re-assessed	2762***	-84.4	.00133	-.00304
	(573)	(122)	(.0282)	(.00785)
N	16353	15793	13889	16354
Sample	Full	Full	Full	Full
Mean of control group in gen. pop. sample	3928.3	1690.7	0.7	0.6
Rev. vs. Multitasking p.	0.162	0.572	0.206	0.634
Objective vs. Subjective p.	0.280	0.569	0.437	0.277
Equality of Schemes	0.364	0.831	0.443	0.545
Joint significance	0.150	0.755	0.579	0.705

Notes: This table re-estimates Table 6 from the main text, estimating sub-treatments separately and including properties from both phases of the survey. See Notes to Table 6 for additional information. Specification includes a dummy that controls for survey phase (see text for details). * p<0.10, ** p<0.05, *** p<0.01

Table 7-N1: Impacts on Satisfaction and Accuracy by Reassessed Status, Both Phases

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Re-assessed * Treatment	0.022	0.022	0.001	-0.005
	(0.017)	(0.017)	(0.017)	(0.028)
Re-assessed	0.019**	0.017*	-0.061***	0.122***
	(0.010)	(0.010)	(0.009)	(0.015)
N	16354	16354	14173	14173
Sample	Full	Full	Full	Full
Mean of control group in gen. pop. sample	0.548	0.568	0.339	-0.103

Notes: This table re-estimates Table 7 from the main text, including properties from both phases of the survey. See Notes to Table 7 for additional information. * p<0.10, ** p<0.05, *** p<0.01

Table 7-N2: Impacts on Satisfaction and Accuracy by Reassessed Status and by Subtreatment, Both Phases

	(1) Quality	(2) Satisfaction	(3) Inaccuracy	(4) Tax Gap
Re-assessed * Revenue	-0.002	0.011	0.005	-0.014
	(0.023)	(0.022)	(0.026)	(0.040)
Re-assessed * Revenue Plus	0.034	0.034	-0.026	0.011
	(0.023)	(0.023)	(0.022)	(0.041)
Re-assessed * Flexible Bonus	0.034	0.021	0.023	-0.011
	(0.023)	(0.024)	(0.025)	(0.039)
Re-assessed	0.019**	0.017*	-0.061***	0.122***
	(0.010)	(0.010)	(0.009)	(0.015)
N	16354	16354	14173	14173
Sample	Full	Full	Full	Full
Mean of control group in gen. pop. sample	0.548	0.568	0.339	-0.103

Notes: This table re-estimates Table 7 from the main text, estimating sub-treatments separately and including properties from both phases of the survey. See Notes to Table 7 for additional information. * p<0.10, ** p<0.05, *** p<0.01