

Raising Revenues:

Evidence from Property Tax Assessments in India

Kavya Ravindranath*

October 20, 2026

[*Click here for the latest version*](#)

Abstract

How do property-owners respond to increases in property-tax liabilities? Particularly, do these responses differ across the initial-liability distribution, and how does accounting for this heterogeneity affect the revenue-maximizing tax rate? I construct a novel dataset of geolocated properties and digitised parcel-level maps from an Indian municipality to answer this question. I exploit variation in liabilities that arise from area-based guidance values that are fixed as of 2011, known as ‘circle rates’. First, I find that a 1% increase in the circle rate reduces compliance by 0.45% but increases total revenues by 0.85%. This implies that revenues are maximized by scaling every circle rate by a common factor of 4.011. This preserves the existing relative circle rate schedule while increasing liabilities proportionally across properties, and is equivalent to increasing the statutory tax rate from 0.25% to approximately 1.00%. Second, I find statistically significant differences in compliance elasticities across the circle rate distribution, with larger responses at higher rates. Incorporating this reduces the revenue-maximizing scaling factor to 2.217, equivalent to increasing the statutory tax rate to 0.55%. Third, I find that updating liabilities to reflect revenue maximising schedules leads to large gains in revenue, but simply updating outdated circle rates to more recent values recovers a substantial share of these gains.

*George Washington University. I am grateful to my advisors Elira Kuka, Tanner Regan, Paul Carrillo and Remi Jedwab for their constant guidance and support. I also thank Nick Li, Amory Gethin, Pierre Bachas, Neha Betai and Kathryn Nicholson for their valuable comments and suggestions. I am grateful to Yashpal Singh Rajput for his help with data access. All errors are my own.

Property tax revenue as a percentage of GDP is around 0.6% in LMICs, four times lower than the OECD average of 2.5% (Bahl and Martinez-Vazquez, 2007). Among other reasons, low revenues may reflect low tax liabilities (Awasthi et al., 2021). Raising liabilities in low-enforcement environments however, may come with increased non-compliance¹ (Slemrod, 2019). More importantly, this response need not be the same throughout the tax base. If taxpayers facing different initial liabilities respond differently to an increase in taxes, a single average elasticity may mask variation in the revenue consequences of raising liabilities.

This heterogeneity matters for policy. Existing work on income taxation shows both that behavioral elasticities can differ substantially across taxpayers (Gruber and Saez, 2002) and that pooled elasticity estimates combine underlying heterogeneous responses (Kumar and Liang, 2020). Accounting for this heterogeneity has been shown to matter for revenue-maximizing income-tax rates (Holter et al., 2019; Guner et al., 2016) and corporate-tax rates (Serrato and Zidar, 2018). I bring this question to the property tax literature. How much heterogeneity exists in responses to increases in property tax liabilities and how does accounting for it change the amount by which a municipality can raise liabilities before additional increases begin to reduce revenue? I answer this question using variation in *property assessments*. In the municipality I study, the statutory tax rate is common, while assessed values depend on administratively assigned, area-based guidance values known as “circle rates”. A circle rate is a dollar amount per square metre assigned to properties within a given area, or “circle”. Circle rates used for property tax assessments have remained fixed since 2011. This system of assessments generates spatial variation in liabilities at many points of the existing tax base. I use this variation to estimate how compliance and the amount paid, conditional on payment², respond to higher assessed values, and how these responses differ across the circle-rate distribution.

To study how heterogeneity affects these revenue-maximizing updates, I set up a simple conceptual framework that defines total revenues as average compliance times the average amount paid by compliers times the number of properties. Assuming that the number of properties is fixed, I model compliance and log payments among compliers as functions of log circle rates. I first estimate pooled behavioral responses and then allow them to vary across the circle-rate distribution. Using a

¹Non-compliance in this case refers to complete non-payment of the tax. There could be other behavioral responses to consider, typically characterised as changes in the tax base or avoidance (Saez et al., 2012). This includes actions such as filing for property re-assessments, changes in the decision to own property, moving to other low-liability areas or constructing properties in low-liability areas. I don’t find evidence of these behaviors, making the decision to just not pay the tax the most important response.

²This estimate is important as households are not required to pay the full liability and often have penalties imposed due to non-compliance in a previous year, leading to payments much larger than their liability. Penalties are 2% of the liability, so larger liabilities lead to larger penalties.

model-based extrapolation, for both sets of responses, I estimate a common factor by which all existing circle rates can be scaled to maximize revenue. This proportional scaling preserves the relative circle-rate schedule and is therefore equivalent to changing the statutory tax rate uniformly across properties while holding assessed values fixed. Finally, I use these elasticities to evaluate alternative ways of updating the assessment schedule. This includes comparing them with a more conventional policy of updating outdated circle rates to their values in a more recent year. I study these questions in Kolhapur (population 800,000), India, where assessed values for residential, owner-occupied properties have remained unchanged since 2011.

I combine three new sources of data for this analysis. First, I collect administrative property-level data on tax liabilities, payments and demographic information across 12 years. I translate all names and addresses from Marathi to English and mark properties as being residential and owner-occupied. Second, I geolocate a sample of 6,345 properties using data from an official Kolhapur Municipal Corporation (KMC) geotagging exercise. Third, I georeference and digitize scanned parcel and “circle”-level maps of the city and link them to “circle rate” data. Together, these data allow me to observe each property’s tax liability, payment behavior and granular location. I restrict the main sample to residential, owner-occupied properties observed in a balanced panel from 2013–2024, to exclude new properties entering the tax roll in later years.

To causally measure elasticities, I make use of variation that arises from the area-based property valuation formula. Properties are assessed using a hedonic function with a fixed set of weights for each property characteristic and the 2011 circle rate. The tax liability is then 0.25% of the property valuation. The circle rate is a floor price meant to guide the reporting of property sale prices ([Anagol et al., 2024](#))³ and the KMC pegs it at its 2011 value in the property-valuation formula⁴. This has meant that liabilities have remained unchanged since 2011. Since circle boundaries have stayed constant over the years and do not follow local administrative divisions, I define the 2011 circle rate as the baseline circle rate in my main empirical specification.

Specifically, two *similar* properties located sufficiently close to each other can lie on opposite sides of a circle boundary, be assigned different circle rates and have different liabilities. I match properties using a nearest-neighbour matching algorithm and define property-pairs using a threshold of 16 metres⁵. In my main specification, I then compare properties using a

³A circle typically comprises a few parcels and could be as big as 1–2 blocks. The word ‘circle’ is a bit of a misnomer as these areas are not actually shaped as circles.

⁴Circle rates are, however, set and updated by an external state-level commission every two years.

⁵This narrow distance helps compare near-neighbours on the same side of a road. This allows me to hold fixed

within-property-pair-by-year fixed effect with the logarithm of the 2011 circle rate as the independent variable⁶. This allows me to estimate differences in the outcome within a pair of properties for the same year, due to a 1% increase in the circle rate.

The main identifying assumption is then that within a property-pair-year, differences in baseline circle rates are unrelated to unobserved determinants of compliance and payments. I cannot directly test this assumption, but I can check if there are differences in *observable* property characteristics within a matched property-pair. Since liabilities are a hedonic function of the circle rate and observed property characteristics, I construct a “residual” term by residualizing liabilities with respect to the circle rate. This term captures characteristics such as the size of the property, the height and age of the building and any amenities located within an apartment complex. I show that properties within a pair are balanced on this residual term. I also show that properties are balanced on observable characteristics such as the number of floors, the type of construction and the area in square metres of the property, variables I obtain from the Google Footprints database.

I first estimate a pooled response, where I define compliance as an indicator equal to one if any payment is made in a given year. I find that a 1% increase in the baseline circle rate reduces compliance by 0.32 percentage points, implying an elasticity of -0.45⁷. This is then an extensive-margin response to higher liabilities.

Is this fall in compliance offset by higher payments among those who do pay? On the intensive margin using only the sample of complier property-pairs, payments rise substantially. A 1% increase in the circle rate increases taxes paid by 1.30%. Based on the model, the sum of the two elasticities should equal 0 at the revenue-maximizing point. Combining the two responses, I get a sum of 0.85, suggesting that the average property lies on the rising portion of the revenue function and that a 1% rise in the circle rate leads to a 0.85% increase in revenues.

Since I restrict the intensive margin to complier-property-pairs, a concern may be that this effect is driven by property-pairs differentially selecting into the sample. In a robustness check, I relax the fixed-effect structure so that pairs with only one payer also contribute to the estimate, and find a similar elasticity of 1.20. I also use a PPML specification⁸ and find a statistically similar but imprecisely estimated total revenue elasticity of 0.55. Together, this suggests that selection into characteristics shared by properties that may drive compliance, such as access to public services, quality of roads, and other neighbourhood amenities.

⁶This design closely follows Dube et al. (2010)’s design comparing county-pairs with differences in minimum wages.

⁷I show in a robustness check that this estimate is very close to estimates I get from a classic Spatial Regression Discontinuity Design, following Best et al. (2025).

⁸Since this sample now includes lots of 0s from non-compliant properties, I don’t use a logged dependent variable.

payment marginally affects the magnitude of the intensive-margin estimate, but does not overturn the conclusion that higher circle rates increase total revenues.

Extrapolating these elasticities⁹, I find that the pooled RMCR is 858 USD. This amounts to all circle rates being scaled by a factor of 4.011, or an increase of about 301%. This is equivalent to the statutory tax rate increasing to approximately 1.00% from 0.25%, while keeping assessments unchanged¹⁰.

Any enforcement activity could be underestimating the baseline compliance effect. I next make use of a policy rule, according to which the KMC sends reminder letters and threatens enforcement action once accrued liabilities cross a threshold of 117 USD. I restrict my baseline sample to properties below this threshold and reassuringly find that these estimates are not statistically different from those estimated at baseline¹¹. This enforcement by the KMC is very light-touch and my elasticity estimates should be interpreted as those in settings with weak enforcement (Bachas and Soto, 2021; Best et al., 2015; Emran and Stiglitz, 2005).

I next proxy for house prices using contemporaneous circle rates and find that my results are not driven by differential growth in house prices. I also proxy for construction using new property entries to the tax roll and find that an increase in the circle rate leads to more houses being constructed on average within that circle, suggesting the opposite of avoidance behavior. I also find a substantially large negative compliance response during 2020 that returns toward its earlier level afterwards, consistent with tighter income constraints during the lockdown (Brockmeyer et al., 2021).

I now ease the assumption that a single elasticity describes the whole city. Following the continuous interaction specification in Allcott et al. (2020), I allow responses to vary with the local within-pair mean log circle rate, and estimate a response function. I summarize these responses in terciles of the pair-level distribution¹². Compliance elasticities differ significantly across terciles (joint $p < 0.001$), with point estimates of -0.077 (imprecisely estimated) in the low tercile, -0.413 in the middle and -0.667 in the high tercile¹³. These elasticities measure proportional changes in compliance and reflect estimated response slopes as well as baseline compliance levels. Conditional-payment

⁹And assuming compliance changes linearly

¹⁰Although this is approximately a fourfold increase, the new rate is not unusually high by international standards. The IMF reports that rates in market-value-based property tax systems typically range from 0.5–2%, the average effective tax rate on a median-valued owner-occupied home across the largest U.S. cities was 1.21% in 2025, and São Paulo and Rio de Janeiro levy residential property taxes at approximately 1% of assessed property value

¹¹The KMC only does enforcement action via tax officers when total accrued liabilities cross a threshold of 10,000 USD, which is so high that very few residential properties are threatened with it. This is reassuring in my setting as it is not plausible that enforcement officers choosing to enforce in higher circle-rate areas biases my results

¹²I count each pair once when setting the cutpoints

¹³I confirm balance on observables using this specification

elasticities, though increasing across terciles, are not statistically distinguishable from each other (joint $p=0.345$). The main evidence of heterogeneity thus comes from the extensive margin compliance elasticities.

I use the two elasticities and their respective response functions in a model-based numerical exercise to find the common scaling factor that maximizes aggregate revenue. For each baseline circle-rate tercile, I use mean compliance and predict compliance as rates rise, using the fixed estimated elasticity of that tercile. I then search for a scaling factor that maximises aggregate revenues¹⁴. I find that this RMCR is now lower than the one suggested by the pooled exercise, at 474 USD, corresponding to a common scaling factor of 2.217, or an increase of about 122%. This is equivalent to a statutory tax rate of approximately 0.55%. In a separate exercise, I measure separate revenue-maximizing rates for each circle rate, using these elasticities. In line with [Gruber and Saez \(2002\)](#), I find that this results in a declining tax-rate schedule.

Is the heterogeneous-response scaling factor significantly different than the pooled scaling factor of 4.011? To check, I quantify the uncertainty in the two estimates using a clustered bootstrap¹⁵. The 95% percentile intervals for the factors are [1.636, 3.118] under heterogeneous responses and [2.615, 7.354] under pooled responses. Although these intervals overlap, comparing two individual confidence intervals is not a test of equality between the estimates. The relevant question is whether their difference is statistically distinguishable from zero. I therefore calculate the pooled-minus-heterogeneous difference within each bootstrap, accounting for the covariance between the two estimates. This difference is 1.794, with a 95% percentile interval of [0.316, 5.193]. Since this interval excludes zero, accounting for heterogeneity significantly lowers the scaling factor at the 5% level.

Finally, I compare revenue changes under alternative assessment updates using a separate first-order approximation. Alongside the two model-implied revenue-maximizing schedules, I consider updating outdated circle rates to a more recent base year. Scaling by the heterogeneous factor of 2.217 gives an approximate revenue increase of 33% in this separate exercise, compared with about 100% under the pooled schedule. Updating the base year produces gains of a similar order to the heterogeneous schedule. These approximations are distinct from the exact predicted-revenue objective used to select the constrained RMCR.

¹⁴In practice to do this, I have to extrapolate predicted compliance over the CR distribution support, and assume that elasticities can be linearly extrapolated

¹⁵I resample property pairs and estimate both rates on the same resampled data

This paper makes three contributions. First, it contributes to the literature on revenue-maximizing tax rates. A large literature shows that revenue-maximizing tax rates depend on taxpayers' behavioral responses to taxation (Saez, 2001; Badel and Huggett, 2017; Saez et al., 2012). However, aggregate elasticity estimates mask heterogeneity in underlying responses (Kumar and Liang, 2020). This heterogeneity matters for revenue-maximizing taxation. Holter et al. (2019) and Guner et al. (2016) use structural models to show how household heterogeneity affects income-tax Laffer curves, while Serrato and Zidar (2018) show that the structure of the corporate-tax base affects revenue responses and increases revenue-maximizing rates. I contribute by considering heterogeneous behavioral responses in revenue-maximizing taxation in property taxes. I show that accounting for this heterogeneity substantially lowers the revenue-maximizing common scaling factor. This exercise holds the relative assessment schedule fixed, so the change can also be expressed as a lower revenue-maximizing statutory property-tax rate.

Second, it contributes to the literature on property-tax assessment systems. I study an area-based assessment system, while most of the literature focuses on individual property-assessments. For instance, in a very similar setting of a city in India, Duflo et al. (2026) find that updating individual property tax assessments leads to a large 30–60% increase in revenues. Avenancio-León and Howard (2022) and Hou et al. (2025) show that imperfect and infrequent property assessments lead to regressivity. Knebelmann et al. (2026) show that algorithmic valuation improves assessment accuracy and equity in Senegal.

Third, it contributes to the literature on tax compliance and fiscal capacity in developing countries. Ajzenman et al. (2024) and Tourek et al. (2026) find that increasing the rate for richer properties and reducing the rate for poorer properties leads to increases in revenues. Existing work studies how enforcement, information, technology, social norms, tax rates, lottery-based rewards, collector incentives and progressivity affect compliance (Bergeron et al., 2024; Brockmeyer et al., 2021; Knebelmann et al., 2026; Best et al., 2025; Manwaring and Regan, 2024; Okunogbe and Tourek, 2024; Ferraz et al., 2024; Bergeron et al., 2024; Brockmeyer et al., 2021; Carrillo et al., 2021; Balán et al., 2022; Khan et al., 2016; Aparicio et al., 2011; Pomeranz, 2015; Ajzenman et al., 2024; Tourek et al., 2026). I contribute by showing that compliance significantly fell during COVID-19.

I begin by describing the institutional setting in Section 1, describe my data in Section 2, and go over the conceptual framework, main results and heterogeneity in Sections 3, 4 and 5. In Section 8, I conclude.

1 Urban Municipal Finances

In this section, I first describe the increasing importance of property taxes in India. Next, I describe the features of the property tax in Kolhapur, focusing on the ones I exploit to measure elasticities.

1.1 Recent developments

There are two broad sources of revenues for urban municipalities in India: own-source and government transfers. Historically, property taxes have formed about 30% of own-source revenue and 50% of own-tax-revenue, with some variation across cities (Awasthi et al., 2021). However, post-2016, after the passage of a Goods and Services Tax law that centralized sales tax collection, other sources of own-tax-revenue have significantly decreased. This has increased municipalities' reliance on property taxes and has brought local tax collection to the forefront as a policy issue¹⁶.

This increased reliance has come along with reforms in property taxes including digitization of valuation exercises, digitizing village-level tax collection records, harmonizing tax bands, etc¹⁷. The paucity of easily-available data has however made studying the effects of these reforms difficult. In 2011, the Government of Maharashtra passed a Maharashtra Municipal Property Tax Board Act and *suggested* that all municipalities move away from the Rental Value Assessment¹⁸ of properties and into the Capital Value Assessment method. The state also prescribed a partly hedonic formula with assigned weights to be used when assessing properties. Only two municipalities immediately took up this suggestion and made it law; Kolhapur was one of them. Thus, starting in 2011, Kolhapur moved to assessing *owner-occupied properties* using the Capital Value Assessment Method¹⁹.

1.2 Property Tax in Kolhapur: Valuation of Owner-Occupied Properties

On switching to the Capital Value Assessment method, each property 'a' in structure 'b' in year 't' was assessed based on the following formula:

$$\text{Valuation}_{a,b,t} = \text{Circle Rate}_{2011,c} \times \theta_{\text{Category},b} \times \kappa_{\text{Type},b} \times \mu_{\text{Age},b} \times \lambda_{\text{Height},b} \times \text{Square feet}_a \quad (1)$$

θ, κ, μ and λ are all weights assigned based on specific rules that I show in Table A1. Notice that all

¹⁶See: <https://fincomindia.nic.in/asset/doc/commission-reports/15th-FC/reports/studies/State%20of%20Municipal%20Finances%20in%20India.pdf>.

¹⁷This includes policies such as SVAMITVA, for instance: <https://svamitva.nic.in/>

¹⁸The Rental Value Assessment method is resource-intensive and based on valuator discretion, as officials must estimate the actual or notional rent for all properties.

¹⁹Rental properties switched to using this method only in 2019

weights are common for all properties in structure ‘b’. ‘Square feet’ refers to the area of the property ‘a’ and is the only feature that is property-specific, and lastly, ‘Circle Rate’ is a \$ per square metre amount set by the state government and assigned to all properties in a ‘circle’ or a ‘survey area’.

Circle Rates serve as a proxy for market property values and though they are likely much lower (Awasthi et al., 2021) they closely follow house price trends and have been trending closer to actual house prices over the past decade (Anagol et al., 2024). I present an excerpt of the document assigning circle rates to survey areas in Panel C and snapshots of a few survey areas in Panels A and B of Figure A1. Circle rate assignments are made by a state-level commission based on the demand and supply of properties in an area and are updated every other year ²⁰.

However, circle rates in the property tax formula in Equation 1 are fixed at their 2011 value. Combined with the fact that properties were assessed for their characteristics last in 2007, the valuation of all owner-occupied properties has remained *unchanged* since 2011. The statutory property tax rate is then 0.25% of $\text{Valuation}_{a,b,t}$. This rate is constant across all properties in the municipality, though as valuation amounts differ, the liability or in other words, the effective tax rate for each property, differs.

1.3 Details on Circles or Survey Areas

Since I use the difference between two circle rates as my main identifying variation, it is important to note that circles or survey areas are not administrative divisions and not explicitly drawn for purposes of the property tax. They often criss-cross over roads and are fairly small, often the equivalent of one block in a standard city in the United States. In Figure A2, Panel A, I present a sample of circles superimposed on a zoomed-in Google Maps screenshot of the city. Outlines in red represent circles as drawn by the state commission. Notice that while circles adhere to the shape of large roads, they cross over smaller roads. This means that there is likely no difference in local government services or political institutions across neighbouring circles.

In Figure 1, Panel A, I present a shaded map of circles within the city. Darker colours present areas with a higher circle rate in 2011 and lighter colours are areas with a lower circle rate. I present a zoomed-in version of this map focusing on a portion of one ward in the city in Panel B, showing both the compact nature of the circles and the jumps in rates. As rates are set based on the demand and

²⁰Their purpose is to serve as a floor for the rates at which properties in the circle must be sold and are reported for the purposes of a separate sales transaction tax. Anagol et al. (2024), for instance, use these rates to check for bunching around transaction prices.

supply of properties within a circle, it is likely that even within neighbouring circles, the one with the higher rate has more construction. As new properties get constructed, they get added to the roll and get freshly assessed. I constrain my main results to a balanced panel of properties to avoid any bias from newly constructed properties.

It is also possible that even between neighbouring circles, properties in the high-rate circle are different from properties in a lower-rate circle. However, at the border of two circles, this is less likely; houses at circle-borders that neighbour each other are likely to be observationally similar.

1.4 Building Codes and Property-Width Regulations

These distances may seem very small but under building codes in India and the state of Maharashtra, they are not implausible. Building code regulations allow for no gaps to exist between two adjacent plots up to an area of 450 metres² ²¹. Average plot widths in mid-sized urban residential areas are about 6 metres, making it possible (and likely) that adjacent or near-adjacent plots have a distance of 8 metres between their centroids. Since the same building code also specifies that roads must be around 9–12 metres wide, it is also likely that the coordinates I compare are on the same side of the road. I flesh out details on my distance thresholds, the main specification I use and the identifying assumptions that need to be satisfied for causality in Section 2.

1.5 Tax Bills and Enforcement

Tax Bills are generated on the 1st of April for the whole city, with a deadline to pay by the 30th of June, post which monthly penalties of 2% are levied. Before 2015, penalties were levied annually at 18% instead of accruing monthly. Tax bills are generated digitally and are physically delivered by post, while payments can be made online through a payment portal. Property tax debts need to be cleared in order to transfer property deeds; however, enforcement action outside of this requirement primarily happens through letters.

Property owners with a cumulative accrued amount of over 10,000 rupees (around 100 USD), including penalties and unpaid bills, are sent enforcement letters reminding them to make a payment. Once this accrued amount crosses 1,000,000 rupees (around 10,000 USD), enforcement officers make visits and threaten to lock up the premises of non-payers. This second threshold is high enough that most properties that face this threat are large commercial properties. The first threshold allows me to

²¹See page 221: https://www.mmrda.maharashtra.gov.in/sites/default/files/2023-10/UDCPR_compriessed_2.pdf

separately check the effect of enforcement actions once the threshold of 100 USD is crossed. However, outside of this, there does not seem to be any evidence of enforcement officers making visits and asking people to pay, as is seen in other low and middle income countries.

In the next section, I describe the various datasets used in this paper and the construction of all the primary variables used for analysis. I also describe the process of collecting and constructing the geographic variables necessary for my baseline spatial specification.

2 Data

I use three different sources of data to construct the main dataset with property-level variables: (i) a property-level tax bills and payments history dataset, (ii) property-level coordinates data for a subset of properties and (iii) a Google Footprints database with property-level characteristics. I further augment these with a dataset of circle rates and circle-level maps, digitized from scanned documents.

2.1 Property Tax Bills and Payment History

I start by collecting property-level bills and payment data from the Kolhapur Property Tax defaulters [website](#), the purpose of which was to publicize names of defaulters. I then collect information from the Kolhapur Municipal Corporation's (KMC) property tax payment [website](#) once in December 2024 and use a different [website](#) to re-collect data on missing bills in April 2026 ²². On all these sites, I sequentially search for property numbers starting from 100 and collect all the information displayed. This gives me detailed demographic information such as the names of the property assessee and the occupier of the property, the full address, the circle number, ward number and division number; as well as information on bills and payments.

2.1.1 Tax Bills

Using the first source on defaulters, I obtain a unique property assessment number for each property, the ward it belongs to and information on tax bills such as the liability, the penalty accrued due to non-payment, the amount of the municipal tax owed and any discount offered on the bill by the municipality. I collect this information for all available years, though it starts getting well-populated only after 2011. This source for defaulters however, only contains tax bills for *never-payer* properties.

For those who always pay the tax (always-payers) or pay intermittently (ever-payers) the tax bill

²²The two sites I collected data from originally in December 2024 are not hosted by the KMC anymore. However, since the tax year had not ended in 2024 when I first collected the data, payments are incomplete for Financial Year 2025, which allows me to retain my main sample only until FY 2024.

on the website disappears once the assessee pays the bill. This means that I don't observe any tax bills for always-payers and for ever-payers that paid their bill in FY 2025 ²³.

I do however see the original bill for never-payers and for ever-payers who hadn't paid in FY 2025 ²⁴. I start with populating data for these two groups and out of 75,473 owner-occupied properties in the sample, this step gives me data for 23,069 properties ²⁵.

Next, I use the tax bills data I re-collected in April 2026. New bills for properties are generated on the 1st of April, and re-collecting it immediately ensures that I have data on bills for properties *before* they are able to make a payment. This step gives me a large portion of the remaining data on liabilities and I am able to populate bills for 43,559 properties.

I am now left with 8845 owner-occupied properties with no data on bills, likely always- and ever-payers who paid very early in April 2026. For these properties, I simulate their bills using their payments and a few assumptions ²⁶, however, I note that none of these properties with simulated bills are retained in my final geo-located sample.

2.1.2 Tax Payments

I collect information on the amount of the tax paid, the mode of payment, the exact date of payment and the receipt number.

2.1.3 Processing demographic information

Since all names and addresses are stored in Marathi using the Devanagari script, I first translate everything from Marathi to English using a Google Translate API. I then use the names of the owner or the assessee, and the name of the occupier to construct a variable marking properties as being owner-occupied. I also use assessee names along with regular expressions to classify properties as being 'commercial' or 'institutionally-owned' properties ²⁷. I make these distinctions as (i) rental properties underwent a valuation change in 2019 before which their valuations were not dependent on the circle rate and are not constant over the past 15 years; (ii) commercial properties use a

²³Even if they didn't pay a bill in one intermittent year, but paid in the next year, the bill for the previous year disappears

²⁴I see bills for never-payers for every single year, obtained from the defaulter list. For ever-payers, if I see a bill in FY2025, I am able to assume that the bill was the same for all years prior as liabilities have remained unchanged.

²⁵Properties can ask to be re-evaluated, however, there are only about 1500 properties that have ever asked for re-evaluation. I am not able to see which properties these are, unless they are never-payers

²⁶I use the rule that until 2015, penalties only accrued yearly on the entire amount defaulted. This means that the payment on the bill could be made at any point during the year and with no penalties. I now assign bills in two steps: (i) I tag all payments made per property before 2015 and (ii) assign them as a bill if payments were made for the same exact amount for 2 consecutive years. Using these assumptions for properties for which I have the actual tax bill, I find that I predict the bill accurately 77% of the time. This last step allows me to populate a tax bill for all properties in my sample.

²⁷For instance, all properties ending with a 'Ltd', 'Limited', 'Corporation', 'Corpn' etc. are marked as commercial. All schools, trusts, religious places are marked as institutional

different circle rate, typically one meant for commercial spaces and since this rate is much higher, using the same rate as in the residential one may generate bias and (iii) institutional properties are those that are owned by the Kolhapur Municipal Corporation or other governments, these properties may have different incentives to comply or evade than purely residential properties.

2.2 Coordinate-level data

I obtain and construct two sets of coordinate-level data for a smaller sample of my data, (i) a Google Footprints database and (ii) a geolocated property-database. I then merge these two datasets so that each property in the final sample has a set of coordinates and a set of covariates attached to it.

2.2.1 Geolocating properties

I construct the geolocated property database using three sources. First, I obtain a sample of coordinates from an official KMC source. The KMC developed an app in 2026 which allowed the official property tax bill delivery officers to geotag each property they visited. I obtain coordinates for 4141 properties in my sample through this source. Since coordinates were located at the doorstep of each property, these are very accurate.

Second, I use the Maps Places API and find addresses for coordinates from the Google Footprints dataset. The Places API gives me an exact address of the coordinate, as well as a name of the building. I then manually match these addresses with the addresses from the property-tax dataset and am able to find coordinates for 670 properties this way. While the places API is useful, due to the poor quality of addresses in India, there are a lot of addresses with no matches and it is often only able to find names of the closest apartment building.

I then move on to the third method of collecting coordinates where I manually find coordinates for addresses in my data using Google Maps²⁸. In Figure A2 Panel B, I provide an example of a property on Google Maps that I manually geolocate. This gives me an additional 1534 properties and a total of 6345 properties with coordinates. I present all coordinates in my geolocated sample in Figure A4.

All coordinates in my data have a minimum precision of six decimal points and a maximum precision of 11 decimal points. To construct pairs, I round up all coordinates to 6 decimal points.

²⁸This method is very flexible and is particularly useful since several addresses in the data are structured in a non-standard way; some are specific and state the name of the building, the plot number or a colony number mixed with a plot number; some are very vague, with just the name of a street or a landmark. I am unable to geolocate properties with these very vague addresses but am able to use plot numbers and other details to get coordinates

2.2.2 Google Footprints data

The Google Footprints database for Kolhapur gives me useful information on the structures located on each coordinate. This includes the type of the structure - residential, commercial, religious, etc.; the number of floors; the area of the structure; the name of the road it is located on; the width of the road it is located on; the ward and the name of the area. I present a snapshot of the shapefile I obtain in Panel A, Figure A6 and a zoomed-in snapshot in Panel B.

I then use a nearest-neighbour matching exercise between this data and the set of geolocated properties and retain characteristics only for those properties that find a match at or below 0.0005 miles.

2.3 Digitizing Circle-level shapefiles

Since there are no readily available shape files for the city at the circle-level, I first collect scanned maps of the city that mark circle areas using the city's 'Development Plan'. In Panels A and B of Figure A1, I present snapshots of one of the PDF maps I digitize.

I collect 18 maps from the KMC website, all of which join to make the whole city. Nine of these are for denser areas as dense areas tend to have closer circles. All circle areas are marked by dashed red lines. I first georeference all maps using control points on QGIS so as to be able to use their coordinate-level information. I then programmatically extract the red boundary linework from the PDFs and reconstruct circle polygons from these dashed outlines. Because the original boundaries are dashed rather than continuous, I use automated gap-closing steps to connect nearby dash segments before forming polygons. I separately extract the printed survey numbers from the PDF text layer, clean and standardize these labels, and spatially assign them to the corresponding polygons. While this step allows me to make maps as in Figure 1, there are still missing and mislabeled parcels. Due to this issue, I don't rely on these maps for my main analysis, but use them for exploratory purposes.

2.4 Digitizing Circle Rates

The other crucial piece of data I need is the circle-rates for 2011. I obtain hard copies of this data from the Stamps and Registration Department in Kolhapur and digitize it. I present a snapshot of the pdf I digitize in Panel C Appendix Figure A1. Each circle has a different rate for residential constructions, open land, commercial constructions and office spaces. I only extract the rate meant for residential purposes as that is the rate used in the property tax formula and thus the rate relevant to me. I

also collect this data from 2015 to 2024 as it is available on the Maharashtra Registration & Stamps Department's eASR portal. While I only use the 2011 rate for identification, the longer panel of the change in rates allows me to test for heterogeneity in outcomes between areas with higher and lower circle rate growth rates.

Circle rates are set by an external committee and the formula used to determine and update rates is not publicly available. There could then be a concern that boundaries are drawn and rates are set based on area characteristics. While my design of comparing houses located very close to each other should take care of this concern, in Section [A](#), I conduct an exercise to check what predicts the circle rate. I find that no concerning factors emerge.

2.5 Final Sample

Out of the 75,473 owner-occupied properties in my initial sample, 38812 properties have a balanced panel and are not commercial or institutional properties. Out of these, I obtain coordinates for 6345 properties.

The highest number of properties in the full balanced panel of properties lie in the middle, with an average circle rate of 183 USD. In Figure [A5](#), Panel A, I present the distribution of the number of properties across circle rate bins. In Panel B, I present the corresponding amount of revenue collected for each circle rate bin and the potential revenue that could be collected if compliance was full. The largest amount of potential revenue to be collected also lies in the middle of the circle rate distribution, though there is a large amount of uncollected revenue even at the highest bin. In total, just from the balanced panel, there is 8 mn USD in potential uncollected revenue. In Section [A.1](#), I plot descriptive relationships between key variables.

I also note that since the property tax formula uses rates from 2011, there is a large growth in circle rates over the past 15 years. In Figure [A9](#), I present the distribution of circle rates in 2011 in the city in Panel A and the distributions for 2017, 2019, 2021 and 2023 in Panel B. The median circle rate, which is about 180 USD in 2011, is nearly 450 USD in 2023. This points at the likely large under-estimation of property taxes due to using rates from 16 years ago.

3 Conceptual Framework and Empirical Strategy

From the perspective of a policy maker, there are two things she could care about: compliance and the revenue collected ²⁹. The final revenue collected is then dependent on the compliance *and* the amount they pay while complying.

The amount by which compliance and taxes paid change with circle rates then determines the revenue function for this city. If it is necessary that taxpayers pay their whole liability every time, the amount of tax paid would be redundant as a singular measure for compliance would capture them both. However, assesseees need not pay the entire amount while complying. In addition, since penalties increase with liabilities, taxpayers who comply with a delay often pay an amount much higher than their liability.

I start by modifying the framework in Bergeron et al. (2024) ³⁰ and add an additional term for the amount of the tax paid by compliers. The revenue collected at a circle rate 'cr', is then defined by

$$R(cr) = N(cr) \times C(cr) \times A(cr)^+ \quad (2)$$

where $A(cr)^+$ is the average dollar amount paid by compliers and $C(cr)$ is the average compliance ranging between 1 and 0 at circle rate 'cr' and $N(cr)$ is the number of properties in a circle. Using the standard formula for elasticity, the elasticity of compliance is then defined as $\epsilon_{compliance} = \frac{\Delta C(cr)}{C(cr)} \times \frac{cr}{\Delta cr}$ or $\frac{dC(cr)}{dcr} \frac{cr}{C(cr)}$ and the elasticity of payments by compliers as $\epsilon_{payments} = \frac{\Delta A^+(cr)}{A^+(cr)} \times \frac{cr}{\Delta cr}$ or $\frac{dA^+(cr)}{dcr} \frac{cr}{A^+(cr)}$, where 'cr' is the dollar amount of the circle rate.

To obtain the circle rate at which revenue is the highest, I differentiate Equation 2 with respect to 'cr',

$$\frac{dR(cr)}{dcr} = N(cr)[C'(cr) \times A(cr)^+ + C(cr) \times A'(cr)^+] \quad (3)$$

Holding $N(cr)$ fixed, factoring out $\frac{C(cr)A^+(cr)}{cr}$ and substituting in the terms for elasticity, we get:

$$\frac{dR(cr)}{dcr} = [\epsilon_{compliance} + \epsilon_{payments}] \frac{C(cr)A(cr)^+}{cr} \quad (4)$$

Using the First Order Condition, revenue is maximized when $\frac{dR(cr)}{dcr} = 0$. Thus, at the circle rate

²⁹She could care about compliance as a standalone parameter, outside of total revenues collected, in order to address concerns about horizontal equity

³⁰In their context, liabilities are fixed and compliers have to pay the full amount. Revenues are thus only a function of compliance and revenues are maximized when the elasticity of compliance equals -1

at which revenue is the highest, it must be that:

$$\epsilon_{compliance} + \epsilon_{payments} = 0 \quad (5)$$

Then, revenue is increasing when $\frac{dR(cr)}{dcr} > 0$ or when $\epsilon_{payments} > |\epsilon_{compliance}|$ ³¹.

3.1 Obtaining Elasticities

In a standard setting, to measure the elasticity of compliance to liabilities, assuming a level-log model, we can estimate the following equation:

$$\text{Compliance}_{it} = \alpha_0 + \alpha_1 \text{Log}(\text{Tax Liability})_i + \tau_t + \varepsilon_{it} \quad (6)$$

Where the elasticity at the mean for compliance is then given by:

$$\text{Elasticity} = \frac{\hat{\alpha}_1}{\text{Compliance}} \quad (7)$$

And the elasticity of taxes paid can be estimated using:

$$\text{Log}(\text{Tax paid})_{it} = \lambda_0 + \lambda_1 \text{Log}(\text{Tax Liability})_i + \tau_t + \varepsilon_{it} \quad \text{for Compliance}_{it} = 1 \quad (8)$$

Where the elasticity is given by λ_1 since Eq 8 is a log-log regression.

However, there could be several omitted variables leading to bias in α_1 and λ_1 . Properties with high tax liabilities could be located in areas with better public services, causing property-owners to comply. It could also be possible that property-owners with a high moral incentive or ability to pay the tax choose to locate in larger houses, leading to an upward bias in the estimate. In Figure A7, Panel D, I show that this hypothesis is likely true. It could also be possible that property-owners with means to evade the tax could be locating in larger houses, leading to a downward bias.

3.2 Causally identifying elasticities

To deal with this issue of endogeneity, I make use of the property tax valuation formula (Equation 1) and compare two properties that are located very close to each other but fall in different circles with differing circle rates in 2011. Properties that are located very close to each other have the same access

³¹This result is in principle similar to the one shown by Saez (2001), where the revenue effect is captured by summing the average elasticities derived from the behavioral and mechanical effects.

to public services, and likely look similar to each other.

The identifying assumption is that *within* a pair of neighbouring properties, the assignment to either side of the circle border is random, thus randomly generating a difference in their compliance.

I restrict my baseline results to properties located very close to each other and that have a balanced panel. Formally, for a property ‘i’ in a circle with a circle rate ‘cr’ and a closely located property ‘i’ with a circle rate ‘cr’ :

$$\text{Tax Liability}_{i,i'} = \text{Circle rate}_{cr,cr'} \times v_{i,i'} \quad (9)$$

where $v_{i,i'}$ represents all the other weights on hedonic property characteristics used to calculate the tax liability as specified in Equation 1.

The identifying assumption here is then that for properties where $v_i = v_{i'}$, all differences in the liability are captured by the difference in circle rates. In my baseline results, I construct pairs of properties that lie within 16 metres of each other in bordering 2011 circle rate-varying circles, and test if this is true.

To do this, I match all *coordinates* in my data using a nearest-neighbour matching algorithm. This results in distances ranging from 0.002 miles to 0.2 miles as it could be possible that the nearest-neighbour for a property is not a neighbouring one. However, I present my baseline results using 16 metres and show robustness to using 8 metres and 0.05 miles. I then construct pairs of properties between matched coordinates. Appendix Figure A10 illustrates this process in detail. I match coordinates A and B if they meet two conditions: (i) they are nearest neighbours and (ii) they have different circle rates in 2011. Then, if coordinate A has three properties and coordinate B has two properties, I form 6 separate property-pairs. Since I restrict my baseline results to a balanced panel from 2013–2024, this results in 72 (6 x 12) observations for this sample.

Within each property-pair-year, I am now interested in the change in compliance due to the difference in baseline circle rates. This setup is similar in essence to Dube et al. (2010), where there are discontinuous changes in minimum wages across contiguous county-border-pairs in the United States. In their setting, the only thing changing across a pair was the minimum wage, in mine it is the circle rate in 2011 ³². Then, using a pair-year fixed effect (δ_{pt}), I can measure compliance *within* a

³²I also note that the institutional setting here is similar to that of Best et al. (2025), however, I do not use a regression discontinuity for my main specification as that would only allow me to capture a change in compliance due to having a higher circle rate. I show these results as robustness

property-pair-year. Formally, I estimate for all properties:

$$\text{Compliance}_{i(p)t} = \beta_0 + \beta_1 \text{Log(Circle Rate)}_{ip} + \delta_{pt} + \varepsilon_{i(p)t} \quad (10)$$

and

$$\text{Ln(Tax Paid)}_{i(p)t} = \gamma_0 + \gamma_1 \text{Log(Circle Rate)}_{ip} + \delta_{pt} + v_{i(p)t}, \quad \text{for Compliance}_{i(p)t} = 1 \quad (11)$$

Where ‘p’ refers to a property-pair, for property ‘i’ in year ‘t’. β_1 captures the marginal effect on compliance in percentage points and γ_1 captures the percentage change in taxes paid, due to a 1% change in the circle rate.

The identifying assumption for β_1 and γ_1 is that, within a property-pair and year, assignment to the higher baseline circle rate is unrelated to potential compliance and payment outcomes, conditional on the pair-year fixed effects and included controls. In other words, unobserved determinants of these outcomes must not vary systematically with the circle rate within a pair. I cannot directly test whether these unobserved determinants are similar. I therefore test for balance on observable property characteristics as a check of the identifying assumption. The close proximity of the properties and the pair-year fixed effects also help account for shared characteristics such as access to and quality of public services.

The average compliance elasticity then, from Eq 10 is $\frac{\beta_1}{\text{Compliance}}$ and the average elasticity of taxes paid from Eq 11 is γ_1 . I can use these elasticities in Equation 5 to then estimate where the average property in this municipality lies along a revenue function.

3.3 At what circle rate is revenue maximised?

I next calculate the circle rate at which revenues are maximised when I increase all circle rates by the same proportion. I multiply every property’s baseline circle rate by a common factor k , so that

$$\text{Circle Rate}'_i = k \times \text{Circle Rate}_{i,2011}.$$

I call the mean of the resulting distribution the ‘revenue-maximising circle rate’.

Using the condition derived in Eq. 5 and Eq. 10, with the full derivation provided in Section A.2,

the revenue-maximising circle rate is:

$$\text{Circle Rate}^* = \exp \left[\overline{\text{Log}(\text{Circle Rate})}_{2011} + \frac{-\frac{\beta_1}{\gamma_1} - \overline{\text{Compliance}}}{\beta_1} \right]. \quad (12)$$

Here, $\overline{\text{Log}(\text{Circle Rate})}_{2011}$ is the mean log 2011 circle rate. Equivalently, the implied proportional scaling factor is

$$k^* = \frac{\text{Circle Rate}^*}{\exp(\overline{\text{Log}(\text{Circle Rate})}_{2011})},$$

so that a value of $k^* = 1.5$, for example, corresponds to increasing every property's circle rate by 50%.

3.4 What is the total change in revenue from changing circle rates?

I also derive the total change in revenue from changing circle rates in Section A.3, by differentiating Eq 2 by 'cr'. This results in

$$\Delta R \approx \Delta_{cr} A^+(cr) [\beta_1 + \gamma_1 C(cr)] \quad (13)$$

where β_1 is the marginal effect on compliance and γ_1 is the elasticity of taxes paid by compliers. Δ_{cr} is the proportional change in circle rates.

3.5 Heterogeneous Effects

In the average effects I estimate above, there is an inbuilt assumption that all properties through the circle rate distribution respond similarly. The revenue-maximising circle rate is then built off this assumption.

I can go a step further and ease this assumption by allowing properties to respond differentially, based on where they lie in the rate distribution. For instance, there is evidence in the optimal taxation of incomes literature, that elasticities may be larger at both the lower- and the higher-end of the income distribution [Saez \(2001\)](#).

To do this, I allow the effect of an increase in circle rates to vary with the mean circle rate within a pair. Following [Allcott et al. \(2020\)](#), I interact log circle rates with a baseline measure of where the pair lies in the circle rate distribution.

Let ℓ_{ip} denote the log baseline circle rate of property 'i' in pair 'p'. I define

$$Z_{g(p)}^c = \bar{\ell}_{g(p)} - \bar{\ell}, \quad (14)$$

where $\bar{\ell}_{g(p)}$ is the within-pair mean log baseline circle rate and $\bar{\ell}$ is the sample mean. Here, $g(p)$ refers to the matched coordinate group containing property-pair 'p'. Thus, $Z_{g(p)}^c$ measures how far the pair's mean log circle rate lies above or below the sample mean. I use the same measure for payments, so $Z_{g(p)}^a = Z_{g(p)}^c$.

For the extensive margin, I now modify the baseline equation and estimate:

$$\text{Compliance}_{i(p)t} = \alpha_c + \beta_c \ell_{ip} + \theta_c \left(\ell_{ip} \times Z_{g(p)}^c \right) + \delta_{pt} + \varepsilon_{i(p)t}. \quad (15)$$

I continue to use property-pair-year fixed effects (δ_{pt}), as in Eq 10. The interaction coefficient θ_c captures how the compliance response changes as I move across the baseline circle rate distribution.

For the corresponding intensive margin, I estimate the analogous specification among properties that pay positive taxes:

$$\text{Log(Tax Paid)}_{i(p)t} = \alpha_y + \beta_y \ell_{ip} + \theta_y \left(\ell_{ip} \times Z_{g(p)}^a \right) + \delta_{pt} + \nu_{i(p)t}, \quad \text{for Compliance}_{i(p)t} = 1. \quad (16)$$

The marginal effect on compliance is then $\beta_c + \theta_c Z_{g(p)}^c$, and the corresponding effect on log taxes paid is $\beta_y + \theta_y Z_{g(p)}^a$. I evaluate these effects across the circle rate distribution and summarize them in terciles.

My object of interest is the elasticity, which is the proportional change in compliance or payments for a proportional change in the circle rate. For compliance, I divide the marginal effect by mean compliance in the corresponding tercile. These elasticities therefore reflect both the estimated responses and baseline compliance levels. The effect on log taxes paid is however, already an elasticity.

3.6 Heterogeneous revenue-maximizing circle rate

I now use these responses to calculate the revenue-maximising circle rate, increasing all circle rates by the same factor k , as in Section 3.3. I keep each pair's baseline position in the distribution fixed as I increase rates.

I start by using compliance at the mean for each pair's tercile and use the estimated marginal effect to predict how it changes as rates increase. For payments, I start at the mean amount paid by

compliers and use the estimated payment elasticity. Thus,

$$C_i(k) = \bar{C}_i + (\beta_c + \theta_c Z_{g(p)}^c) \log k, \quad (17)$$

$$A_i^+(k) = \bar{A}^+ k^{\beta_y + \theta_y Z_{g(p)}^a}, \quad (18)$$

where $\bar{C}_i$ is mean baseline compliance in the tercile containing property ‘i’, including nonpayers, and $\bar{A}^+$ is mean baseline payment among compliers. As in the pooled exercise, predicted compliance changes as rates rise, so the compliance elasticity changes too. The payment elasticity stays fixed.

I multiply predicted compliance by predicted payments and add these revenues across properties. I then choose the common scaling factor that gives the highest total revenue:

$$k^* = \arg \max_k \sum_i C_i(k) A_i^+(k), \quad 0 < C_i(k) < 1. \quad (19)$$

I solve this numerically, considering only changes that keep predicted compliance between zero and one for every property. I multiply the baseline mean circle rate by k^* to obtain the revenue-maximising circle rate. At an interior maximum, the revenue-weighted sum of the two elasticities equals zero. If compliance reaches a boundary first, that constraint limits the increase in circle rates. I provide the full derivation in Section [A.2](#).

3.7 A distribution of revenue maximizing circle rates

I can also allow a different revenue-maximising circle rate at each point in the distribution. Let b_c and b_y denote the compliance marginal effect and payment elasticity at that point, as defined above. Using the same logic as in Eq [12](#), I obtain

$$CR_i^* = CR_{i0} \exp \left[\frac{-b_c/b_y - \bar{C}_i}{b_c} \right], \quad (20)$$

where CR_{i0} is the baseline circle rate and $\bar{C}_i$ is the corresponding tercile’s mean compliance. This gives a distribution of revenue-maximising circle rates across the city. The formula applies when the implied compliance is between zero and one and the solution is a maximum.

4 Main results

4.1 Defending the choice of the baseline distance

I start with estimating Equations 10 and 11 and present the pooled results at my preferred distance of 0.01 miles, or about 16 metres, in Panel A, columns (1)–(3), of Table 1. Results at 0.005 miles, or about 8 metres, are in Appendix Table A9. I define compliance in two ways. One is my baseline extensive margin measure, where compliance is a binary variable equalling 1 if the assessee pays *any* amount of the tax. The second is a ‘liability-clearing’ measure, where compliance equals 1 if the assessee pays *at least* the recorded liability. I discuss the heterogeneous coefficients in Panel B and tercile elasticities in Panel C in Section 5.

Before analysing my main results in Table 1, I examine the nearest-neighbour coordinate matched distances in my data as the thresholds I present results for are far narrower than usually seen in the literature. My econometric specification rests on comparing properties that are similar to each other, on account of being located very close to each other, such that they lie at the borders of their respective circle areas. In dense cities of India, properties often start looking very different from each other even if they are just one street away and this would introduce bias in my results. In Figure 2 Panel A, I present a sample of coordinate pairs, overlaid on circle areas where darker colors present higher circle rates. The two green circles and the two black ‘Xes’ that are numbered ‘658’ and ‘670’, respectively, represent two coordinate pairs. The pair labeled ‘658’ has a distance less than 16 metres between them, which is also equivalent to taking about 20 steps between the centroid of two properties.

It is clear from this figure that the two coordinates in green are likely only separated by 1 property, they lie on the same side of the street and have varying circle rates. On the other hand, the property-pair labeled ‘670’ has a distance of 0.011–0.05 miles between properties and is separated by 1 whole block. This pair is not in my main estimating sample. Despite being located in areas with varying rates and having a distance that is objectively not very large, in this context, they (i) are separated by several properties and (ii) do not lie at the border of two circles. I discuss and show that they are likely not similar looking properties in a later sub-section when I discuss the validity of my results. Once I constrain my sample to between-property distances below 16 metres, the number of observations that contribute to my results are a smaller percentage of the total observations in my data. In Panel B, I present the distribution of the share of observations at each distance, 16 metres corresponds to a share of 23%, and 8 metres to around 13%.

4.2 What is the average effect on compliance?

How much does compliance on the extensive margin change with a change in circle rates? Using the pooled result in Panel A, column (1), of Table 1, a 1% rise in the baseline circle rate leads to a 0.316 percentage point (pp) fall in compliance. The resultant elasticity using a mean of 0.699 is -0.452. Thus a 1% rise in the baseline circle rate leads to a 0.45% fall in compliance. This point estimate is similar for distances between 8 and 16 metres. In Figure 4, I divide this distance range into 10 percentiles and estimate Equation 10 within each percentile. In Figure A12, I also present estimates using a spatial Regression Discontinuity Design, with distance to the circle border as the running variable and a linear polynomial fit on both sides of the cut-off, similar to Best et al. (2025). Reassuringly, these estimates are similar to those from the within-property-pair specification and attenuate toward zero as the distance grows.

While this extensive margin effect is important, even a partial payment allows the property to be marked as compliant. In Panel A, column (2), I therefore use liability-clearing compliance, where property owners who pay less than their full liability are marked as non-compliant. A 1% rise in the baseline circle rate leads to a 0.233 pp fall in this measure. Although this point estimate is smaller than in column (1), the mean is also lower, at 0.388. The implied elasticity is -0.599. Aside from not paying entirely, another response to higher rates is making incomplete payments.

4.3 What is the average effect on the amount of taxes paid?

The third margin I am interested in is now the elasticity of taxes paid itself, or the intensive margin effect. The dollar amount collected is very important for municipal tax policy and estimating revenue, as discussed in Section 3. Here I ask, between a pair of *compliant* properties how does the amount paid vary as the circle rate increases by 1%? Note that this is different than a pure mechanical increase as properties pay penalties.

I present results on the intensive margin in Panel A, column (3), of Table 1, using Equation 11 and restricting to compliant properties³³. These coefficients should be interpreted as elasticities as this is a log-log specification. A 1% increase in the baseline circle rate leads to a 1.30% increase in taxes paid. Relative to the sample mean of USD 79.19, this amounts to approximately USD 1.03.

This effect, however, combines the mechanical effect with properties selecting into payment. The composition of property-pairs that pay could thus affect this estimate. In Appendix Table A3, Panel

³³The smaller number of observations reflects both this restriction and the exclusion of property-pair-year cells with only one compliant property, which are singleton observations.

A, column (1), I ease the fixed effect structure and allow property-pairs with one payer to remain in the regression. I find a similar elasticity of 1.20.

4.4 Comparing the two effects

The reduction in compliance is completely offset by this increase in the amount of taxes paid due to a rate increase. Plugging the elasticity estimates of compliance and taxes paid on the intensive margin at 16 metres into Equation 5, with $\epsilon_{compliance} = -0.45$ and $\epsilon_{payments} = 1.30$, I get 0.85.

This is higher than the revenue-maximizing condition of zero, suggesting that the average property lies on a rising portion of the revenue function. The payment elasticity is also greater than the absolute value of the compliance elasticity. This remains true using liability-clearing compliance: combining -0.599 with 1.299 gives 0.70.

Overall, a 1% increase in the circle rate leads to a model-implied 0.85% increase in revenues. In Appendix Table A5, I also estimate the response of total taxes paid, including zeros. Panel A, column (3), using PPML, gives an elasticity of 0.550, which is positive but not statistically significant at the 10% level (p-value = 0.106).

4.5 Are these estimates causal?

For these estimates to be causal, differences in baseline circle rates within a property-pair and year must be unrelated to unobserved determinants of compliance and payments. I cannot directly test this assumption, so I use balance on observables as a partial check. In Panel A, columns (1)–(3), of Table 2, I present the pooled balance tests for the log residual defined in Equation 9, the number of floors and construction type. The corresponding coefficient plot is in Figure 5.

The pooled coefficient on the log residual at 0.01 miles is not statistically different from zero, supporting my estimation strategy. In Figure A8, Panel A, I also present this coefficient for distance thresholds through 0.05 miles. Beyond 16 metres the point estimate becomes more negative, consistent with properties becoming less similar as the distance between them grows, though the estimates remain imprecise³⁴. Coefficients for number of floors and construction type in Panel A, columns (2)–(3), are also statistically insignificant at 0.01 miles. ‘Number of floors’ is an integer variable going from 1 to 12 and construction type is a categorical variable that is equal to 1 if the property is built with cement, 2 if it is partially built with cement, and 3 if it is built using other

³⁴The residual turning negative is also indicative that as the circle rate increases, the liability increase for the higher circle rate property happens at a slower rate.

less-durable materials. I thus find that properties are balanced across all three characteristics. I return to the corresponding heterogeneous balance tests in Section 5.

4.6 Is there an impact of enforcement?

Due to enforcement, compliance responses to rate increases could be muted, causing the elasticities I find above to be under-estimates. I check this using the KMC policy rule that reminder letters are sent once cumulative unpaid liabilities exceed INR 10,000, about USD 117 at the exchange rate used in the analysis. This is a light touch enforcement action. Stricter action by enforcement officers only applies when the liability crosses 10000 USD, this often applies only to commercial properties or very large residential properties, both of which do not exist in my sample.

I then estimate Equation 10 using observations that fall below the enforcement threshold and present the pooled results in Panel A, columns (1)–(3), of Table 3. Since both properties must remain in a pair-year, the sample is smaller. A 1% increase in the circle rate leads to a 0.369 pp fall in compliance, implying an elasticity of -0.499, compared with -0.452 in the restricted baseline sample. The conditional-payment elasticity is 1.658, compared with 1.299 in the restricted baseline sample. These point estimates remain qualitatively similar to the main results.

I consider house price growth, construction and effects during COVID, and discuss them in detail in Appendix Section A.6.

4.7 Robustness

Lastly, I check robustness in Appendix Table A8, with the preferred 16-metre estimates in column (2). In Panel B, I weight observations inversely by each property’s number of distinct matched pairs in that year. The compliance coefficient is -0.238, compared with -0.316 in the restricted baseline sample, and is less precisely estimated. In Panel C, I allow an unbalanced panel and obtain -0.341. In Panel D, I add fixed effects for floors and construction type and obtain -0.290. These checks retain the negative direction of the compliance response.

5 Heterogeneous responses across rates

Pooled results suggest that higher circle rates reduce compliance but increase the amount paid by compliers. I now ease the assumption that a single elasticity describes the whole city. Using the continuous interaction specification in Section 3.5, I allow these responses to vary with the local within-pair mean log baseline circle rate. I present the coefficients in Panel B of Table 1 and

summarize the elasticities in terciles of this distribution in Panel C.³⁵ Figure 7 plots the corresponding compliance and conditional-payment elasticities.

5.1 How does the elasticity of compliance change as circle rates increase?

Compliance becomes more elastic at higher circle rates. In the low tercile, a 1% increase in the circle rate leads to a 0.08% fall in compliance. This estimate is imprecise and is not statistically different from zero. In the middle and high terciles, the same increase leads to falls of 0.41% and 0.67%, respectively, both statistically significant at the 1% level. I reject equality of the three elasticities (joint $p < 0.001$). Thus, the pooled elasticity of -0.452 masks significant differences in proportional compliance responses: raising circle rates leads to a larger proportional loss of payers at the top of the distribution.

I calculate these compliance elasticities by dividing the fitted marginal effect by mean compliance in each tercile. They measure proportional changes in compliance and reflect both the estimated response and the baseline compliance level.³⁶ Liability-clearing compliance also differs across terciles (joint $p = 0.008$), but does not show the same pattern of increasingly negative elasticities. Its point estimates are -0.694, -0.654 and -0.562 in the low, middle and high terciles. The response of paying anything and the response of clearing the full liability therefore need not vary in the same way across the circle rate distribution.

5.2 How does the elasticity of payments change as circle rates increase?

The point estimates for payments among compliers rise across terciles as well. A 1% increase in the circle rate raises conditional payments by 0.30% in the low tercile, 1.14% in the middle and 1.89% in the high tercile. The middle and high estimates are statistically different from zero, while the low estimate is imprecise. However, I cannot reject equality of the three payment elasticities (joint $p = 0.345$). The main evidence of heterogeneity in Table 1 thus comes from the extensive margin compliance response. The payment estimates suggest larger increases among compliers at higher rates, but do not establish that these responses differ across terciles.

The amount paid can also change because property owners pay a different share of what they

³⁵Each distinct property pair contributes once when I set the tercile cutpoints. I assign both properties in a pair to the same tercile. I also do this to be able to define a compliance elasticity as the original response function is a marginal effect. For the elasticity, I need to divide by the mean.

³⁶The interaction test in Panel B does not reject a common marginal effect on compliance ($p = 0.259$). This tests a different hypothesis from the equality of elasticities in Panel C, since elasticities also depend on tercile-specific mean compliance. The joint elasticity test in Panel B does not establish that every pair of terciles differs significantly.

owe. In Appendix Table A4, I find that the share-paid response becomes larger at higher circle rates, with significant differences across terciles in both OLS and PPML (joint $p < 0.001$). This suggests that the payment response involves changes in the share paid as well as the higher liability itself. Payments can exceed recorded liability because they include penalties.

I also estimate the responses on total taxes paid, including zeros for nonpayers. This combines the loss of payers with the change in the amount paid by those who continue to pay. In the PPML specification in Appendix Table A5, the elasticity rises from -1.287 in the low tercile to 0.196 in the middle and 1.570 in the high tercile. Only the high-tercile estimate is statistically different from zero but I cannot reject equality across terciles (joint $p = 0.207$). Figure A20 presents the PPML profiles for total payments and the share paid among payers. The total-payment estimates suggest larger revenue gains at higher rates, but remain too imprecise to establish differences across the distribution.

A possible explanation for this is that liability increases at lower circle rates discourage marginal payers, leaving only owners with greater ability or willingness to pay.

5.3 Checking for balance on observables and effects of enforcement

I next check whether the compliance pattern could reflect differences in property characteristics or enforcement. In Panel C of Table 2, I find no statistically significant differences across terciles in the balance estimates for the residual, number of floors or construction type. In Table 3, I restrict attention to properties below the enforcement threshold. Compliance elasticities still become more negative at higher circle rates, and I reject their equality (joint $p < 0.001$). Conditional-payment elasticities again do not differ significantly. These checks suggest that the compliance pattern is not solely a response to the threat of reminder letters at higher liabilities.

For revenues, these responses work in opposite directions. More elastic compliance makes a rate increase more costly in lost payers, while larger conditional-payment responses increase collections from those who pay. I use both fitted responses in the revenue-maximization exercise that follows, including payment responses whose differences across the distribution are imprecisely estimated.

6 What is the revenue maximising circle rate?

I use the fitted response functions to maximize aggregate predicted revenue under a common proportional scaling of the schedule. This RMCR exercise and its paired bootstrap retain the full-support sample, while Table 1 and Figures 7, A19 and A20 use the separate USD 280 restriction.

6.1 Using pooled elasticities

I start by using Eq 12 to calculate the Revenue Maximising Circle Rate (RMCR). I consider a uniform proportional increase in all circle rates, so that the whole circle-rate schedule shifts up by a common scaling factor k . The RMCR summarizes the resulting schedule.

Using the pooled compliance and conditional-payment estimates and the relevant sample moments, I obtain an RMCR of 858 USD. The corresponding scaling factor is $k^* = 4.011$. The RMCR schedule is thus about 301% higher than the current circle-rate schedule.

I plot this schedule in Figure 8. The blue line shows the RMCR schedule using pooled elasticities. The dashed grey line is the 45-degree line representing the baseline circle rate.

6.2 Using heterogeneous elasticities

I numerically maximize aggregate predicted revenue using the fitted compliance and conditional-payment response functions. Each observation is assigned its tercile-level mean compliance, estimated conditional payment elasticity, and estimated compliance semi-elasticity. As circle rates rise, mean compliance changes from its tercile starting level. Thus as circle rates rise, compliance elasticity within the tercile changes as per predicted mean compliance. I go over this process in some detail in Section 3.

Revenues are maximised when $k^* = 2.217$, corresponding to an increase of about 122% and an RMCR of USD 474. I plot the resulting schedule in Figure 8.

I use Eq 20 to do a version of this exercise that uses local elasticities to measure local RMCRs, such that each baseline circle rate gets a different RMCR and the change is not just a proportional rise in the schedule. I plot this exercise in Appendix Figure A21. Due to a low compliance response at smaller circle rates, it assigns a very high RMCR value, making this plot appear regressive.

6.3 Inference on revenue-maximizing circle rates

I quantify uncertainty using a clustered bootstrap that resamples property pairs. Each of the 499 successful replications reconstructs the pair-level terciles, recomputes their compliance means, re-estimates both response models and estimates the pooled and heterogeneous optima on the same resampled data. The heterogeneous specification gives $k^* = 2.217$ (bootstrap SE 0.364; 95% percentile CI [1.636, 3.118]), corresponding to an RMCR of \$474 (95% CI [\$353, \$666]). The pooled specification gives $k^* = 4.011$ (SE 1.234; 95% CI [2.615, 7.354]), corresponding to \$858 (95% CI [\$561, \$1,570]).

The relevant comparison however, is the difference, not whether the individual confidence intervals overlap. The pooled-minus-heterogeneous difference is 1.794 (SE 1.293; 95% CI [0.316, 5.193])³⁷. This interval excludes zero, indicating a significantly lower heterogeneous scaling factor at the 5% level. The RMCR difference is \$384, with a 95% percentile CI of [\$68, \$1,105]. These intervals describe uncertainty in the constrained model-based exercise.

6.4 Using heterogeneous elasticities by ward

In a separate ward-level illustration, I use fitted response functions and the elasticities measured in Section 5 to find scaling factors and RMCRs separated by ward. Ward 15 ends up with the largest change, amounting to 350%, while ward 48 actually suggests a schedule 2.6% lower than the current schedule. I present these figures in Figure A22. Within each ward, not all circle rates are identified, leading to differences in the range of the predicted schedule.

7 Impact on revenues due to changing area-based assessments

Next, I use Eq 13 and calculate the potential revenue to be gained under 8 different scenarios. Five of these are circle rate updates to alternate years, 2016, 2018, 2020, 2022 and 2024. Then, I consider a yearly 2% increase, scaling to the pooled RMCR, which raises the schedule by about 301%, and the heterogeneous RMCR, which raises it by about 122%. For each of these, I use both pooled and heterogeneous elasticities to measure the total change in revenues. I plot these changes in Figure 9.

In Panel A, using pooled elasticities, I find that scaling to the Revenue Maximising Circle Rate has a large 100% increase in revenues. The yearly 2% increase, which also represents a much smaller change to circle rates than suggested by the RMCR, results in the smallest revenue gain of 11%. Updating the base year results in changes ranging from 56–63%, but these do not linearly rise due to the way circle rates in the city grew. Since circle rates dropped in 2024, the change in 2024 is also the lowest.

These patterns are somewhat consistent in panel B, which uses heterogeneous elasticities as well. However, scaling to the heterogeneous RMCR now has a much smaller 33% increase in revenues. This first-order approximation is quantitatively very similar to the increase suggested by updating the base year. Thus, allowing properties to respond heterogeneously seems to have a substantial impact on total revenues collected. In Figure A23 I plot revenues from ward-level RMCRs. Interestingly, ward no. 48, which had a decrease in the RMCR schedule, continues to experience a 1.04% increase in total

³⁷The bootstrapped distribution is right-skewed, which causes the CIs to be above 0 despite the SE of 1.293

revenues.

8 Conclusion

In this paper I examine how higher property-tax liabilities affect the decision to pay and the amount paid conditional on payment. In a setting with near-stagnant assessed values, I use spatial variation in baseline circle rates to estimate the revenue consequences of updating area-based assessments. I find that higher circle rates reduce compliance, but the pooled decline is more than offset by higher payments among those who comply, placing the average property on the rising portion of the revenue curve.

I test for heterogeneity and allow properties that lie on different parts of the circle rate distribution to respond differentially. I find that compliance elasticities differ significantly across the baseline circle-rate terciles, with more negative point estimates at higher rates. However, I find that though conditional-payment elasticities increase with the circle rate, I cannot reject that they are statistically similar to each other. Using pooled estimates results in an extrapolated revenue-maximizing common scaling factor of 4.011, or a 300% tax rate increase. However, using heterogeneous estimates lowers this scaling factor to 2.217, or a 122% tax rate increase. Estimating a local revenue maximising rate at each baseline circle using these estimates results in a regressive tax-rate schedule.

Finally, I use these estimates to measure the implied revenue response, and find that although updating to the model-implied revenue-maximizing schedule yields the largest gains, simply updating outdated circle rates to a more recent base year recovers substantial revenue.

References

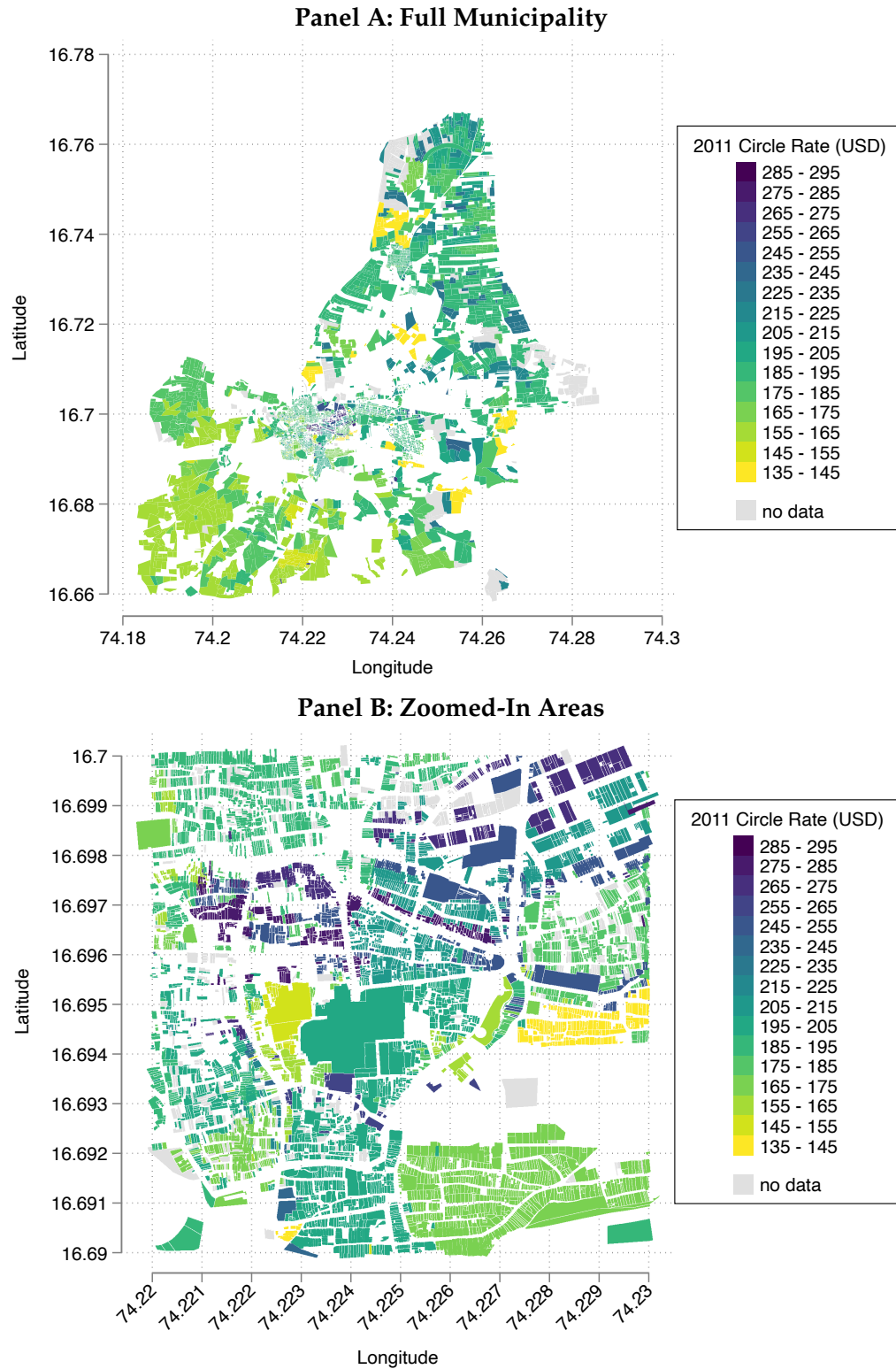
- Ajzenman, Nicolas, Guillermo Cruces, Ricardo Perez-Truglia, Darío Tortarolo, and Gonzalo Vazquez-Bare, "From Flat to Fair? The Effects of a Progressive Tax Reform," December 2024.
- Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow, "The Welfare Effects of Social Media," *American Economic Review*, 2020, 110 (3), 629–676.
- Amornsiripanitch, Natee, "Why Are Residential Property Tax Rates Regressive?," Working Paper 22-02, Federal Reserve Bank of Philadelphia 2022.
- Anagol, Santosh, Allan Davids, Benjamin B Lockwood, and Tarun Ramadorai, "Diffuse bunching with frictions: Theory and estimation," Technical Report, National Bureau of Economic Research 2024.
- Aparicio, Gabriela, Paul E. Carrillo, and M. Shahe Emran, "Taxes, Prisons, and CFOs: The Effects of Increased Punishment on Corporate Tax Compliance in Ecuador," *SSRN Electronic Journal*, 2011.
- Avenancio-León, Carlos F. and Troup Howard, "The Assessment Gap: Racial Inequalities in Property Taxation," *The Quarterly Journal of Economics*, 2022, 137 (3), 1383–1434.
- Awasthi, Rajul, Mohan Nagarajan, and Klaus W Deininger, "Property taxation in India: Issues impacting revenue performance and suggestions for reform," *Land use policy*, 2021, 110, 104539.
- Bachas, Pierre and Mauricio Soto, "Corporate taxation under weak enforcement," *American Economic Journal: Economic Policy*, 2021, 13 (4), 36–71.
- Badel, Alejandro and Mark Huggett, "The Sufficient Statistic Approach: Predicting the Top of the Laffer Curve," *Journal of Monetary Economics*, 2017, 87, 1–12.
- Bahl, Roy and Jorge Martinez-Vazquez, *The property tax in developing countries: Current practice and prospects*, Lincoln Institute of Land Policy., 2007.
- Balán, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel, "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo," *American Economic Review*, March 2022, 112 (3), 762–797.
- Bergeron, Augustin, Gabriel Tourek, and Jonathan L Weigel, "The state capacity ceiling on tax rates: Evidence from randomized tax abatements in the drc," *Econometrica*, 2024, 92 (4), 1163–1193.
- Best, Michael C., Luigi Caloi, François Gerard, Evan Plous Kresch, Joana Naritomi, and Laura Zoratto, "Greener on the Other Side: Inequity and Tax Compliance," Working Paper 34062, National Bureau of Economic Research 2025. Revised August 2026.
- Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem, "Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan," *Journal of political Economy*, 2015, 123 (6), 1311–1355.

- Brockmeyer, Anne, Alejandro Estefan, Karina Ramírez Arras, and Juan Carlos Suárez Serrato**, “Taxing property in developing countries: Theory and evidence from Mexico,” Technical Report, National Bureau of Economic Research 2021.
- Carrillo, Paul E, Edgar Castro, and Carlos Scartascini**, “Public good provision and property tax compliance: Evidence from a natural experiment,” *Journal of Public Economics*, 2021, 198, 104422.
- Dube, Arindrajit, T. William Lester, and Michael Reich**, “Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties,” *The Review of Economics and Statistics*, 2010, 92 (4), 945–964.
- Duflo, Esther, Rema Hanna, Benjamin A. Olken, and Shreya Tandon**, “The Limits of Government Outsourcing: Property Tax Re-Assessments in India,” Working Paper 35610, National Bureau of Economic Research 2026.
- Emran, M Shahe and Joseph E Stiglitz**, “On selective indirect tax reform in developing countries,” *Journal of public Economics*, 2005, 89 (4), 599–623.
- Ferraz, Claudio, Dirk Foremny, and Juan Francisco Santini**, “Revenue Slumps and Fiscal Capacity: Evidence from Brazil,” NBER Working Paper 32440, National Bureau of Economic Research 2024.
- Gruber, Jon and Emmanuel Saez**, “The elasticity of taxable income: evidence and implications,” *Journal of public Economics*, 2002, 84 (1), 1–32.
- Guner, Nezih, Martin Lopez-Daneri, and Gustavo Ventura**, “Heterogeneity and Government Revenues: Higher Taxes at the Top?,” *Journal of Monetary Economics*, 2016, 80, 69–85.
- Holter, Hans A., Dirk Krueger, and Serhiy Stepanchuk**, “How Do Tax Progressivity and Household Heterogeneity Affect Laffer Curves?,” *Quantitative Economics*, 2019, 10 (4), 1317–1356.
- Hou, Yilin, Lei Ding, David J. Schwegman, and Alaina G. Barca**, “Assessment Frequency and Equity of the Property Tax: Latest Evidence from Philadelphia,” *Journal of Policy Analysis and Management*, 2025, 44 (2), 483–507.
- Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken**, “Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors *,” *The Quarterly Journal of Economics*, February 2016, 131 (1), 219–271.
- Knebelmann, Justine, Victor Pouliquen, and Bassirou Sarr**, “Discretion versus algorithms: bureaucrats, tax equity and acceptability,” Technical Report, Institute for Fiscal Studies 2026.
- Kumar, Anil and Che-Yuan Liang**, “Estimating Taxable Income Responses with Elasticity Heterogeneity,” *Journal of Public Economics*, 2020, 188, 104209.
- Manwaring, Priya and Tanner Regan**, “Public Disclosure and Tax Compliance: Evidence from Uganda,” report, The Institute of Development Studies and Partner Organisations September 2024.
- Norregaard, John**, “Taxing Immovable Property: Revenue Potential and Implementation Challenges,” IMF Working Paper 13/129, International Monetary Fund 2013.

- Okunogbe, Oyebola and Gabriel Tourek**, “How Can Lower-Income Countries Collect More Taxes? The Role of Technology, Tax Agents, and Politics,” *Journal of Economic Perspectives*, February 2024, 38 (1), 81–106.
- Pomeranz, Dina**, “No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax,” *American Economic Review*, 2015, 105 (8), 2539–2569.
- Saez, Emmanuel**, “Using Elasticities to Derive Optimal Income Tax Rates,” Technical Report 2001.
- , **Joel Slemrod, and Seth H. Giertz**, “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review,” *Journal of Economic Literature*, 2012, 50 (1), 3–50.
- Serrato, Juan Carlos Suárez and Owen Zidar**, “The structure of state corporate taxation and its impact on state tax revenues and economic activity,” *Journal of Public Economics*, 2018, 167, 158–176.
- Slemrod, Joel**, “Tax Compliance and Enforcement,” *Journal of Economic Literature*, December 2019, 57 (4), 904–954.
- Tourek, Gabriel Z, Arthur Laroche, Augustin Bergeron, Joana Naritomi, Jonathan L Weigel, and Marina Mavungu Ngoma**, “Does Progressivity Raise Tax Capacity? Experimental Evidence from the DR Congo,” Technical Report, National Bureau of Economic Research 2026.

9 Figures

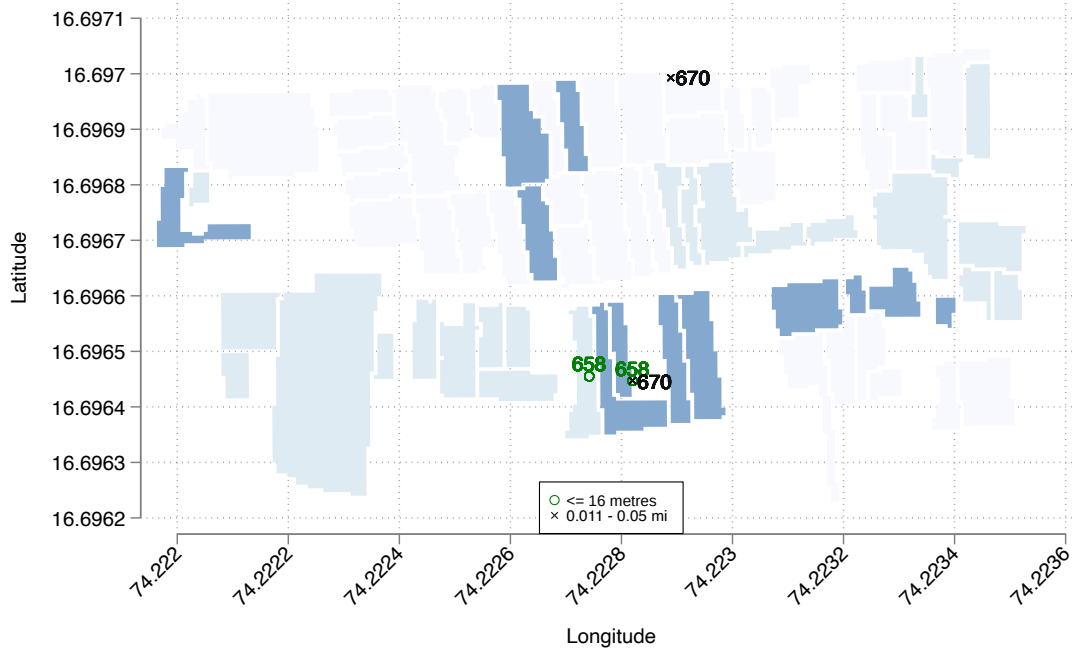
Figure 1
Discontinuities in Baseline Circle Rates



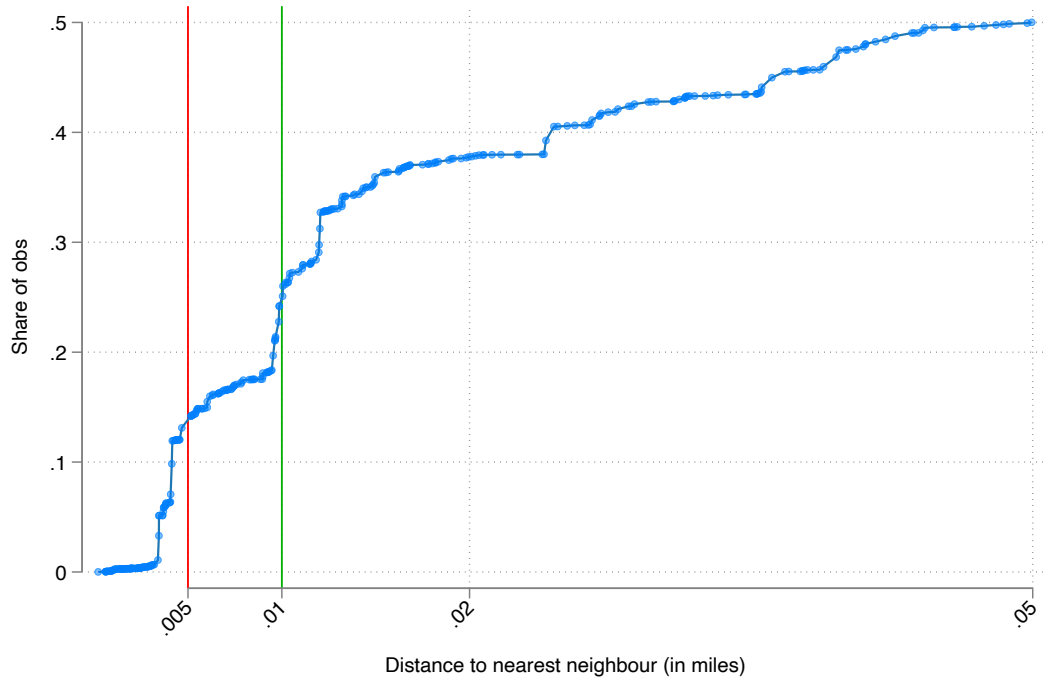
Notes: This figure presents survey areas shaded according to their 2011 circle rates. Panel A shows the full municipality of Kolhapur, while Panel B presents a zoomed-in view. Darker areas have higher circle rates than lighter areas. Circle rates range from USD 135 to USD 295. Areas shown in grey could not be matched to the circle-rate data.

Figure 2
Picking the Threshold Nearest-Neighbour Matched Distance

Panel A. An Example of property pairs at different distances



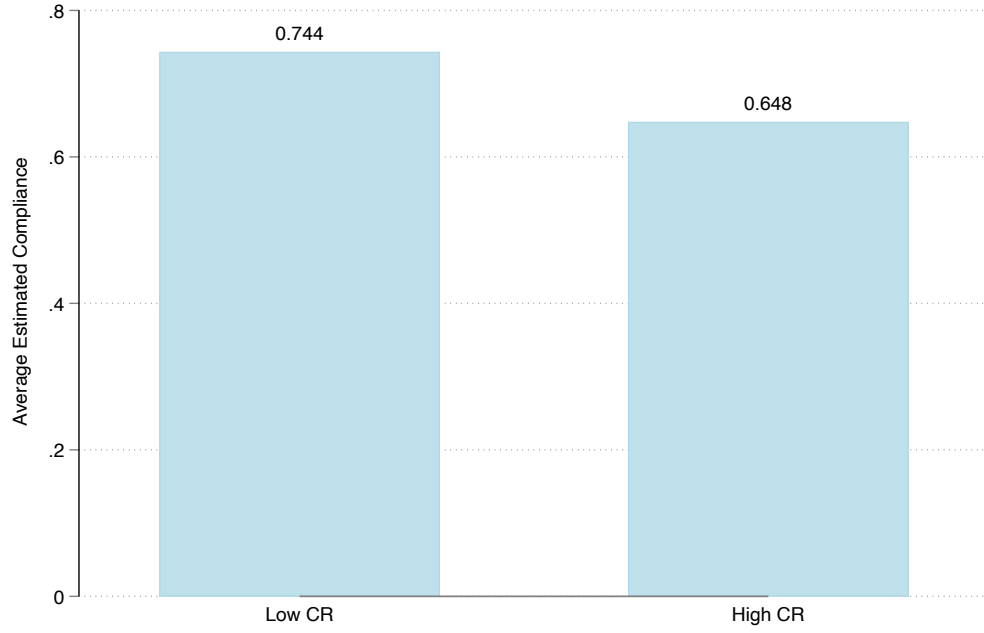
Panel B. Cumulative share of observations at each distance



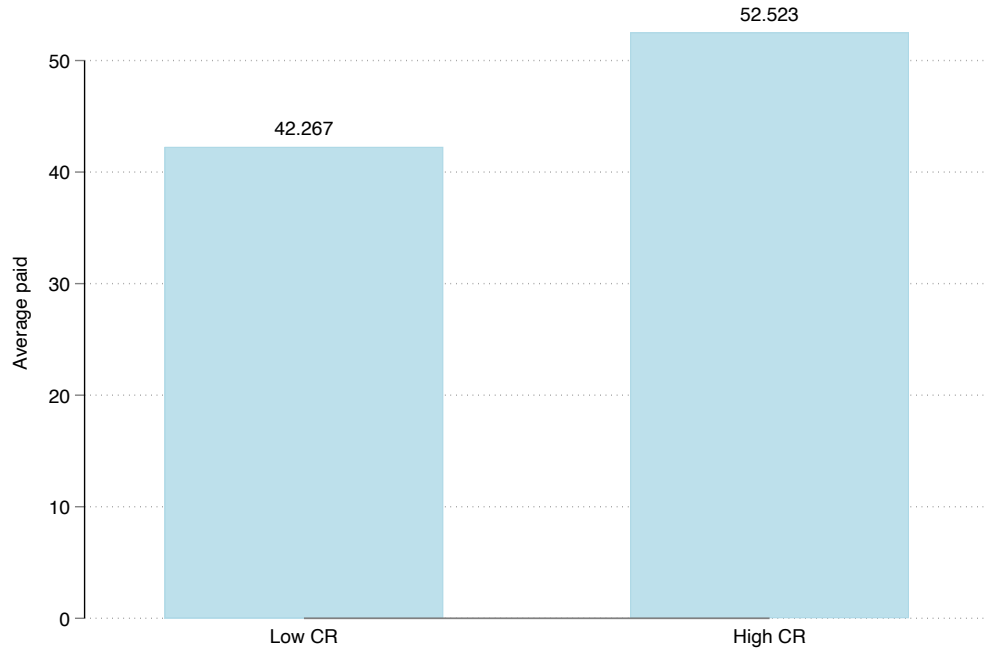
Notes: This figure presents a sample of property pairs at varying distances in Panel A and the cumulative share of observations at each distance in Panel B. In Panel A, the green labels present a pair of coordinates with a distance of 16 metres between them and the black labels present a pair of coordinates with a distance of 0.011 to 0.05 miles between them. Coordinates are overlaid on circle areas, with darker shades representing higher circle rates in 2011. In Panel B, each point is the cumulative share of observations at that distance threshold. All distances between coordinates are a result of a nearest neighbour matching algorithm. In Panel B, the red and green lines mark 8 metres (appendix tables) and 16 metres (main tables), respectively.

Figure 3
Mean Differences between High and Low CR properties

Panel A. Compliance



Panel B. Amount Paid



Notes: This figure presents average predicted binary compliance and the average predicted amount of taxes paid in Panels A and B, respectively. Compliance is measured using the following equation:

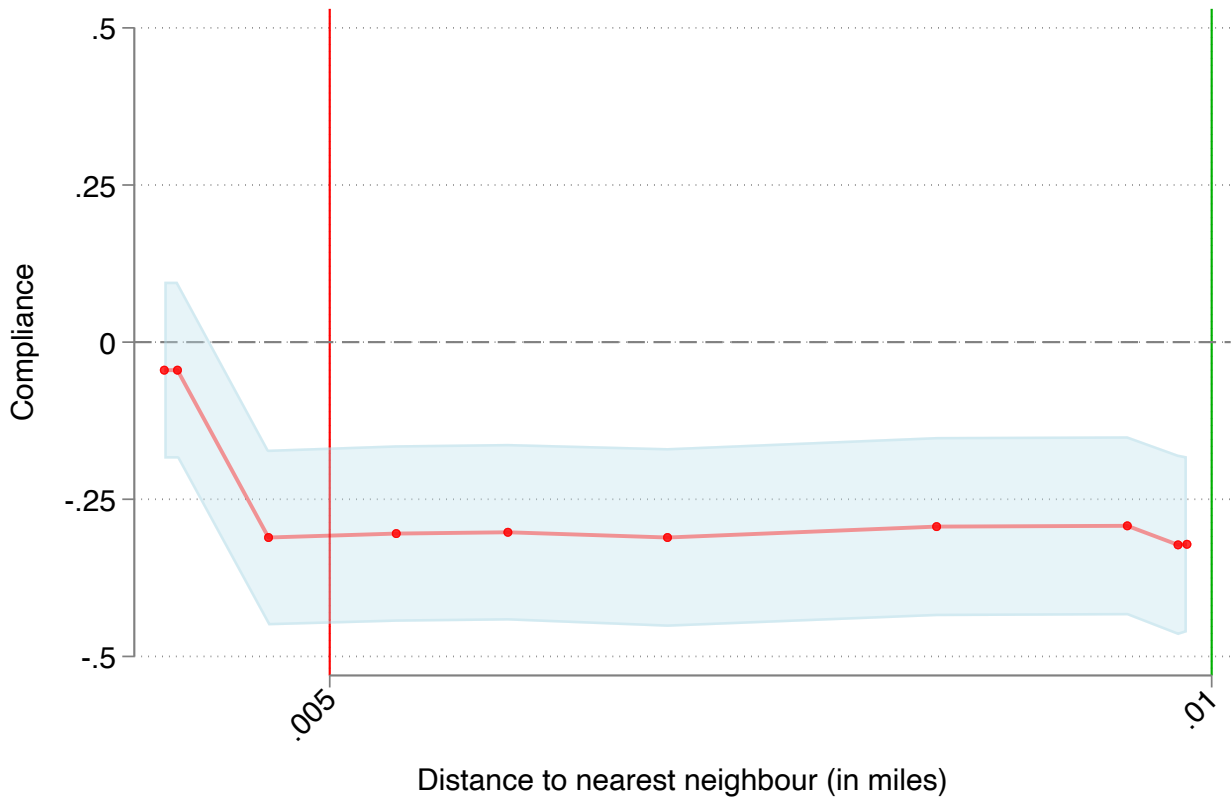
$$\text{Compliance}_{i(p)t} = \beta_0 + \beta_1 \text{Log}(\text{Circle Rate})_{ip} + \delta_{pt} + \varepsilon_{i(p)} \quad (21)$$

and taxes paid for compliers using:

$$\text{Log}(\text{Tax Paid})_{i(p)t} = \beta_0 + \beta_1 \text{Log}(\text{Circle Rate})_{ip} + \delta_{pt} + \varepsilon_{i(p)} \quad (22)$$

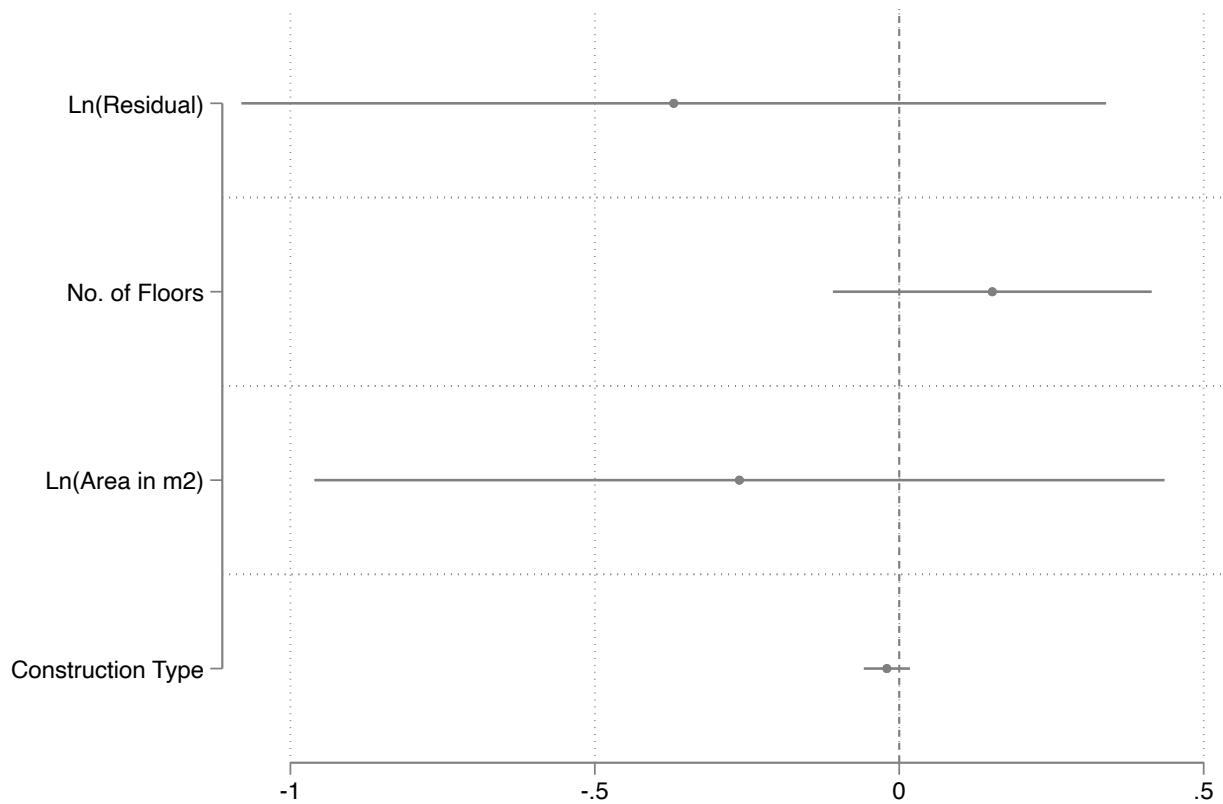
Within each pair ‘p’, I mark the property with the higher circle rate as being a ‘High CR’ property and the one with the lower circle rate as being a ‘Low CR’ property. I estimate $\widehat{\text{Compliance}}$ and $\widehat{\text{Log}(\text{Tax Paid})}$, exponentiate $\widehat{\text{Log}(\text{Tax Paid})}$ and present means separately for High and Low CR properties.

Figure 4
Point Estimate on Compliance stays constant through 16 metres



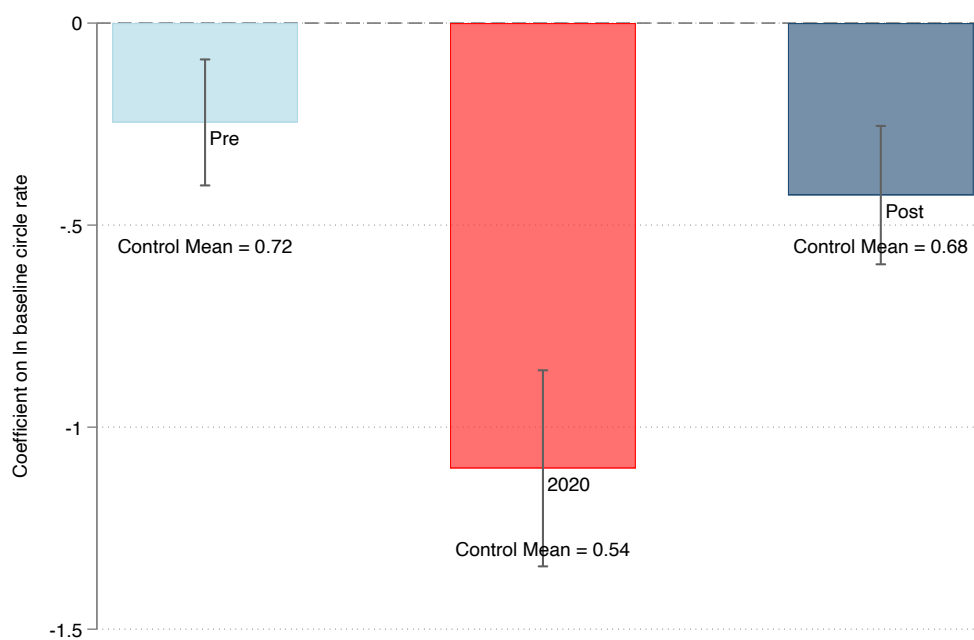
Notes: This figure presents the β_1 from estimating Equation 21 for varying distance thresholds. $Compliance_{i(p)t}$ refers to the compliance status of a property 'i' in year 't' within property-pair 'p'. Property pairs are formed using a nearest-neighbour distance matching algorithm. Each point in this graph represents the estimated β_1 from changing the distance threshold that defines the property-pairs. Compliance is defined as an indicator for whether a property paid *any* tax in a given year. β_1 should be interpreted as the change in compliance due to a 1% change in the 2011 Circle Rate. All specifications include property-pair-by-year and ward fixed effects. The red and green vertical lines indicate the two distance thresholds used for the main results, 8 metres and 16 metres, respectively. The shaded light-blue region denotes 95% confidence intervals.

Figure 5
No differences in Observables within a Property-Pair

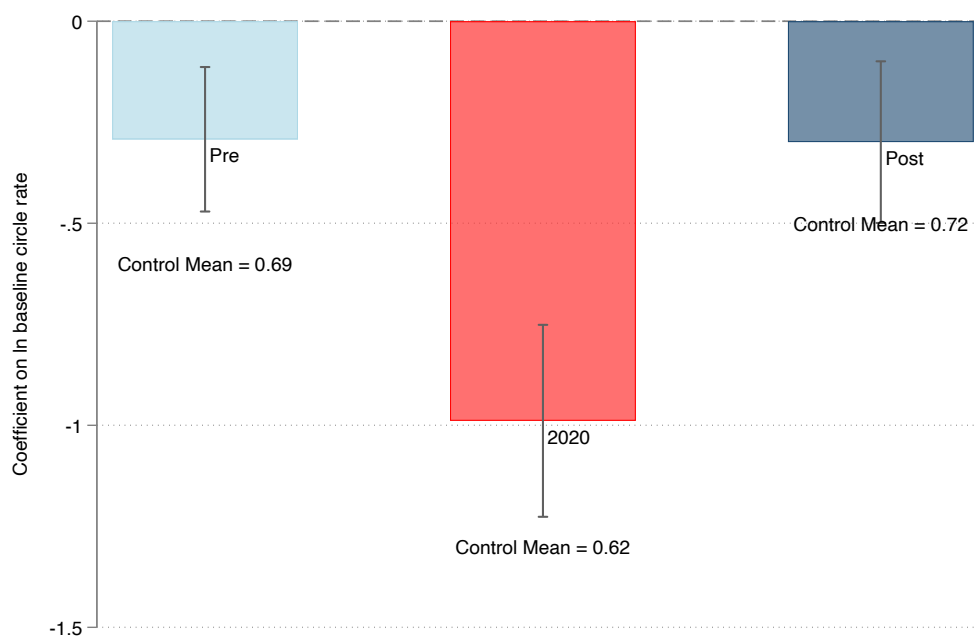


Notes: This figure reports balance tests for observable property characteristics within matched property pairs. Each point is the coefficient on log baseline 2011 circle rate from the main property-pair-by-year and ward fixed-effects specification, estimated separately with the listed characteristic as the dependent variable. Ln(Residual) is the residual component of the valuation formula after accounting for the circle rate, as defined in Equation 9. Number of floors, construction type, and built-up area are from Google Footprints where available. Whiskers show 95 percent confidence intervals.

Figure 6
A Large Fall in Compliance during COVID-19



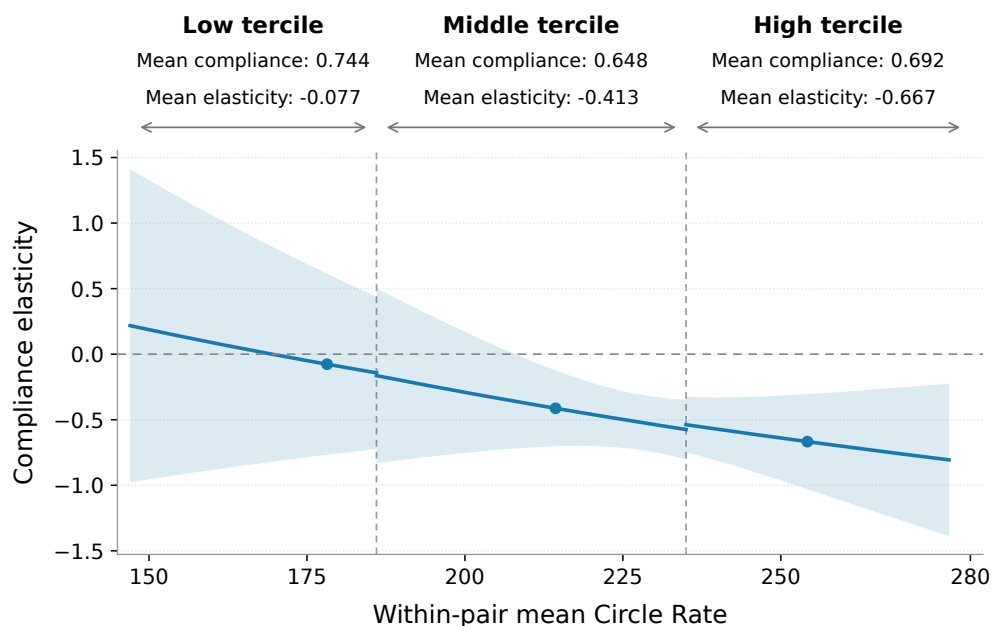
Panel A. 8 metres



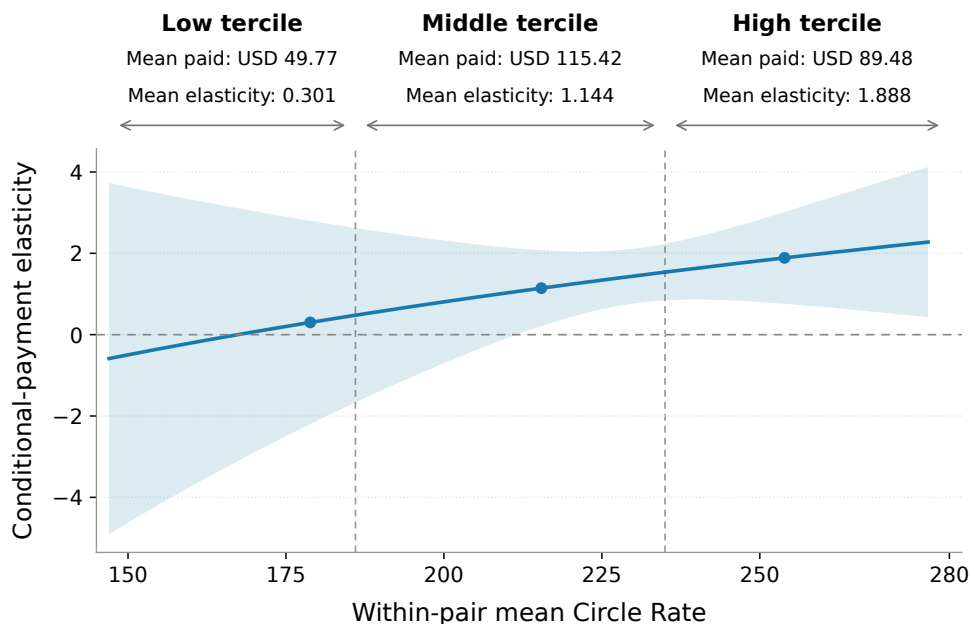
Panel B. 16 metres

Notes: This figure presents the coefficient from estimating Equation 21 separately for years pre-2020, post-2020 and 2020. For 2020, the specification uses only property-pair and ward fixed effects. Panel A presents estimates using a threshold distance of 8 metres and Panel B presents estimates using a distance of 16 metres. Confidence intervals are presented at the 95% level. Fig A13 presents a version of this figure where Equation 21 is estimated separately for each year.

Figure 7
Elasticity of Compliance and Taxes Paid: Heterogeneous Responses across the CR distribution



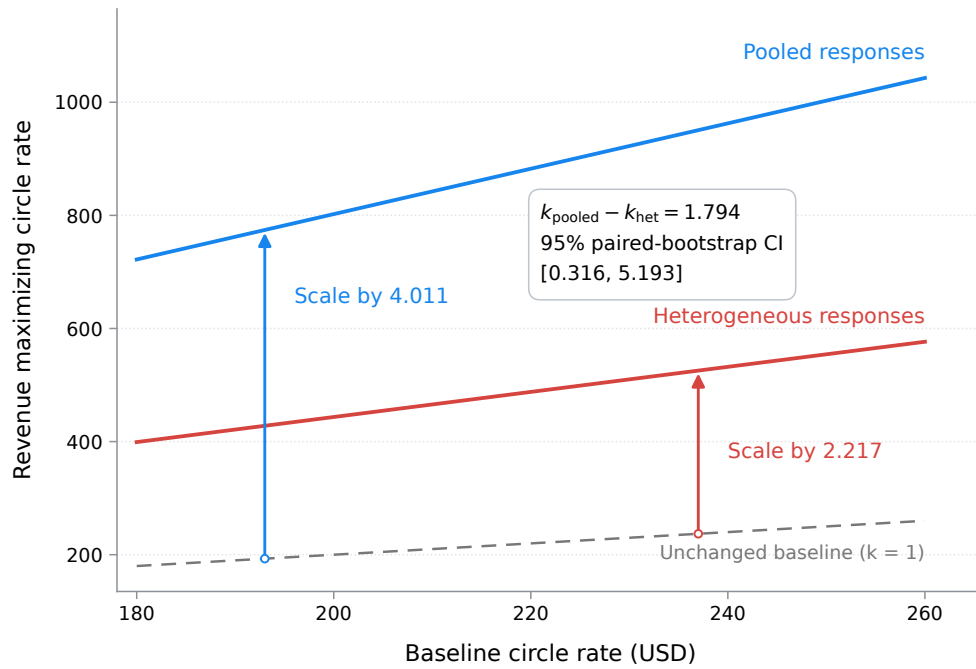
Panel A. Elasticity of Compliance



Panel B. Elasticity of Taxes Paid

Notes: This figure presents the continuous interaction specifications in Equations 15 and 16. Panel A reports the compliance elasticity; Panel B reports the log-payment response among payers. Both panels use terciles recomputed from property-pairs, counting each pair once (cutpoints: USD 186 and USD 235). Labels and dots summarize each outcome's estimation sample and match Table 1, Panel C. Panel A divides compliance slopes by the corresponding tercile's mean compliance; normalization changes generate boundary jumps. Panel B reports log-payment slopes, which are elasticities directly; mean payments are descriptive. Shading shows 95% confidence intervals. The joint tercile tests give $F(2,142)=12.781$ ($p<0.001$) for compliance and $F(1,129)=0.897$ ($p=0.345$) for payments. Inference uses the full two-way clustered covariance matrix with means and cutpoints fixed.

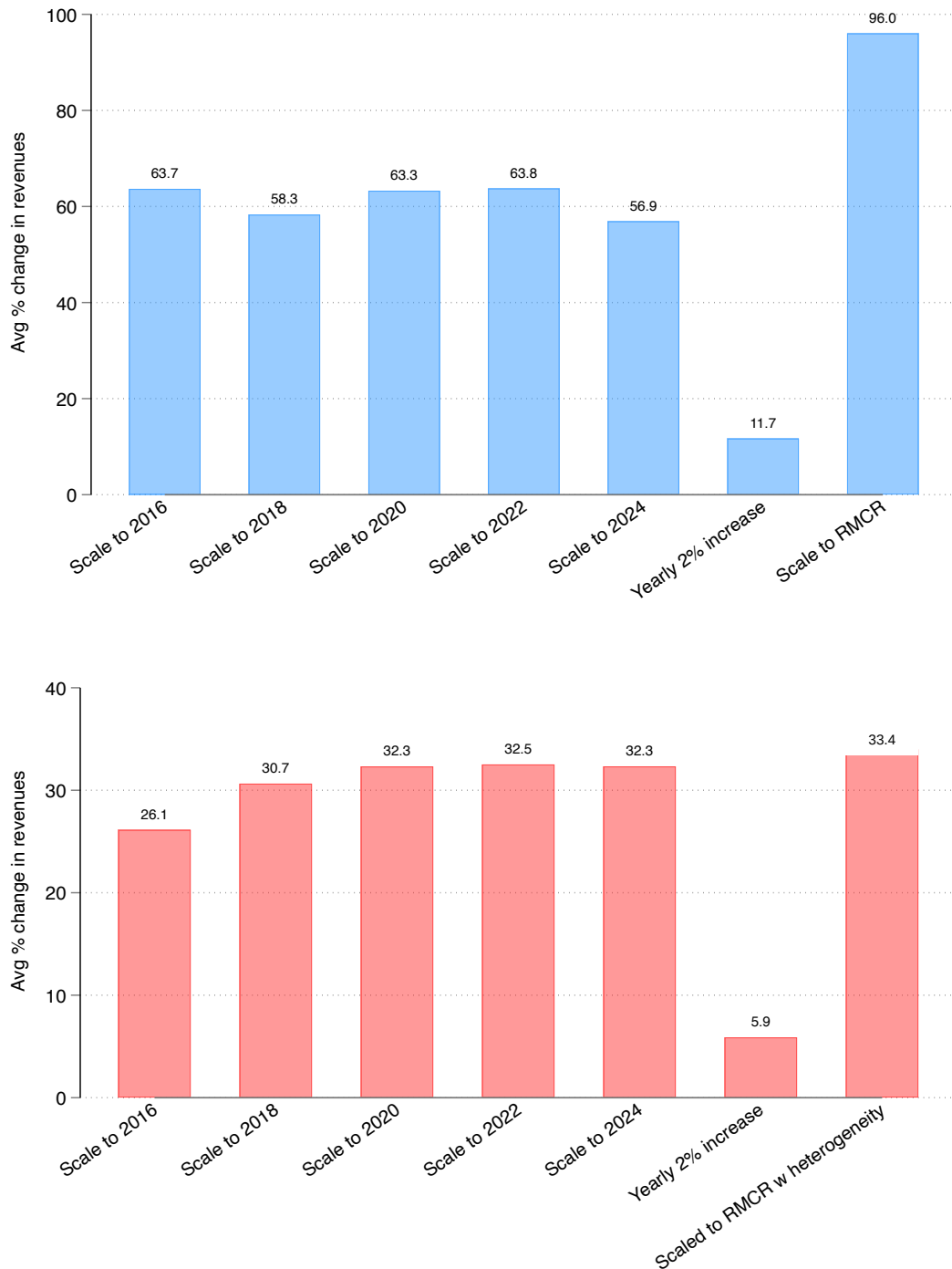
Figure 8
How should the circle rate schedule be scaled such that revenue is maximised?



Common scaling of the baseline circle-rate schedule

Notes: This figure presents model-implied revenue-maximizing circle-rate schedules. It compares a uniform citywide scaling based on pooled elasticities (blue line) with a uniform scaling that allows compliance and payment elasticities to vary over the baseline circle-rate distribution (red line). The arrows indicate multiplying every baseline rate by 4.011 under pooled responses and 2.217 under heterogeneous responses, corresponding to RMCRs of USD 858 and USD 474, respectively. The box reports the paired difference in scaling factors, 1.794, and its 95% percentile interval, [0.316, 5.193]. Both optima are estimated on the same 499 property-pair bootstrap samples; terciles, compliance means and response models are reconstructed in each draw. The paired interval excludes zero. Individual-factor intervals are [2.615, 7.354] and [1.636, 3.118], respectively. The RMCR difference is USD 384, with a 95% interval of [USD 68, USD 1,105].

Figure 9
Percentage Change in Revenue



Notes: First-order predicted revenue changes use Equation 13. Panel A uses pooled responses and Panel B heterogeneous responses. The base-year scenarios replace 2011 circle rates with the observed 2016, 2018, 2020, 2022 or 2024 rates. Other scenarios apply a 2 percent annual update or scale every rate to the corresponding RMCR, using factors of 4.011 and 2.217, respectively. These revenue-change approximations are separate from the exact aggregate predicted-revenue objective used in the constrained RMCR search.

10 Tables

Table 1

Compliance and payments: pooled and heterogeneous responses: 0.01 miles (about 16 metres)

	(1) Any payment	(2) Liability clearing	(3) Log tax paid
Panel A: Pooled specification			
Log(Baseline Circle Rate)	-0.316*** (0.071)	-0.233*** (0.076)	1.299*** (0.400)
Implied pooled elasticity	-0.452*** (0.102)	-0.599*** (0.197)	1.299*** (0.400)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	-0.234** (0.117)	-0.257* (0.131)	0.988 (0.598)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-1.139 (1.005)	0.346 (1.112)	4.529 (4.781)
F-test: $\theta = 0$	1.284	0.097	0.897
Panel C: Implied elasticities by tercile			
Low tercile	-0.077 (0.350)	-0.694 (0.652)	0.301 (1.264)
Middle tercile	-0.413*** (0.147)	-0.654** (0.279)	1.144** (0.472)
High tercile	-0.667*** (0.183)	-0.562 (0.393)	1.888*** (0.571)
Joint tercile F-test	12.781***	5.038***	0.897
Observations	18,738	18,738	9,036
Clusters (minimum)	143	143	130
Mean outcome (levels)	0.699	0.388	79.191
Low tercile mean (levels)	0.744	0.449	49.772
Middle tercile mean (levels)	0.648	0.378	115.421
High tercile mean (levels)	0.692	0.335	89.475
Pair-by-year and ward-division FE	Y	Y	Y
Payers only	N	N	Y

Notes: Any payment is an indicator for a positive tax payment. Liability clearing requires payment of at least the recorded liability. Log tax paid is restricted to payers; pair-year cells with only one payer are dropped as singletons. Payments are in USD and may include penalties. All columns include property-pair-by-year and ward-division fixed effects. The sample excludes entire property pairs with a geometric-mean baseline circle rate above USD 280 and is otherwise restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. The moderator is re-centered after this restriction. Tercile cutpoints are recomputed in the retained pair distribution (USD 186 and USD 235). They rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table 2

Balance between pairs of properties: pooled and heterogeneous responses: 0.01 miles (about 16 metres)

	(1) Log residual	(2) Floors	(3) Constr. type
Panel A: Pooled specification			
Log(Baseline Circle Rate)	-0.363 (0.360)	0.154 (0.133)	-0.020 (0.019)
Implied pooled elasticity	-0.363 (0.360)	0.089 (0.077)	-0.020 (0.019)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	-0.341 (0.669)	0.390 (0.280)	-0.019 (0.033)
Log(Baseline Circle Rate) $\times$ Centered within-pair mean log CR	-0.305 (5.851)	-3.284 (2.471)	-0.016 (0.336)
F-test: $\theta = 0$	0.003	1.766	0.002
Panel C: Implied elasticities by tercile			
Low tercile	-0.294 (1.531)	0.416 (0.306)	-0.016 (0.079)
Middle tercile	-0.350 (0.524)	0.155 (0.110)	-0.020 (0.025)
High tercile	-0.402 (0.664)	-0.223 (0.186)	-0.022 (0.044)
Joint tercile F-test	0.003	1.011	0.575
Observations	18,738	18,504	18,504
Clusters (minimum)	143	139	139
Mean outcome (levels)	0.941	1.727	1.015
Low tercile mean (levels)	0.920	2.156	1.029
Middle tercile mean (levels)	1.112	1.875	1.010
High tercile mean (levels)	0.830	1.194	1.006
Pair-by-year and ward-division FE	Y	Y	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. Log residual is the log of the liability component after accounting for the circle rate in Equation 9. Floors and construction type come from the matched property-characteristics data. Construction is coded 1 for cement, 2 for partial cement and 3 for less durable materials; its normalized slope depends on this coding and is not an economic elasticity. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table 3
Compliance and payments below the enforcement threshold: 0.01 miles (about 16 metres)

	(1) Any payment	(2) Liability clearing	(3) Log tax paid
Panel A: Pooled specification			
Log(Baseline Circle Rate)	-0.369*** (0.067)	-0.298*** (0.084)	1.658*** (0.356)
Implied pooled elasticity	-0.499*** (0.091)	-0.701*** (0.197)	1.658*** (0.356)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	-0.284** (0.117)	-0.325** (0.158)	1.546*** (0.519)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-1.142 (0.933)	0.361 (1.348)	1.584 (4.066)
F-test: $\theta = 0$	1.497	0.072	0.152
Panel C: Implied elasticities by tercile			
Low tercile	-0.141 (0.309)	-0.762 (0.701)	1.310 (1.057)
Middle tercile	-0.464*** (0.132)	-0.727*** (0.278)	1.603*** (0.417)
High tercile	-0.710*** (0.150)	-0.711* (0.419)	1.856*** (0.522)
Joint tercile F-test	20.874***	7.377***	0.152
Observations	14,884	14,884	8,106
Clusters (minimum)	133	133	122
Mean outcome (levels)	0.740	0.425	66.388
Low tercile mean (levels)	0.798	0.498	28.305
Middle tercile mean (levels)	0.696	0.430	100.882
High tercile mean (levels)	0.719	0.356	86.571
Pair-by-year and ward-division FE	Y	Y	Y
Payers only	N	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. The outcomes and payer restriction are defined as in Table 1. The sample retains observations coded as not enforced in the analysis: observations with recorded cumulative dues above INR 10,000 (about USD 117 at INR 85 per USD) are excluded. Missing cumulative dues are also coded as not enforced in the source. Cutpoints are inherited from the full matched sample at this distance. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

A Appendix

A.1 Descriptive Plots for the municipality

Since the relationship between circle rates and liabilities is key to my analysis, I plot this relationship in Figure A7, Panel C. Properties at higher circle rates have higher liabilities. I also show in Panels A and B that properties located in areas with higher circle rates are also more compliant and pay more tax. In Panel D, I show that compliance also increases as liabilities increase. These plots are purely descriptive as they are binned scatterplots, with each dot representing the mean of the respective bin.

A.2 Deriving the circle rate at which revenue is maximized

A.2.1 Using Average Effects

At the optimal circle rate, substituting the two elasticities derived from Equations 10 and 11 into Equation 5, we get:

$$\frac{\beta_1}{C^*} + \gamma_1 = 0,$$

where C^* denotes compliance at the revenue-maximizing circle rate.

Using the estimating equation for compliance, this can be written as:

$$-\frac{\beta_1}{\gamma_1} = \beta_0 + \beta_1 \log(\text{CircleRate}^*) + \bar{\delta},$$

where $\bar{\delta}$ denotes the average pair-year fixed effect.

Since

$$\bar{C} = \beta_0 + \beta_1 \overline{\log(\text{CircleRate})} + \bar{\delta},$$

we can recover the effective intercept as

$$\beta_0 + \bar{\delta} = \bar{C} - \beta_1 \overline{\log(\text{CircleRate})}.$$

Then,

$$-\frac{\beta_1}{\gamma_1} = \bar{C} - \beta_1 \overline{\log(\text{CircleRate})} + \beta_1 \log(\text{CircleRate}^*).$$

Rearranging this, we get

$$\log(\text{CircleRate}^*) = \overline{\log(\text{CircleRate})} + \frac{-\frac{\beta_1}{\gamma_1} - \bar{C}}{\beta_1}.$$

The implied revenue-maximizing circle rate is therefore

$$\text{CircleRate}^* = \exp \left[\overline{\log(\text{CircleRate})} + \frac{-\frac{\beta_1}{\gamma_1} - \bar{C}}{\beta_1} \right].$$

A.2.2 Using heterogeneous effects

Let M_i denote the within-pair mean log baseline circle rate, centered at the matched-sample mean log baseline rate. The same baseline moderator is used for compliance and conditional payments and remains fixed when rates are scaled. The

fitted response slopes are

$$b_{Ci} = \beta_c + \theta_c M_i, \quad b_{Yi} = \beta_y + \theta_y M_i. \quad (23)$$

I form terciles of the within-pair mean log baseline circle-rate distribution before restricting to an outcome's estimation sample, counting each distinct pair once when setting the cutpoints and assigning both members to the same group. Let $g(i)$ denote the resulting tercile and let $\bar{C}_{g(i),0}$ be its mean compliance in the full compliance estimation sample, including nonpayers. The counterfactual starts each observation at its tercile's compliance mean. The common baseline conditional-payment level, $\bar{A}_0^+$, is the overall mean payment among observations in the payer regression sample.

A common factor k multiplies every baseline circle rate:

$$cr_i(k) = k cr_{i0}, \quad \log cr_i(k) - \log cr_{i0} = \log k. \quad (24)$$

Holding the baseline moderator fixed, predicted mean compliance is

$$C_i(k) = \bar{C}_{g(i),0} + b_{Ci} \log k. \quad (25)$$

The compliance semi-elasticity is held constant within each baseline type, but its elasticity changes with predicted compliance:

$$\epsilon_{Ci}(k) = \frac{b_{Ci}}{C_i(k)}. \quad (26)$$

For payments conditional on paying, the fitted log-payment response implies

$$A_i^+(k) = \bar{A}_0^+ k^{b_{Yi}}, \quad \epsilon_{Ai}(k) = b_{Yi}. \quad (27)$$

These equations extrapolate the fitted semi-elasticity of compliance and elasticity of conditional payments beyond baseline rates. They do not hold the baseline compliance elasticity fixed.

Expected revenue from observation i is $R_i(k) = C_i(k) A_i^+(k)$. Summing across observations gives

$$R(k) = \sum_i \left[\bar{C}_{g(i),0} + b_{Ci} \log k \right] \bar{A}_0^+ k^{b_{Yi}}. \quad (28)$$

Define the feasible grid

$$\mathcal{K} = \left\{ \exp \left[\log(0.5) + \frac{j}{5000} \log(20) \right] : j = 0, \dots, 5000; 0 < C_i(k) < 1 \text{ for every } i \right\}. \quad (29)$$

The numerical exercise chooses

$$k^* = \arg \max_{k \in \mathcal{K}} R(k). \quad (30)$$

Thus the linear probability model is extrapolated only over common scaling factors that keep every type's predicted compliance strictly between zero and one. Predictions are not clipped at the boundaries.

For an interior optimum of the continuous revenue function, differentiation with respect to $\log k$ gives

$$\frac{dR(k)}{d \log k} = \sum_i A_i^+(k) [b_{Ci} + C_i(k) b_{Yi}] = 0. \quad (31)$$

Equivalently, the revenue-weighted sum of the two elasticities is zero:

$$\sum_i \frac{R_i(k)}{R(k)} [\epsilon_{Ci}(k) + \epsilon_{Ai}(k)] = 0. \quad (32)$$

This condition applies to an interior optimum. The constrained grid solution can occur at the feasibility boundary while aggregate predicted revenue is still increasing; in that case, the zero-sum condition need not hold. Because the moderator is fixed, the calculation does not involve a cubic equation in the counterfactual log circle rate.

The common factor preserves relative baseline circle rates. I summarize the resulting schedule by its geometric mean:

$$CR_{\text{het}}^* = k^* \exp(\overline{\log cr_0}). \quad (33)$$

For bootstrap inference, I resample complete property pairs, reconstruct the moderator and pair-level terciles in each draw, recompute the tercile compliance means, and re-estimate both response functions. Repeated copies of a resampled pair receive distinct pair identifiers for fixed effects and for counting pairs when setting cutpoints. Pooled and heterogeneous optima are calculated on the same resampled data, and their difference is formed within each draw.

A.3 Deriving the total change in revenue due to a change in rates

To calculate the change in revenue from a 1 percent increase in circle rates, start from

$$R(cr) = C(cr) \times A^+(cr), \quad (34)$$

where $C(cr)$ is compliance and $A^+(cr)$ is the amount paid by compliers.

Differentiating revenue with respect to the circle rate gives

$$\frac{dR(cr)}{dcr} = C'(cr)A^+(cr) + C(cr)A^{+'}(cr). \quad (35)$$

Since the compliance equation is estimated with $\log(cr)$, the compliance semi-elasticity is

$$\frac{dC(cr)}{d \log(cr)} = \beta_1. \quad (36)$$

Using the fact that

$$\frac{d \log(cr)}{dcr} = \frac{1}{cr},$$

we get

$$C'(cr) = \frac{dC(cr)}{dcr} = \frac{\beta_1}{cr}. \quad (37)$$

Similarly, since the payment equation among compliers is log-log,

$$\frac{d \log A^+(cr)}{d \log(cr)} = \gamma_1. \quad (38)$$

Therefore,

$$A^{+'}(cr) = \frac{dA^+(cr)}{dcr} = \frac{\gamma_1 A^+(cr)}{cr}. \quad (39)$$

Substituting these terms into the revenue derivative gives

$$\frac{dR(cr)}{dcr} = \frac{\beta_1}{cr} A^+(cr) + C(cr) \frac{\gamma_1 A^+(cr)}{cr}. \quad (40)$$

Factoring out $\frac{A^+(cr)}{cr}$, this becomes

$$\frac{dR(cr)}{dcr} = \frac{A^+(cr)}{cr} [\beta_1 + \gamma_1 C(cr)]. \quad (41)$$

For a 1 percent increase in circle rates,

$$\Delta cr = 0.01cr. \quad (42)$$

Using a first-order approximation,

$$\Delta R \approx \frac{dR(cr)}{dcr} \times 0.01cr. \quad (43)$$

Substituting the expression for $\frac{dR(cr)}{dcr}$,

$$\Delta R \approx 0.01 A^+(cr) [\beta_1 + \gamma_1 C(cr)]. \quad (44)$$

Evaluated at the sample mean, this becomes

$$\Delta R \approx 0.01 \bar{A}^+ [\beta_1 + \gamma_1 \bar{C}], \quad (45)$$

where $\bar{A}^+$ is mean tax paid among compliers and $\bar{C}$ is mean compliance.

A.4 Was the circle rate in 2011 driven by political economy effects?

There could be a concern that the circle rate in 2011 is driven by political factors.

In Table A2, I present results from a multivariate regression with electoral-ward level characteristics as independent variables and the circle-level circle rate in 2011 as the dependent variable. There are 77 electoral wards in the municipality, so each ward could contain multiple circle-rates, however, due to an absence of covariates at the parcel-level, I use aggregate electoral-ward estimates. These estimates should then be interpreted as ward-level averages. No concerning pattern emerges from this estimation.

One concern could be that a politically powerful elected representative helps set higher rates for the circle. I use ward-level elections from 2009, to construct a ‘party alignment’ variable, which is a dummy equal to 1 if the elected representative in 2009 belonged to the same party as the elected party at the state-level (the Indian National Congress). Party Alignment does not seem to drive the circle rate as the coefficient is indistinguishable from 0.

A.5 Regression Discontinuity

An alternate specification that could be used to measure the effect on compliance due to a higher circle rate could be given by:

$$Y_i = \alpha + \tau \text{High}_i + \beta_1 D_i + \beta_2 (\text{High}_i \times D_i) + \varepsilon_i, \quad (46)$$

where D_i is the running variable, which is the distance of property ‘i’ to a *shared border*. High_i is a binary variable that equals 1 when property ‘i’ lies in the higher circle rate side of the border. τ is then the coefficient of interest, and can be

interpreted as the effect on the dependent variable, due to being in a higher circle-rate circle. I estimate this specification by clustering standard errors at the circle-level and using triangular kernels.

In order to estimate this, I have to now structure my data differently. Instead of matching each property with a neighbouring property, I now form coordinate pairs. I restrict pairs to those that lie in neighbouring circles with varying circle rates. Within a coordinate pair, I mark all properties that are in the higher circle-rate circle as being *High_i*, and consequently use the distance of the coordinate to the shared border as the running variable. There are thus a few notable differences in this data structure relative to my baseline. One, the distance measure is now not a distance to the paired property but rather a distance to the shared border and two, since each property-year only appears in my data once, this reduces the power of my estimating sample.

In Figure A11, I present RD estimates on compliance, plotted using linear polynomials on both sides of the running variable. I restrict the distance bandwidth to 0.018 miles, 0.024 miles, 0.031 miles and 0.049 miles in Panels (a)–(d), and a similar pattern as the baseline emerges. Compliance reduces in the high circle-rate circles, and the effect disappears as the sample bandwidth increases to 0.05 miles. Since I use triangular weights, properties that lie closer to the border have higher weights in the estimation than properties that lie farther. From the plots, it seems likely the effect is driven by these very closely located properties. In Figure A12, I present both the conventional and bias-corrected point estimates at different bandwidths ranging from 16 metres to 0.050 miles.

In Table A6, I present the resultant point estimates for compliance in Panel A, taxes paid in Panel B and the residual term in Panel C. The large drop in the number of observations and the resultant lack of power lead to most estimates being not statistically different from 0 at the 95% level of confidence. However, the signs and point estimates of all coefficients are in line with my baseline estimates. Properties at the higher circle-rate side of the discontinuity have 0.18 pp and 0.48 pp lower compliance using 0.012 miles and 0.018 miles as bandwidths. They simultaneously pay 1.30% and 2.311% more tax.

A.6 Robustness

A.6.1 Considering the dynamics

House Price Growth A mechanism at play that would explain these results would be house price growth. If house prices in higher 2011 circle-rate areas grew faster than their neighbouring house prices, then this wealth effect could affect both the fall in compliance and the amount paid by compliers. For instance, it is possible that compliers that see higher house price growth pay all their penalties or always pay the full amount of the liability.

While I don't have direct house price data, I use contemporaneous circle rates as proxies (Anagol et al., 2024). I add circle-rate growth and its interaction with the baseline rate to the main specification in Appendix Table A7, columns (1)–(2). Panel A uses all available years and Panel B excludes 2023–2024, when the relationship between baseline and contemporaneous circle rates changes, as shown in Figure A17. The compliance coefficient remains close to the pooled estimate in both panels.

The payment estimates vary more with the growth controls. Column (3) also suggests that areas with higher baseline circle rates did not grow faster: a 1% higher baseline rate is associated with a 0.52% higher contemporaneous rate over 2015–2024, and 0.98% over 2015–2022. Taken together, these results do not suggest that faster house price growth in high-baseline-rate areas explains the main compliance response.

Construction I also check whether more properties enter the tax roll in lower baseline circle-rate areas, as an avoidance response might suggest. Appendix Table [A7](#), column (4), uses the number of new properties added to the roll within a circle area as the construction measure. The association with baseline circle rates is positive, rather than in the direction this avoidance response would suggest. My main estimates continue to use a balanced panel to exclude properties entering later.

A.6.2 The effect of COVID on compliance

I re-estimate the within-pair specification separately before 2020, during 2020 and after 2020 to isolate the pandemic period. Figure [6](#) presents these estimates; Panel B uses the preferred 16-metre threshold. Appendix Figure [A13](#) presents the year-by-year version.

The negative compliance response is substantially larger in 2020 and returns toward its earlier level afterwards. This is consistent with the lockdown tightening income constraints, in line with [Brockmeyer et al. \(2021\)](#) and [Bergeron et al. \(2024\)](#), though these comparisons do not separately identify the lockdown mechanism.

B Appendix Tables and Figures

B.1 Tables

Table A1
Multipliers by Property Characteristics

Item	Weight
Category	
Bungalow; penthouse; row house; club house / other facilities used by members of a co-operative housing society	1.50
Two-storeyed apartments	1.25
Room or apartment and similar; swimming tank	1.00
Attached garage	0.50
Society office	0.30
Car parking in stilt / basement / podium	0.25
Type	
Superior RCC building	1.50
Other RCC building	1.00
Pucca construction other than RCC	0.70
Semi-permanent building	0.40
Kuchha building	0.30
Age	
Residential buildings after 01.04.2008	1.00
Non-residential buildings after 01.04.2008	1.50
Residential buildings after 1992 and before 01.04.2008	0.75
Non-residential buildings after 1992 and before 01.04.2008	1.25
Residential buildings before 1992	0.50
Non-residential buildings before 1992	1.10
Height	
Basement excluding car parking; lower first floor; upper first floor; ground floor; 1st to 4th floor	1.00
5th to 10th floor	1.05
11th to 20th floor	1.10

Notes: This table reports the multipliers used in Kolhapur's formula-based property valuation system. Multipliers vary by property category, construction type, building age, and floor height. These weights enter the assessed valuation formula together with the circle rate and property area. Source: Kolhapur Municipal Corporation property tax rules, available at <https://web.kolhapurcorporation.gov.in/departments?deptid=2>.

Table A2
Predictors of the circle rate

	(1) Baseline circle rate
Number of Households	4.106 (2.964)
Owner occupied properties	3.181 (2.818)
Commercial properties	-15.19 (20.23)
Institutional properties	1.730 (31.35)
Total population	-1.089 (1.064)
Population age 0–6	-18.16*** (4.902)
Illiterate population	6.851*** (1.765)
Main cultivators	-18.40** (7.248)
Main agricultural labourers	7.089 (9.272)
Total workers	-1.715 (3.583)
Main workers	2.249 (2.937)
Main household industry workers	8.122** (3.668)
Scheduled Tribe population	3.871 (8.410)
Total population	0.0389 (0.274)
Scheduled Caste population	-0.983** (0.465)
Party Alignment	-173.4 (339.1)
Constant	20275.2*** (1919.5)
Observations	702
R-squared	0.159

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table presents estimates from a multivariate regression using the baseline 2011 circle rate at the circle-level as the dependent variable and electoral ward characteristics as independent variables. Each electoral ward contains several circles. This specification includes controls for administrative wards, which are bigger and contain several electoral wards. *Households* is the number of households, *Owner occupied*, *Commercial*, and *Institutional* are the number of properties that are owner-occupied, commercial and institutional, respectively. Other covariates include total population, population ages 0–6, illiterate population, total workers, main workers, main cultivators, main agricultural labourers, main household industry workers, Scheduled Tribe population and Scheduled Caste population. *Party Alignment* is a dummy variable equal to one if the ward-level representative in 2009 belonged to the same party as the majority party in the state.

Table A3
Alternative payment specifications: 0.01 miles (about 16 metres)

	(1) Log tax paid	(2) Tax paid (PPML)
Panel A: Pooled specification		
Log(Baseline Circle Rate)	1.200*** (0.323)	0.611* (0.346)
Implied pooled elasticity	1.200*** (0.323)	0.611* (0.346)
Panel B: Heterogeneous specification		
Log(Baseline Circle Rate)	0.906* (0.471)	-0.042 (0.638)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	4.197 (3.564)	8.968 (6.410)
F-test: $\theta = 0$	1.387	1.957
Panel C: Implied elasticities by tercile		
Low tercile	0.203 (1.023)	-1.575 (1.691)
Middle tercile	1.038*** (0.384)	0.218 (0.480)
High tercile	1.743*** (0.413)	1.750** (0.766)
Joint tercile F-test	1.387	1.957
Observations	17,599	24,845
Clusters (minimum)	149	160
Mean outcome (levels)	77.885	55.351
Low tercile mean (levels)	65.285	47.774
Middle tercile mean (levels)	80.064	55.518
High tercile mean (levels)	93.177	65.350
Pair, year and ward-division FE	Y	Y
Payers only	Y	N
Poisson	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. The log-tax-paid outcome uses positive payments; the PPML outcome includes zero payments. These specifications use separate property-pair, financial-year and ward-division fixed effects, allowing pair-years with only one payer to contribute to the log-payment regression. Payments are in USD. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A4
The share of tax liability paid: 0.01 miles (about 16 metres)

	(1) Share paid (OLS)	(2) Share paid (PPML)
Panel A: Pooled specification		
Log(Baseline Circle Rate)	4.733*** (0.585)	1.509*** (0.145)
Implied pooled elasticity	2.254*** (0.279)	1.509*** (0.145)
Panel B: Heterogeneous specification		
Log(Baseline Circle Rate)	3.156*** (0.472)	1.021*** (0.140)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	22.981*** (5.188)	6.537*** (1.473)
F-test: $\theta = 0$	19.621***	19.700***
Panel C: Implied elasticities by tercile		
Low tercile	-0.238 (0.834)	0.006 (0.328)
Middle tercile	1.211*** (0.116)	1.211*** (0.122)
High tercile	3.568*** (0.369)	2.327*** (0.245)
Joint tercile F-test	92.958***	19.700***
Observations	9,036	17,158
Clusters (minimum)	130	143
Mean outcome (levels)	2.100	1.491
Low tercile mean (levels)	1.386	1.050
Middle tercile mean (levels)	3.257	2.062
High tercile mean (levels)	2.165	1.527
Pair-by-year and ward-division FE	Y	Y
Payers only	Y	N
Poisson	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. Share paid is tax paid divided by recorded liability and can exceed one because payments include penalties. The OLS specification is restricted to payers; PPML includes non-payers, with missing shares set to zero as in the analysis. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A5

Tax paid in levels, including zero payments: 0.01 miles (about 16 metres)

	(1) Tax paid: payers	(2) Tax paid: all (OLS)	(3) Tax paid: all (PPML)
Panel A: Pooled specification			
Log(Baseline Circle Rate)	69.698 (47.035)	26.466 (27.845)	0.550 (0.337)
Implied pooled elasticity	0.880 (0.594)	0.489 (0.514)	0.550 (0.337)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	2.356 (83.179)	-11.410 (58.782)	-0.037 (0.634)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	981.539 (759.080)	526.694 (549.638)	8.046 (6.354)
F-test: $\theta = 0$	1.672	0.918	1.604
Panel C: Implied elasticities by tercile			
Low tercile	-2.945 (3.931)	-2.536 (3.873)	-1.287 (1.581)
Middle tercile	0.312 (0.513)	0.059 (0.637)	0.196 (0.475)
High tercile	2.207** (0.864)	1.554 (0.946)	1.570** (0.755)
Joint tercile F-test	8.036***	3.034*	1.604
Observations	9,036	18,738	17,158
Clusters (minimum)	130	143	143
Mean outcome (levels)	79.191	54.122	59.105
Low tercile mean (levels)	49.772	36.778	39.048
Middle tercile mean (levels)	115.421	68.756	77.515
High tercile mean (levels)	89.475	60.413	66.299
Pair-by-year and ward-division FE	Y	Y	Y
Payers only	Y	N	N
Poisson	N	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. The first outcome is tax paid among payers; the second and third use tax paid including zeros, estimated by OLS and PPML respectively. Payments are in USD and may include penalties. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A6
Compliance, amount paid and balance on residual: Regression Discontinuity

Panel A: Compliance

	(1)	(2)	(3)	(4)
	0.012	0.018	0.031	0.049
	Compliance - Extensive Margin			
High	-0.18 (0.46)	-0.48 (0.47)	-0.11 (0.33)	0.01 (0.23)
Obs	178	499	1645	2746
Control Mean	0.612	0.752	0.693	0.652
Shared Boundary FE	N	N	N	N

Panel B: Amount of Tax Paid

	(1)	(2)	(3)	(4)
	0.012	0.018	0.031	0.049
	Ln(Tax Paid) - Intensive Margin			
High	1.306 (1.241)	2.311*** (0.487)	1.263 (1.195)	0.580 (1.046)
Obs	109	375	1140	1790
Control Mean	148.560	64.192	56.005	70.493
Shared Boundary FE	N	N	N	N

Panel C: Residual

	(1)	(2)	(3)	(4)
	0.012	0.018	0.031	0.049
	Ln(Residual)			
High	0.63 (0.74)	1.37 (0.80)	1.20 (0.91)	0.73 (0.92)
Obs	178	499	1645	2746
Control Mean	138.468	72.513	59.238	71.284
Shared Boundary FE	N	N	N	N

Notes: This table presents regression discontinuity results for the extensive margin measure of compliance, the intensive margin measure of the amount of taxes paid and the residual term, across different bandwidths from columns (1) – (4). Each coefficient comes from estimating the following equation

$$\text{Compliance}_i = \alpha + \tau \text{High}_i + \beta_1 D_i + \beta_2 (\text{High}_i \times D_i) + \varepsilon_i,$$

where High is equal to 1 if a coordinate has the higher circle rate relative to its matched coordinate. D_i is the running variable, which is distance to the circle rate border. I present only τ and use triangular distance weights. These results are then equivalent to using Stata's 'rdrobust'.

Table A7
Circle Rate Growth and Construction

Panel A: Sample: 16 metres

	(1) Ln(Tax Paid)	(2) Compliance	(3) Ln(Circle Rate)	(4) Construction
Ln(Baseline Circle Rate)	0.983** (0.477)	-0.334*** (0.0893)	0.523*** (0.0538)	3.227*** (0.415)
Circle Rate Growth	0.270*** (0.0962)	0.0734*** (0.0222)		
Ln(Baseline Circle Rate) × Circle Rate Growth	-0.0496*** (0.0167)	-0.0132*** (0.00395)		
Obs	5146	10080	12446	13766
Control Mean	3.353	0.719	6.010	0.614
Pair × Year, Ward FE	Y	Y	Y	Y

Panel B: Sample: 16 metres and Year less than 2023

	(1) Ln(Tax Paid)	(2) Compliance	(3) Ln(Circle Rate)	(4) Construction
Ln(Baseline Circle Rate)	1.259** (0.508)	-0.364*** (0.0964)	0.981*** (0.0692)	1.507*** (0.205)
Circle Rate Growth	0.00136 (0.267)	-0.0873* (0.0452)		
Ln(Baseline Circle Rate) × Circle Rate Growth	-0.00193 (0.0513)	0.0177** (0.00844)		
Obs	3208	7066	9432	15700
Control Mean	3.205	0.681	6.073	0.290
Pair × Year, Ward FE	Y	Y	Y	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280. This table examines whether the main effects are explained by subsequent growth in circle rates or by new construction. Columns 1 and 2 use log tax paid and compliance as the main dependent variables and include circle-rate growth and its interaction with the baseline 2011 circle rate. Column 3 tests the relationship between baseline 2011 circle rates and contemporaneous circle rates. Column 4 uses new construction, measured as the number of new properties added to the tax roll in a circle area as the outcome. Panel A uses the 16-metre matched sample with all available years. Panel B restricts to years before 2023. Circle rate data starts from 2015, so all tables are restricted to 2016–2024. All specifications include property-pair-by-year and ward fixed effects. Standard errors are clustered two-way at the circle-rate and coordinate-circle-rate levels. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8
Robustness

	(1) 8 metres	(2) 16 metres
Panel A: Baseline		
Ln(Baseline Circle Rate)	-0.308*** (0.072)	-0.316*** (0.071)
Observations	10322	18738
Control Mean	0.701	0.699
Elasticity	-0.439	-0.452
Panel B: With inverse distinct pair-count weights		
Ln(Baseline Circle Rate)	-0.298*** (0.056)	-0.238 (0.167)
Observations	10322	18738
Control Mean	0.701	0.699
Elasticity	-0.425	-0.341
Panel C: Unbalanced panel		
Ln(Baseline Circle Rate)	-0.362*** (0.073)	-0.341*** (0.069)
Observations	11812	23368
Control Mean	0.691	0.696
Elasticity	-0.523	-0.491
Panel D: Property-characteristic fixed effects		
Ln(Baseline Circle Rate)	-0.305*** (0.072)	-0.290*** (0.074)
Observations	10166	18506
Control Mean	0.706	0.702
Elasticity	-0.432	-0.413

Notes: I exclude pairs with mean baseline circle rates above USD 280. This table reports robustness checks for extensive-margin compliance. Columns 1 and 2 use 0.005 miles (about 8 metres) and 0.01 miles (about 16 metres). Panel A is the baseline. Panel B weights each observation by the inverse of the number of distinct property pairs in which that property appears in that financial year, so properties matched more often receive less weight. Panel C uses an unbalanced panel. Panel D adds fixed effects for the number of floors and construction type, alongside pair-by-year and ward fixed effects. These are property-characteristic fixed effects. The other panels use pair-by-year and ward-division fixed effects. Standard errors cluster two ways by circle area and coordinate-group-by-circle area. Significance levels: * p<0.10, ** p<0.05, *** p<0.01.

Table A9

Compliance and payments: pooled and heterogeneous responses: 0.005 miles (about 8 metres)

	(1) Any payment	(2) Liability clearing	(3) Log tax paid
Panel A: Pooled specification			
Log(Baseline Circle Rate)	-0.308*** (0.072)	-0.235*** (0.088)	1.688*** (0.379)
Implied pooled elasticity	-0.439*** (0.102)	-0.593*** (0.222)	1.688*** (0.379)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	-0.247*** (0.076)	-0.254** (0.108)	1.806*** (0.422)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-2.707*** (0.926)	0.868 (1.778)	-5.191 (5.485)
F-test: $\theta = 0$	8.549***	0.239	0.896
Panel C: Implied elasticities by tercile			
Low tercile	0.101 (0.214)	-0.691 (0.573)	2.404** (0.927)
Middle tercile	-0.534*** (0.116)	-0.622*** (0.223)	1.618*** (0.359)
High tercile	-0.876*** (0.171)	-0.402 (0.567)	1.102* (0.632)
Joint tercile F-test	15.625***	3.940**	0.896
Observations	10,322	10,322	5,036
Clusters (minimum)	61	61	58
Mean outcome (levels)	0.701	0.395	87.966
Low tercile mean (levels)	0.727	0.516	52.671
Middle tercile mean (levels)	0.604	0.370	184.678
High tercile mean (levels)	0.709	0.333	92.564
Pair-by-year and ward-division FE	Y	Y	Y
Payers only	N	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. Any payment is an indicator for a positive tax payment. Liability clearing requires payment of at least the recorded liability. Log tax paid is restricted to payers; pair-year cells with only one payer are dropped as singletons. Payments are in USD and may include penalties. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A10

Balance between pairs of properties: pooled and heterogeneous responses: 0.005 miles (about 8 metres)

	(1) Log residual	(2) Floors	(3) Constr. type
Panel A: Pooled specification			
Log(Baseline Circle Rate)	-0.012 (0.373)	-0.013 (0.016)	-0.000140 (0.000173)
Implied pooled elasticity	-0.012 (0.373)	-0.010 (0.013)	-0.000139 (0.000172)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	0.362 (0.388)	-0.039 (0.048)	-0.000247 (0.000299)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-16.691*** (5.814)	1.167 (1.434)	0.005 (0.006)
F-test: $\theta = 0$	8.243***	0.662	0.705
Panel C: Implied elasticities by tercile			
Low tercile	2.336** (0.949)	-0.091 (0.112)	-0.001 (0.001)
Middle tercile	-0.104 (0.334)	-0.006 (0.012)	-0.000113 (0.000143)
High tercile	-1.948*** (0.689)	0.121 (0.149)	0.000414 (0.000493)
Joint tercile F-test	8.243***	0.331	0.353
Observations	10,322	10,164	10,164
Clusters (minimum)	61	60	60
Mean outcome (levels)	0.763	1.267	1.005
Low tercile mean (levels)	0.685	1.852	1.017
Middle tercile mean (levels)	0.586	1.000	1.000
High tercile mean (levels)	0.849	1.013	1.000
Pair-by-year and ward-division FE	Y	Y	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. Log residual is the log of the liability component after accounting for the circle rate in Equation 9. Floors and construction type come from the matched property-characteristics data. Construction is coded 1 for cement, 2 for partial cement and 3 for less durable materials; its normalized slope depends on this coding and is not an economic elasticity. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A11

Compliance and payments below the enforcement threshold: 0.005 miles (about 8 metres)

	(1) Any payment	(2) Liability clearing	(3) Log tax paid
Panel A: Pooled specification			
Log(Baseline Circle Rate)	-0.323*** (0.068)	-0.300*** (0.090)	1.921*** (0.368)
Implied pooled elasticity	-0.443*** (0.093)	-0.706*** (0.211)	1.921*** (0.368)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	-0.251*** (0.071)	-0.315*** (0.109)	2.054*** (0.395)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-3.341*** (0.858)	0.715 (1.924)	-5.721 (5.472)
F-test: $\theta = 0$	15.167***	0.138	1.093
Panel C: Implied elasticities by tercile			
Low tercile	0.144 (0.183)	-0.694 (0.516)	2.724*** (0.895)
Middle tercile	-0.538*** (0.100)	-0.751*** (0.215)	1.847*** (0.351)
High tercile	-0.968*** (0.146)	-0.615 (0.592)	1.295* (0.647)
Joint tercile F-test	25.032***	6.206***	1.093
Observations	8,838	8,838	4,670
Clusters (minimum)	56	56	56
Mean outcome (levels)	0.728	0.424	84.006
Low tercile mean (levels)	0.761	0.565	45.226
Middle tercile mean (levels)	0.638	0.393	176.927
High tercile mean (levels)	0.732	0.353	90.497
Pair-by-year and ward-division FE	Y	Y	Y
Payers only	N	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. The outcomes and payer restriction are defined as in Table 1. The sample retains observations coded as not enforced in the analysis: observations with recorded cumulative dues above INR 10,000 (about USD 117 at INR 85 per USD) are excluded. Missing cumulative dues are also coded as not enforced in the source. Cutpoints are inherited from the full matched sample at this distance. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A12
Alternative payment specifications: 0.005 miles (about 8 metres)

	(1) Log tax paid	(2) Tax paid (PPML)
Panel A: Pooled specification		
Log(Baseline Circle Rate)	1.601*** (0.258)	1.113*** (0.323)
Implied pooled elasticity	1.601*** (0.258)	1.113*** (0.323)
Panel B: Heterogeneous specification		
Log(Baseline Circle Rate)	1.761*** (0.270)	1.325*** (0.475)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-7.055* (3.781)	-7.475 (6.396)
F-test: $\theta = 0$	3.482*	1.366
Panel C: Implied elasticities by tercile		
Low tercile	2.875*** (0.741)	2.547* (1.416)
Middle tercile	1.463*** (0.258)	1.042*** (0.354)
High tercile	0.797 (0.486)	0.295 (0.652)
Joint tercile F-test	3.482*	1.366
Observations	10,030	13,990
Clusters (minimum)	71	74
Mean outcome (levels)	75.306	54.183
Low tercile mean (levels)	52.345	38.151
Middle tercile mean (levels)	82.758	58.489
High tercile mean (levels)	93.251	66.993
Pair, year and ward-division FE	Y	Y
Payers only	Y	N
Poisson	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. The log-tax-paid outcome uses positive payments; the PPML outcome includes zero payments. These specifications use separate property-pair, financial-year and ward-division fixed effects, allowing pair-years with only one payer to contribute to the log-payment regression. Payments are in USD. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A13
The share of tax liability paid: 0.005 miles (about 8 metres)

	(1) Share paid (OLS)	(2) Share paid (PPML)
Panel A: Pooled specification		
Log(Baseline Circle Rate)	5.355*** (0.489)	1.667*** (0.149)
Implied pooled elasticity	2.038*** (0.186)	1.667*** (0.149)
Panel B: Heterogeneous specification		
Log(Baseline Circle Rate)	4.524*** (0.441)	1.488*** (0.123)
Log(Baseline Circle Rate) $\times$ Centered within-pair mean log CR	36.631*** (6.400)	9.650*** (1.797)
F-test: $\theta = 0$	32.756***	28.849***
Panel C: Implied elasticities by tercile		
Low tercile	0.139 (0.516)	0.327 (0.296)
Middle tercile	0.847*** (0.043)	1.781*** (0.108)
High tercile	4.357*** (0.271)	2.822*** (0.217)
Joint tercile F-test	295.596***	28.849***
Observations	5,036	9,440
Clusters (minimum)	58	61
Mean outcome (levels)	2.627	1.842
Low tercile mean (levels)	2.162	1.749
Middle tercile mean (levels)	6.906	3.350
High tercile mean (levels)	2.178	1.567
Pair-by-year and ward-division FE	Y	Y
Payers only	Y	N
Poisson	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. Share paid is tax paid divided by recorded liability and can exceed one because payments include penalties. The OLS specification is restricted to payers; PPML includes non-payers, with missing shares set to zero as in the analysis. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

Table A14

Tax paid in levels, including zero payments: 0.005 miles (about 8 metres)

	(1) Tax paid: payers	(2) Tax paid: all (OLS)	(3) Tax paid: all (PPML)
Panel A: Pooled specification			
Log(Baseline Circle Rate)	129.763*** (20.990)	57.458*** (14.942)	1.051*** (0.322)
Implied pooled elasticity	1.475*** (0.239)	0.986*** (0.256)	1.051*** (0.322)
Panel B: Heterogeneous specification			
Log(Baseline Circle Rate)	131.309*** (20.614)	67.138*** (15.866)	1.267*** (0.470)
Log(Baseline Circle Rate) × Centered within-pair mean log CR	-68.181 (647.306)	-431.655 (463.483)	-7.712 (6.286)
F-test: $\theta = 0$	0.011	0.867	1.505
Panel C: Implied elasticities by tercile			
Low tercile	2.642 (1.585)	2.791* (1.477)	2.195* (1.130)
Middle tercile	0.698*** (0.140)	0.715*** (0.211)	1.033*** (0.368)
High tercile	1.319 (0.902)	0.118 (0.953)	0.201 (0.649)
Joint tercile F-test	22.316***	9.214***	1.505
Observations	5,036	10,322	9,440
Clusters (minimum)	58	61	61
Mean outcome (levels)	87.966	58.260	63.704
Low tercile mean (levels)	52.671	42.351	45.322
Middle tercile mean (levels)	184.678	76.996	90.702
High tercile mean (levels)	92.564	62.927	68.494
Pair-by-year and ward-division FE	Y	Y	Y
Payers only	Y	N	N
Poisson	N	N	Y

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the terciles. The first outcome is tax paid among payers; the second and third use tax paid including zeros, estimated by OLS and PPML respectively. Payments are in USD and may include penalties. All columns include property-pair-by-year and ward-division fixed effects. The sample is restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025, matched across different baseline circle rates. Panel A reports the pooled coefficient on log baseline circle rate and its implied elasticity. Panel B reports the coefficients on Log(Baseline Circle Rate) and its interaction with centered within-pair mean log CR. CR denotes baseline circle rate. The moderator is the within-pair mean of log baseline circle rate minus the matched-sample mean of log baseline circle rate. The log-rate coefficient in Panel B is the slope when this moderator equals zero. Its main effect is absorbed by the fixed effects. The Panel B F-test is the one-restriction Wald test of $\theta = 0$, where θ is the interaction coefficient. Panel C reports elasticities from the continuous interaction model evaluated at the tercile-specific mean moderator in the outcome's estimation sample, rather than from separate tercile regressions. Tercile cutpoints rank distinct property pairs by their mean log baseline circle rate, count each pair once, assign both pair members to the same group and remain fixed across outcomes and sample restrictions at each distance. Ties can leave unequal group sizes. Level-outcome slopes are divided by the overall outcome mean in Panel A and by the full corresponding tercile outcome mean in Panel C; log-outcome and PPML slopes are already elasticities. Means for log outcomes are reported in levels. The joint tercile F-test compares the three displayed elasticities, accounting for their covariance. For level outcomes with different tercile means, this tests a different hypothesis from $\theta = 0$. Test degrees of freedom use the restriction rank and the smaller cluster count minus one; PPML Wald chi-squared statistics are divided by their restriction rank to report F statistics. Means and cutpoints are held fixed in inference. Standard errors in parentheses cluster two ways by circle area and coordinate-group-by-circle area. Stars on coefficients, elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$. Dashes denote inapplicable or non-estimable entries.

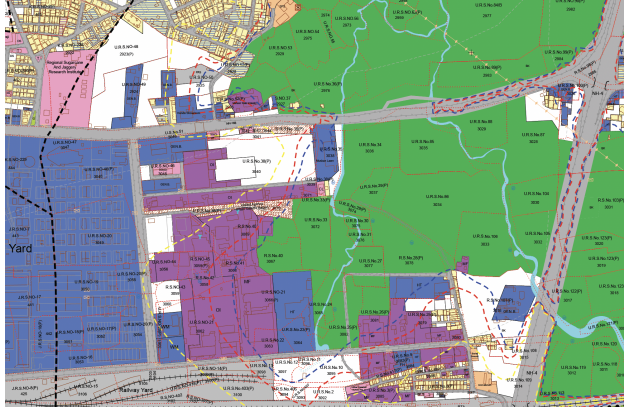
Table A15
Compliance and tax paid: median and tercile heterogeneity

	(1) Compliance elasticity	(2) Tax-paid elasticity: payers
Panel A: Median split		
Low half	-0.506 (0.313)	0.351 (1.251)
High half	-0.451*** (0.109)	1.447*** (0.352)
Joint equality F-test	0.027	0.774
Observations	18,738	9,036
Circle-area clusters	143	130
Panel B: Terciles		
Low tercile	-0.347 (0.442)	-0.619 (1.483)
Middle tercile	-0.341** (0.140)	2.004*** (0.514)
High tercile	-0.676*** (0.145)	0.850** (0.400)
Joint equality F-test	1.457	2.428*
Observations	18,738	9,036
Circle-area clusters	143	130

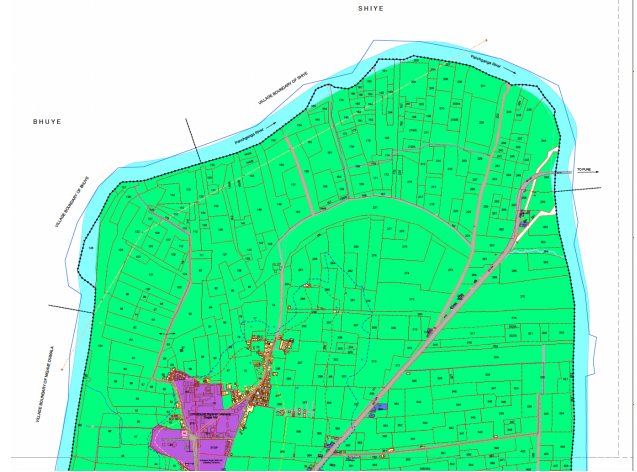
Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the groups. The sample uses matched property pairs within 0.01 miles (about 16 metres), restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025. Groups rank pairs by their mean log baseline circle rate, giving each distinct pair one vote when setting the cutpoints and assigning both members to the same group. Ties can leave unequal group sizes. Each specification interacts log baseline circle rate with the group indicators, allowing a separate slope in each group, and includes property-pair-by-year and ward-division fixed effects. Compliance elasticities divide each group slope by its full estimation-sample mean compliance. Tax paid is estimated in logs among positive payers; its group slope is already an elasticity. These group-specific slope models complement the main continuous interaction model. Standard errors cluster by circle area (parcel ID). Means and cutpoints are held fixed in inference. The joint Wald F-tests compare the displayed elasticities within each outcome using the full covariance matrix. Median split, compliance: $F(1, 142) = 0.027$, $p = 0.870$; tax paid among payers: $F(1, 129) = 0.774$, $p = 0.380$. Terciles, compliance: $F(2, 142) = 1.457$, $p = 0.236$; tax paid among payers: $F(2, 129) = 2.428$, $p = 0.092$. Standard errors are in parentheses. Stars on elasticities and F statistics denote * $p < 0.10$, ** $p < 0.05$ and *** $p < 0.01$.

B.2 Figures

Figure A1
Maps of Survey Areas and Circle Rates



Panel A: Survey area map



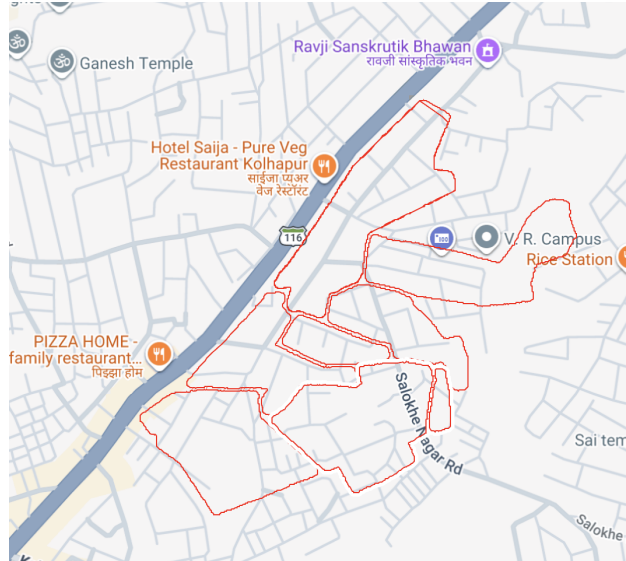
Panel B: Survey area map-2

शहरी विभागतील माहिती						
जिल्हा : कोल्हापूर						
जिल्हा विभाग : (सर्व) कोल्हापूर महानगर पालिका विभाग						
सर्वेक्षक क्र./ विभाग क्र.	सुमारे विभागाला तयार केले	चुकी जमीन	मिळालेले इमारत	एकूण मजगतीतील फाटलव/ व्यावसायिक	सह मजगतीतील दुकाने/ व्यावसायिक	अतिरिक्त क्षेत्र
एकूण क्षेत्र, चौ.मी. वी.ए. कर						
1/1	एकूण क्षेत्र, चौ.मी. वी.ए. कर	15200	18800	20700	67000	स. म. म. म.
1/2	एकूण क्षेत्र, चौ.मी. वी.ए. कर	8700	14500	15500	33000	स. म. म. म.
1/3	एकूण क्षेत्र, चौ.मी. वी.ए. कर	18500	20000	25000	28000	स. म. म. म.

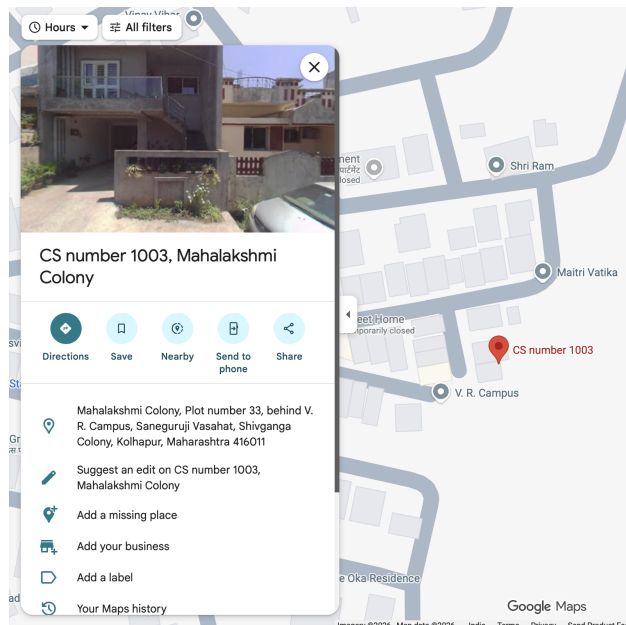
Panel C: Circle rates: 2011

Notes: This figure presents screenshots of survey area PDFs in Panels A and B and a snapshot of the PDF displaying circle rates per survey area in panel C. In Panels A and B, outlines in light red mark survey areas and the text in black marks the survey number. Other dashed red, blue, black, and yellow lines present other administrative divisions within the city.

Figure A2
Parcels and Address Example



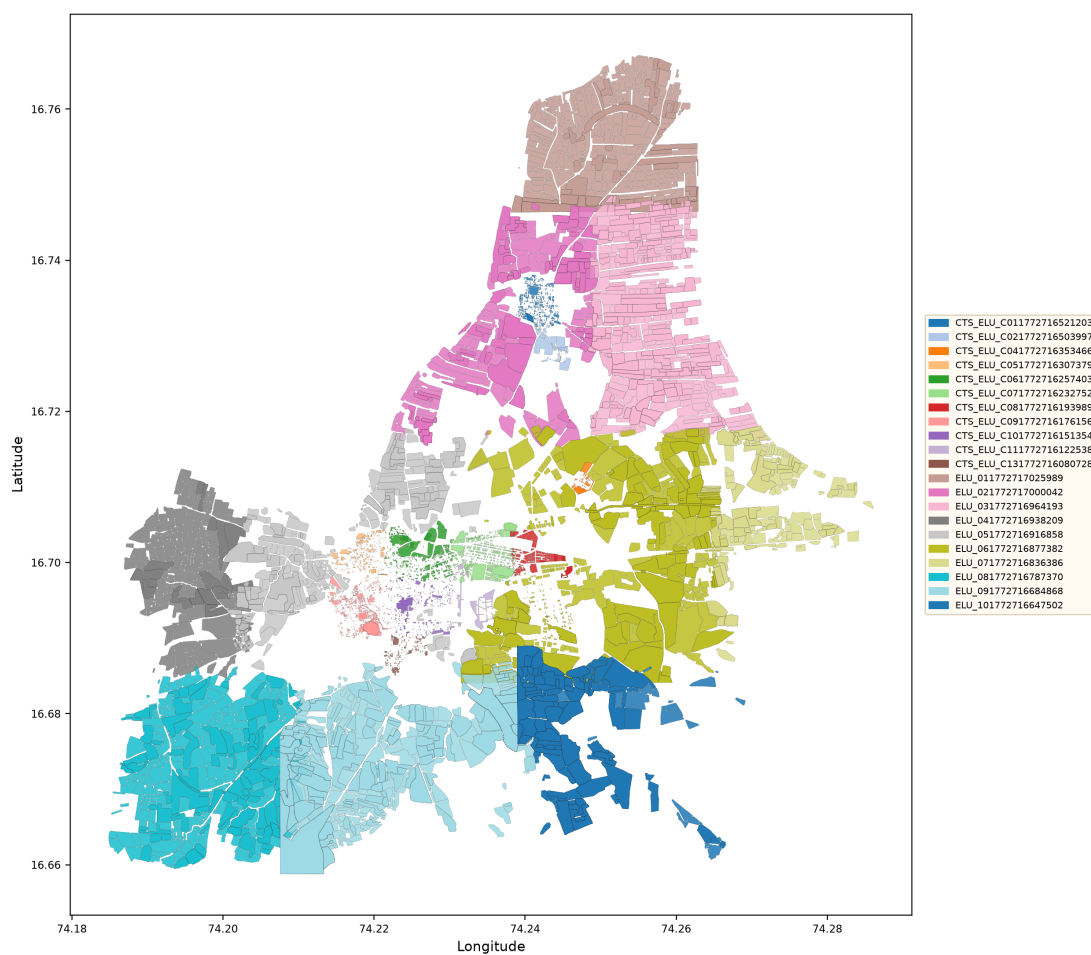
Panel A: Sample of parcels on Google Maps



Panel B: An example address

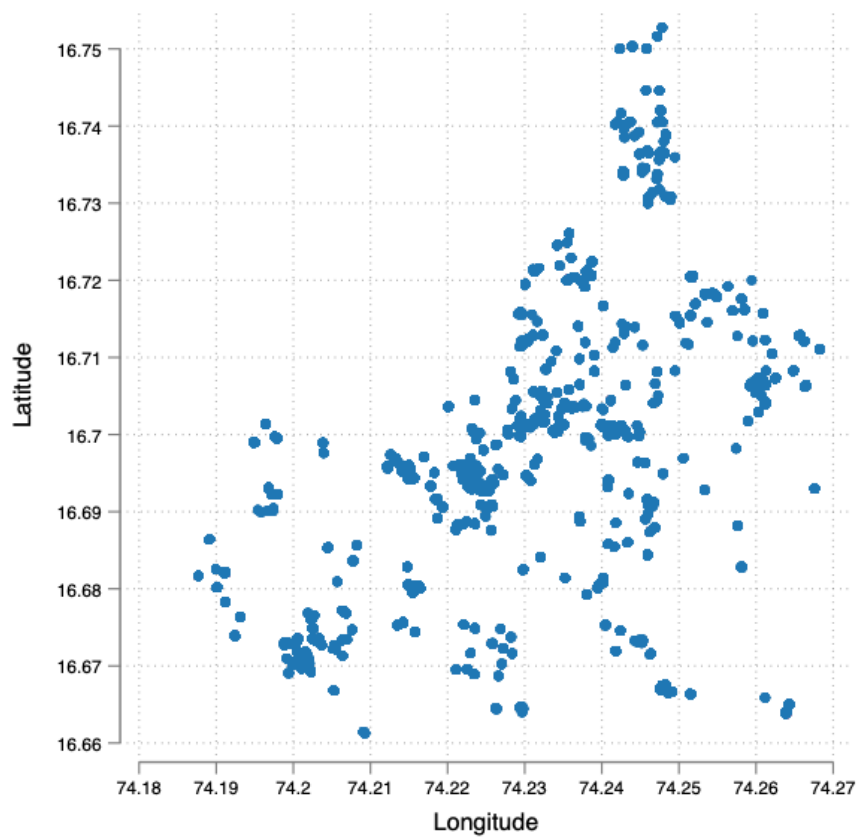
Notes: This figure presents a sample of parcels imposed on a Google Maps screenshot of the city in panel A and an example of a coordinate collected manually, from an address using Google Maps.

Figure A3
Digitising Maps



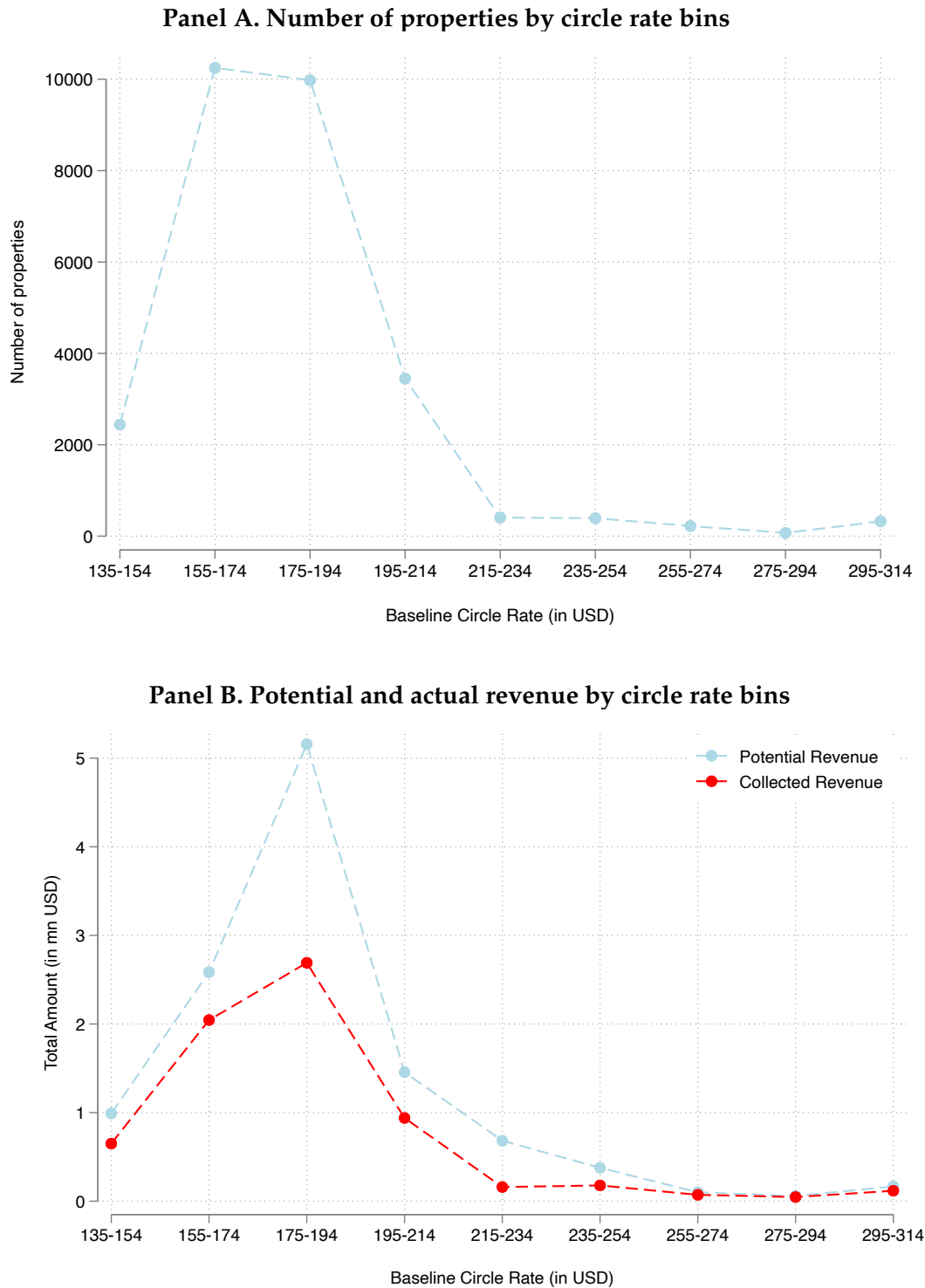
Notes: This figure shows the 18 maps that were combined and digitized to form the full map of circle boundaries. Each map is color-coded to be a different color.

Figure A4
Coordinates in the sample



Notes: Each blue dot in this figure represents a coordinate in the final sample. Each coordinate could have multiple properties.

Figure A5
Distribution of Properties and Potential versus Actual Revenue



Notes: Panel A shows the distribution of owner-occupied properties across 2011 circle-rate bins. Panel B shows actual revenue, defined as the sum of tax payments and penalties collected, and potential revenue, defined as the sum of assessed tax liabilities, within each bin. The sample is restricted to properties observed in the balanced panel from 2013 to 2024. Circle rates are expressed in U.S. dollars using an exchange rate of INR 85 per U.S. dollar. Within this sample, there is about 8 million USD in uncollected, potential revenue.

Figure A6
Google Footprints data



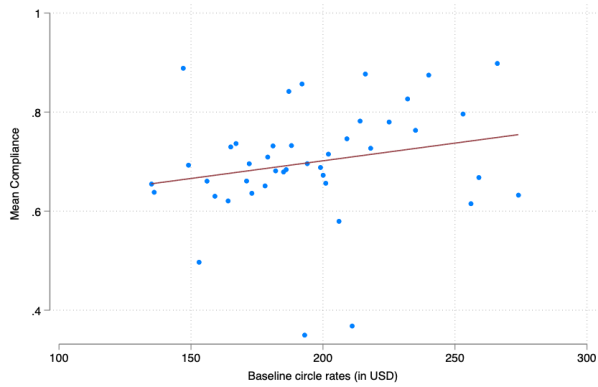
Panel A: Snapshot of the whole city



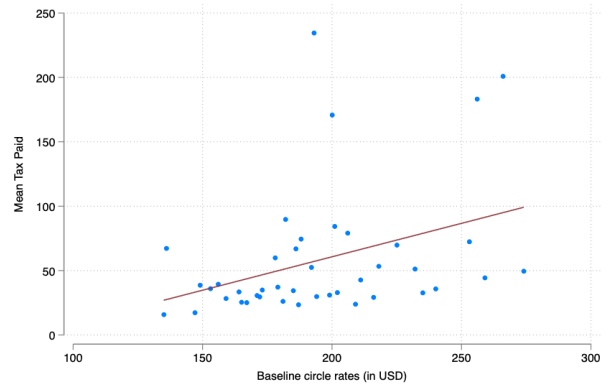
Panel B: Zoomed in

Notes: This figure presents the Google Footprints shapefile for the whole city of Kolhapur in panel A, and a zoomed-in snapshot in Panel B. In panel B, each rectangle or square is a property parcel.

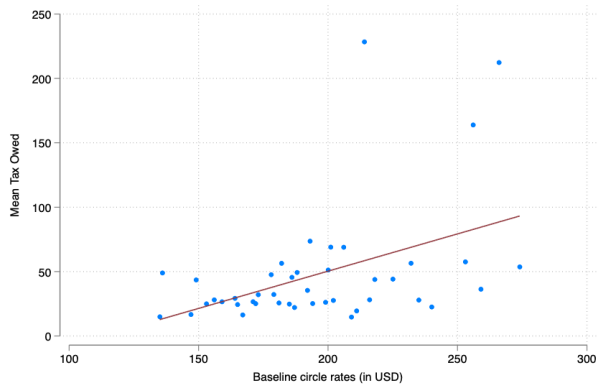
Figure A7
Descriptive Plots



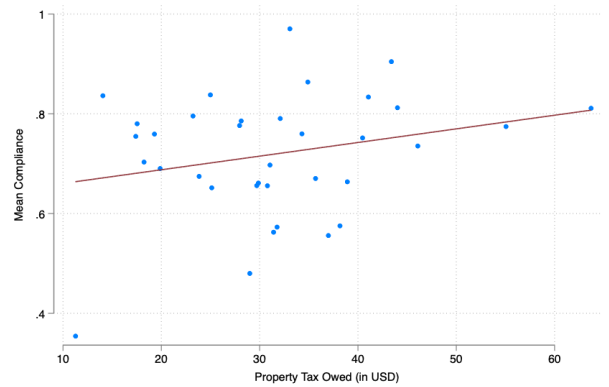
Panel A. Compliance



Panel B. Tax Paid



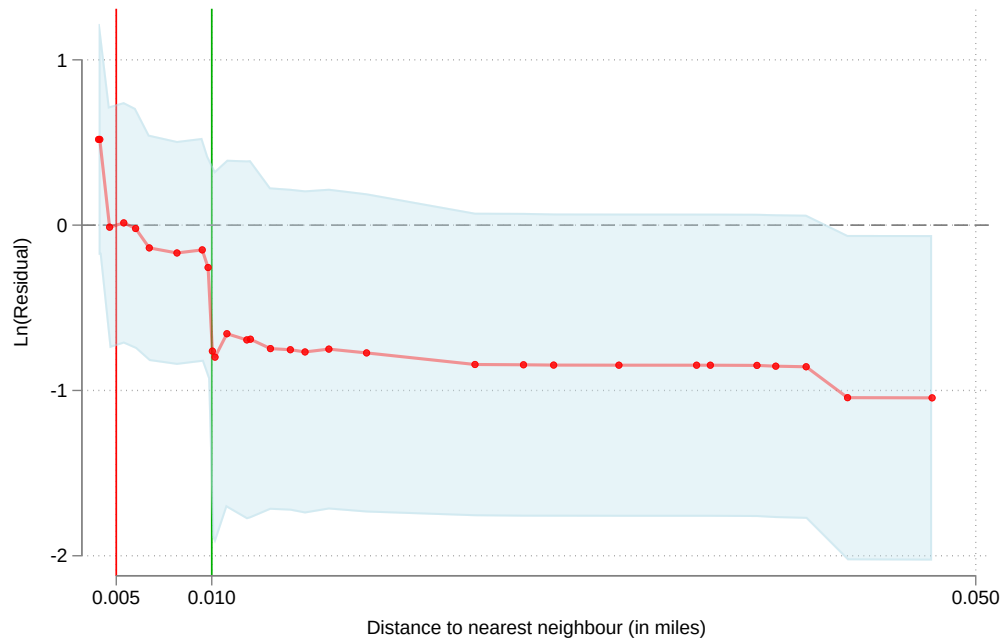
Panel C. Tax Owed



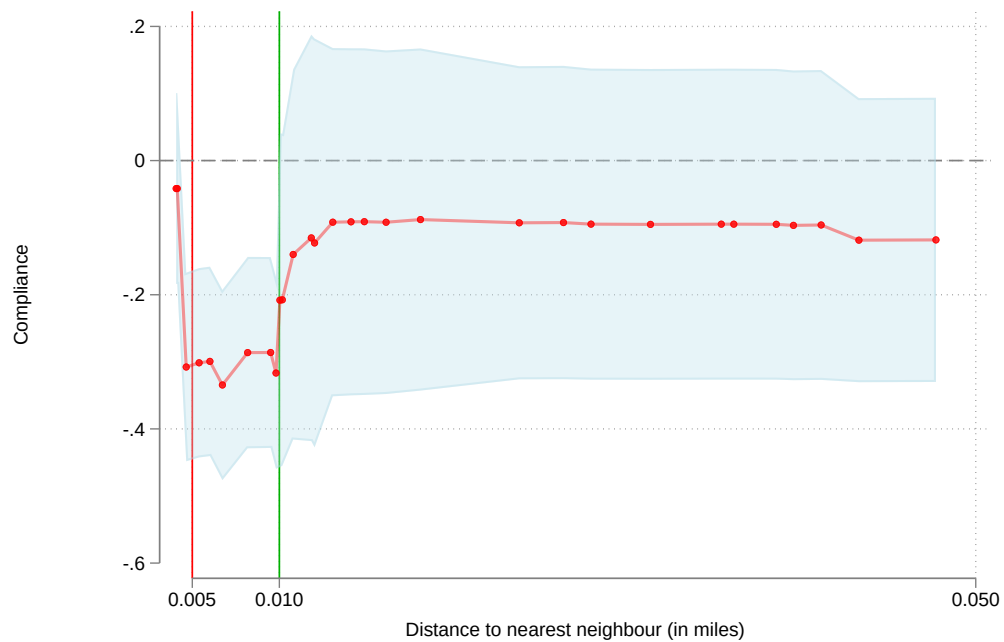
Panel D. Compliance and Tax Owed

Notes: This figure presents descriptive binned scatter plots of the properties in the final sample. Panels A, B and C plot the mean circle rate in 2011 against mean compliance, taxes paid and taxes owed, respectively. Panel D plots mean compliance against the property tax owed.

Figure A8
Distribution of the Residual and compliance after 16 metres



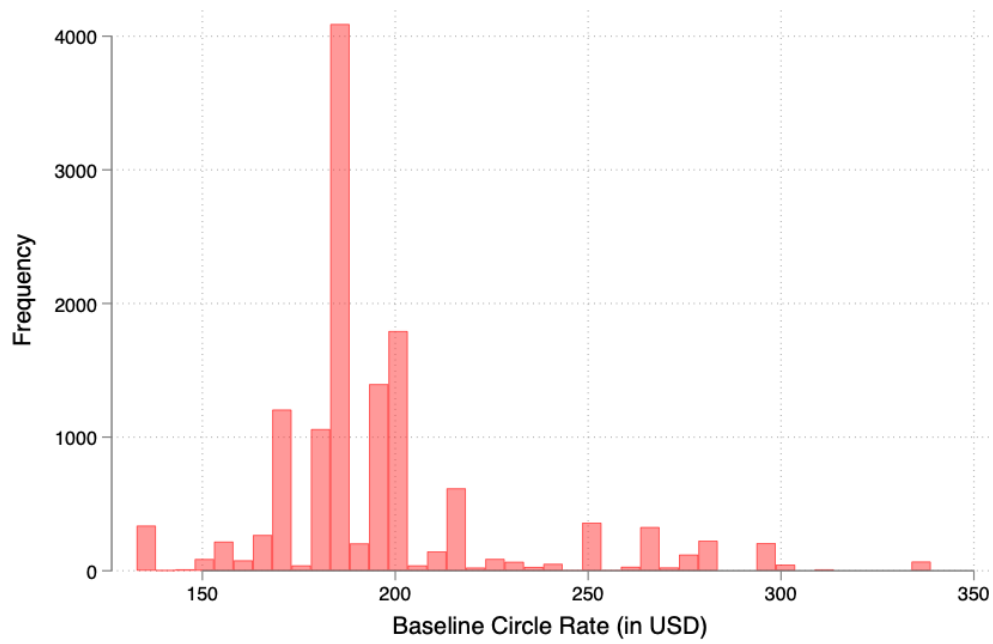
Panel A. Residual



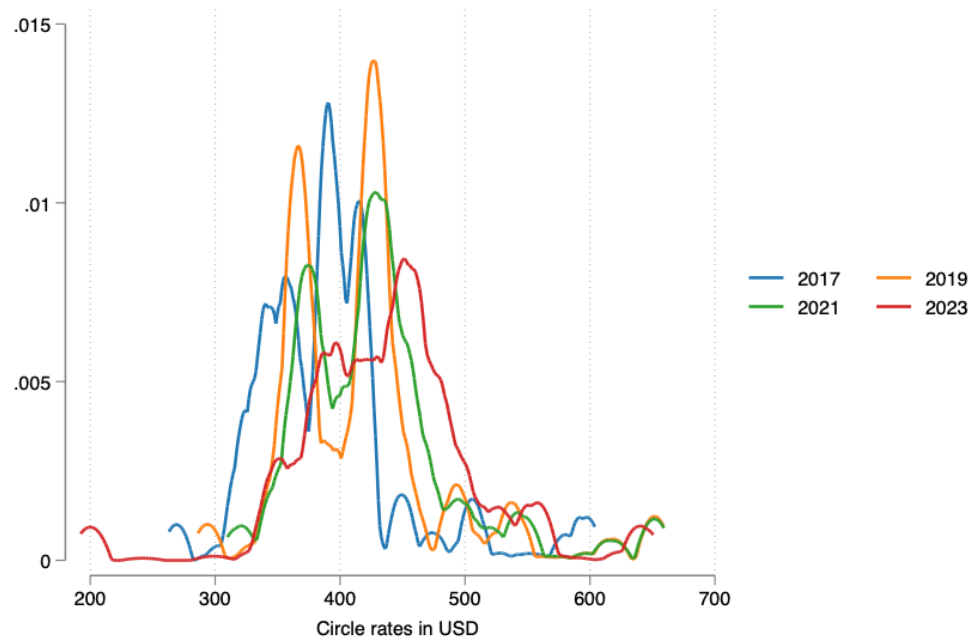
Panel B. Compliance

Notes: I exclude pairs with mean baseline circle rates above USD 280. This figure shows how the balance and compliance estimates vary as the matched-pair distance threshold expands beyond the main 16-metre cutoff. Panel A plots coefficients for the residual from the valuation formula, defined in Equation 9. Panel B plots coefficients for compliance. Confidence intervals are shown at the 95 percent level. The figure illustrates that balance is strongest at short distances and that the compliance effect attenuates as more distant pairs are included.

Figure A9
Circle Rate Distributions



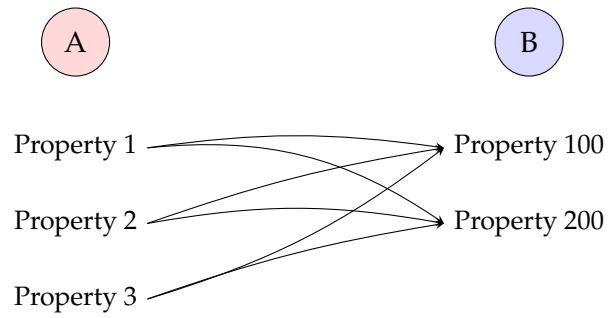
(a) Circle rate distribution: 2011



(b) Circle rate distributions: 2015–2024

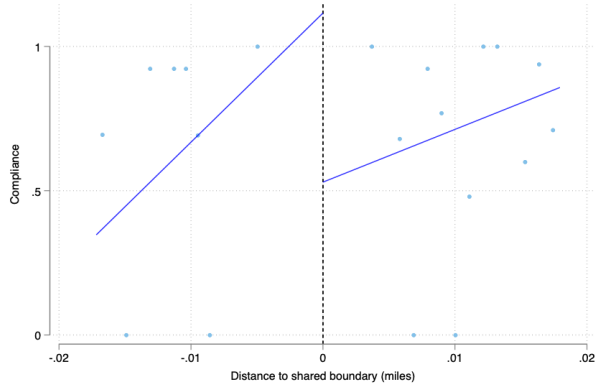
Notes: This figure presents the distribution of circle rates in Kolhapur. Panel A plots the distribution of baseline 2011 circle rates. Panel B plots density distributions for later observed circle-rate schedules in 2017, 2019, 2021, and 2023. Circle rates are expressed in U.S. dollars using an exchange rate of INR 85 per U.S. dollar.

Figure A10
Illustration of pair formation

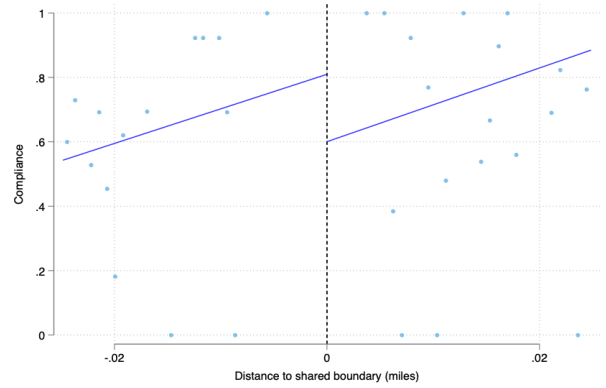


Notes: This figure provides a simplified illustration of the nearest neighbour matching between coordinates and the formation of property-pairs. The two circles A and B represent two different coordinates that got matched using a nearest-neighbor matching algorithm. Properties 1, 2 and 3 are 3 properties in coordinate A and properties 100 and 200 are 2 properties located in coordinate B. The arrows between the 5 properties represent the number of pairs being formed, which is 6 in this illustration.

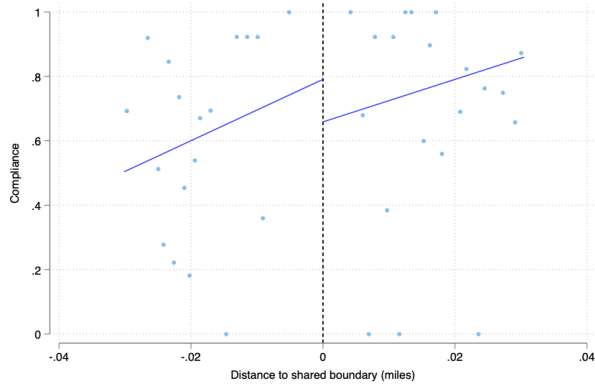
Figure A11
Regression Discontinuity estimates across bandwidths



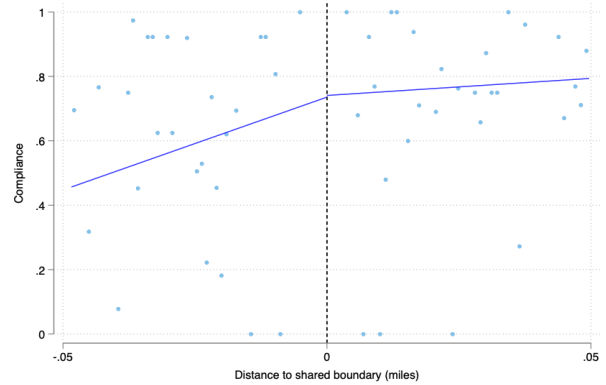
(a) Bandwidth = 0.018 miles



(b) Bandwidth = 0.024 miles



(c) Bandwidth = 0.031 miles



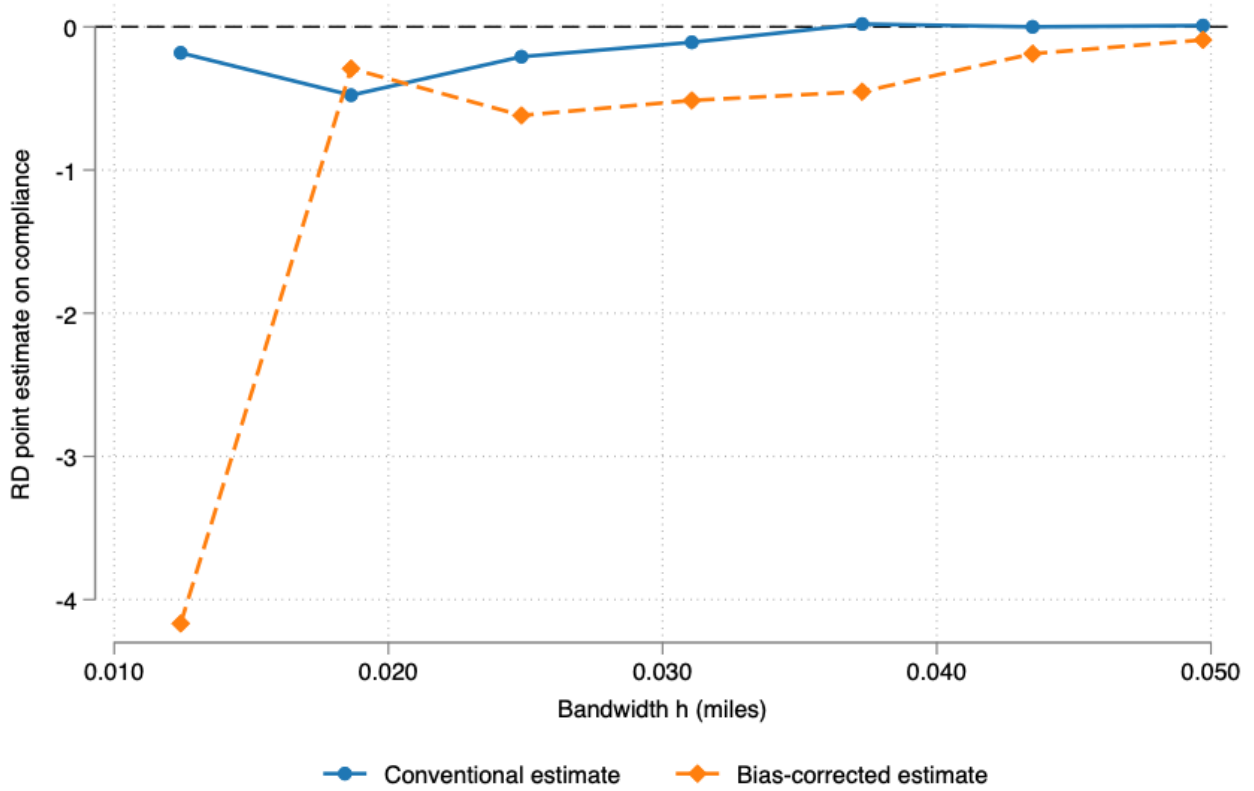
(d) Bandwidth = 0.049 miles

Notes: This figure presents regression discontinuity plots for different bandwidths across Panels (a) – (d). In each figure I estimate

$$\text{Compliance}_i = \alpha + \tau \text{High}_i + \beta_1 D_i + \beta_2 (\text{High}_i \times D_i) + \varepsilon_i,$$

where High is equal to 1 if a coordinate has the higher circle rate relative to its matched coordinate. D_i is the running variable, which is distance to the circle rate border. All specifications use linear polynomials to fit points on both sides of the discontinuity and triangular kernel weights, thus assigning higher weights to observations that lie closer to the border. Observations to the right of the vertical line at 0 are ‘treated’ observations or those where High is equal to 1. Each point is then a binned mean of the outcome variable, compliance. These plots are formed using Stata’s ‘rdplot’ command.

Figure A12
Regression Discontinuity estimates across bandwidths

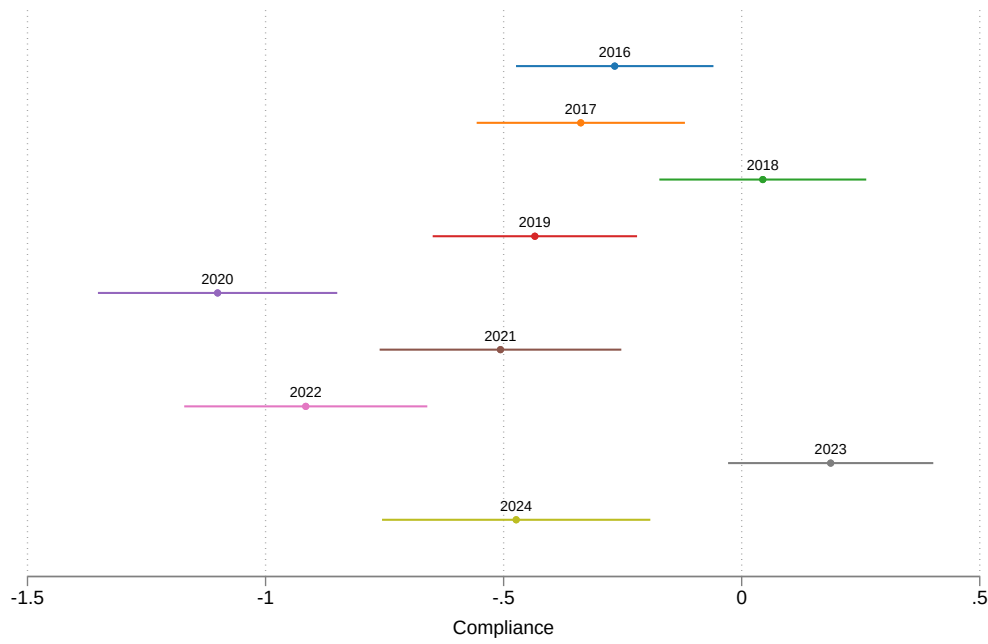


Notes: This figure presents estimates from testing:

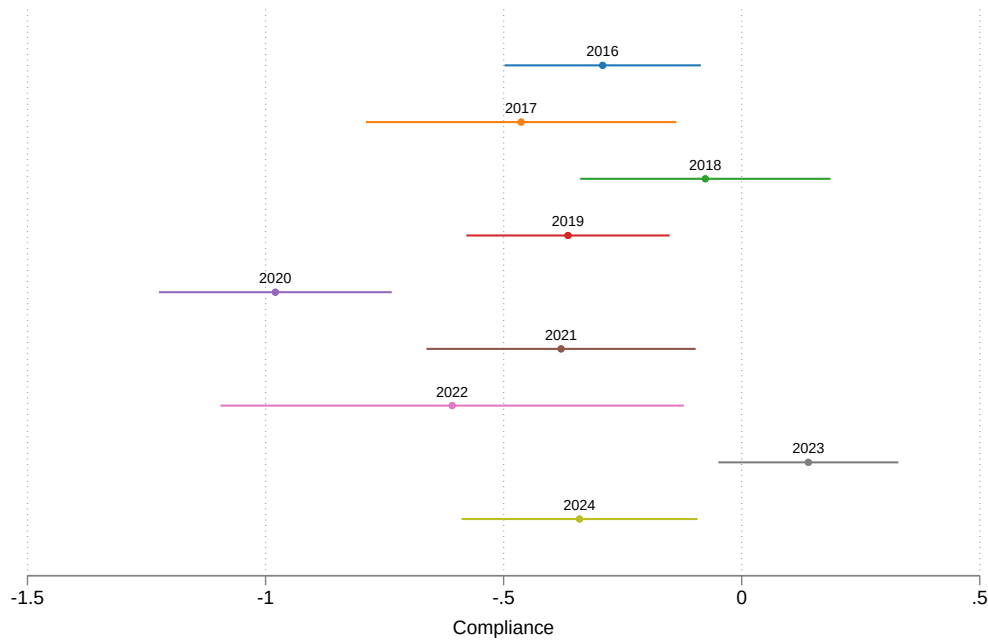
$$\text{Compliance}_i = \alpha + \tau \text{High}_i + \beta_1 D_i + \beta_2 (\text{High}_i \times D_i) + \varepsilon_i$$

for different distance bandwidths. Distances in miles are presented on the x-axis. Each point plots τ , where blue dots are conventional estimates and orange dots are bias-corrected estimates from 'rdrobust' in Stata. All specifications use linear polynomials to fit points on both sides of the discontinuity and triangular kernel weights.

Figure A13
Compliance through time



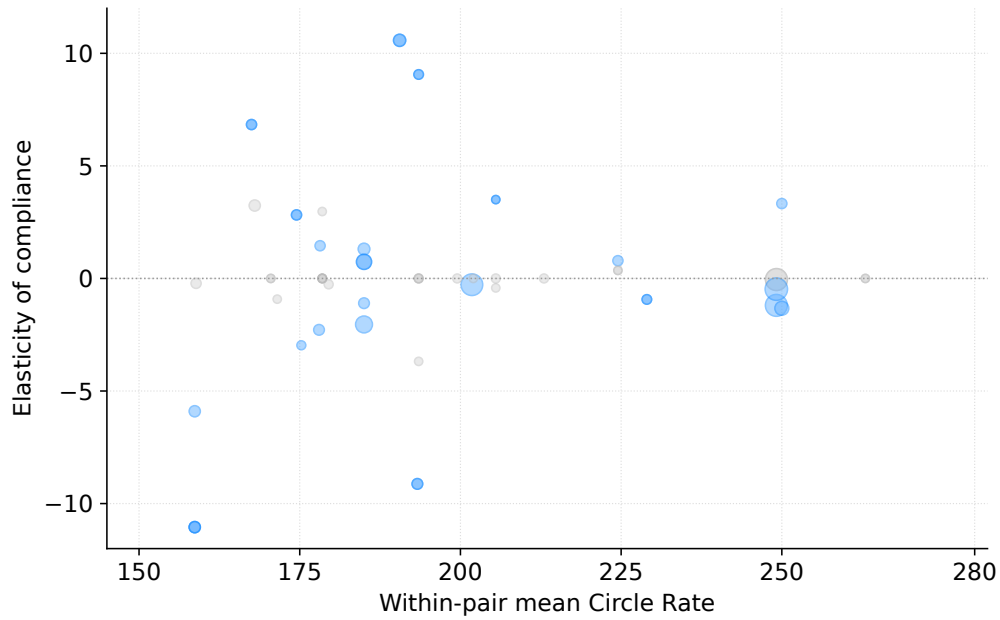
Panel A. 8 metres



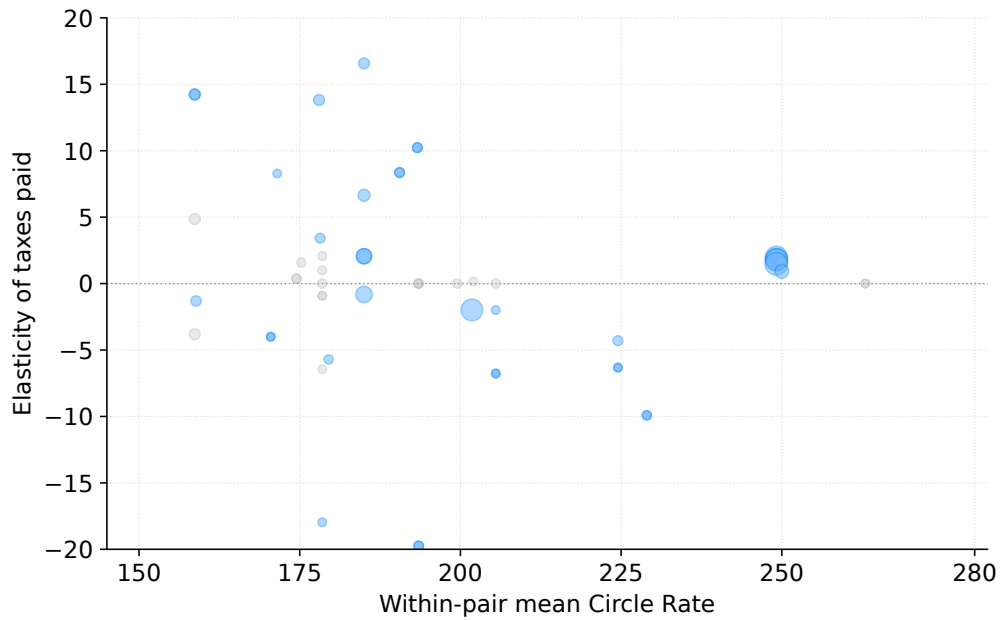
Panel B. 16 metres

Notes: I exclude pairs with mean baseline circle rates above USD 280. This figure presents the coefficient on compliance from estimating a modified version of Equation 10. Each specification uses only property-pair fixed effects and estimates a separate estimate for every year. Panel A presents estimates using a threshold distance of 8 metres and Panel B presents estimates using a distance of 16 metres. Confidence intervals are presented at the 95% level.

Figure A14
Elasticity of compliance and taxes paid: non-parametric



Panel A. Elasticity of Compliance

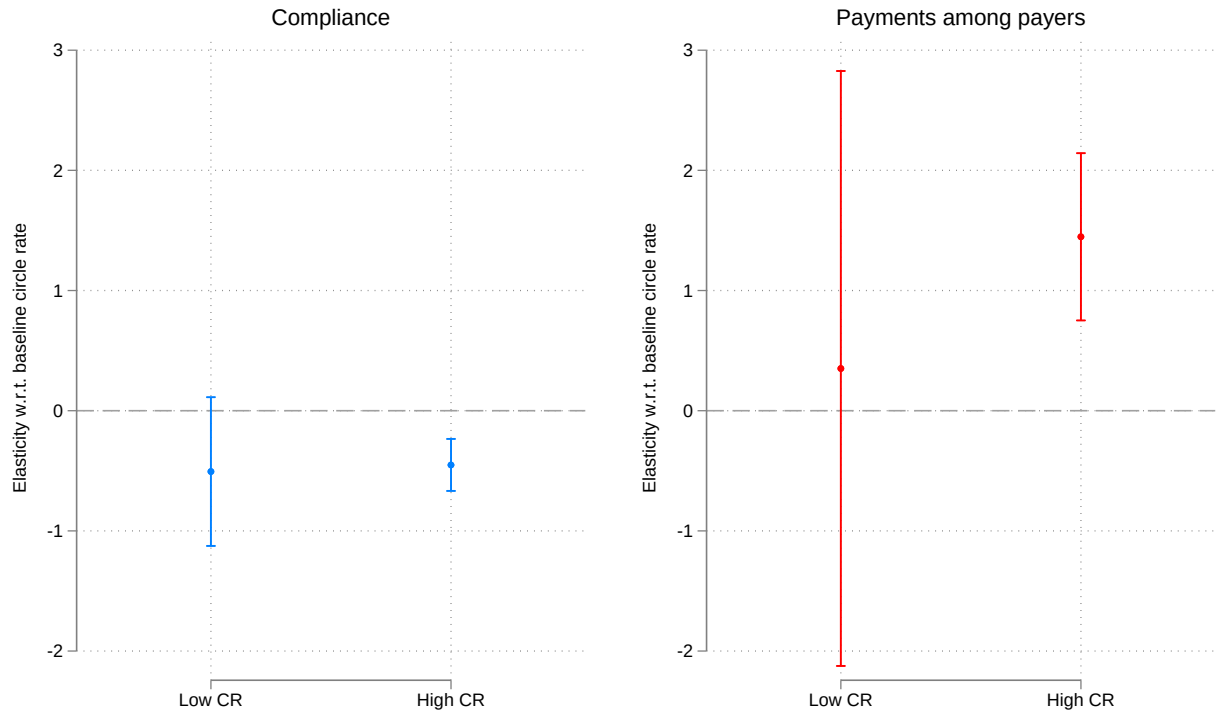


Panel B. Elasticity of Taxes Paid

Notes: I exclude pairs with mean baseline circle rates above USD 280. This figure presents the coefficients and elasticities from estimating specifications analogous to Equation 10. In Panel A, each point presents the corresponding β divided by the mean of compliance within that sample. In Panel B, each point presents the estimated β by using taxes paid as the dependent variable and a sample of compliers. Estimates are presented at the mean of the circle-rate within the pair. Larger circles represent more properties. Blue points are significant at the 10% level and grey points are not. I keep the original display ranges; estimates outside these ranges are not shown.

Figure A15
Compliance and tax-paid elasticities: median split

Median split: baseline circle rate heterogeneity

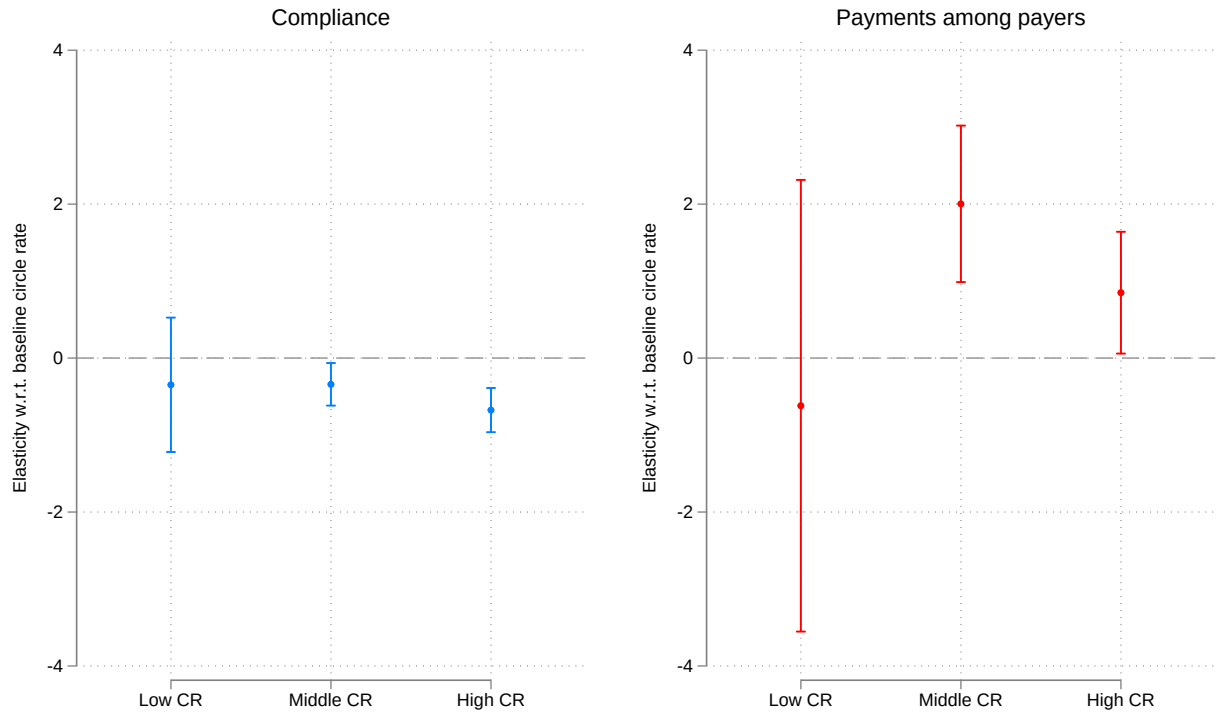


95% CIs; SEs clustered by parcel. Compliance uses full-bin means.

Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the groups. Points show group elasticities with 95% confidence intervals. The left panel reports compliance and the right panel reports tax paid among payers. The sample uses matched property pairs within 0.01 miles (about 16 metres), restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025. Groups rank pairs by their mean log baseline circle rate, giving each distinct pair one vote when setting the cutpoints and assigning both members to the same group. Ties can leave unequal group sizes. Each specification interacts log baseline circle rate with the group indicators, allowing a separate slope in each group, and includes property-pair-by-year and ward-division fixed effects. Compliance elasticities divide each group slope by its full estimation-sample mean compliance. Tax paid is estimated in logs among positive payers; its group slope is already an elasticity. These group-specific slope models complement the main continuous interaction model. Standard errors cluster by circle area (parcel ID). Means and cutpoints are held fixed in inference. The joint Wald F-tests compare the displayed elasticities within each outcome using the full covariance matrix. Joint equality tests, compliance: $F(1, 142) = 0.027$, $p = 0.870$; tax paid among payers: $F(1, 129) = 0.774$, $p = 0.380$. Estimates and tests also appear in Appendix Table [A15](#).

Figure A16
Compliance and tax-paid elasticities: terciles

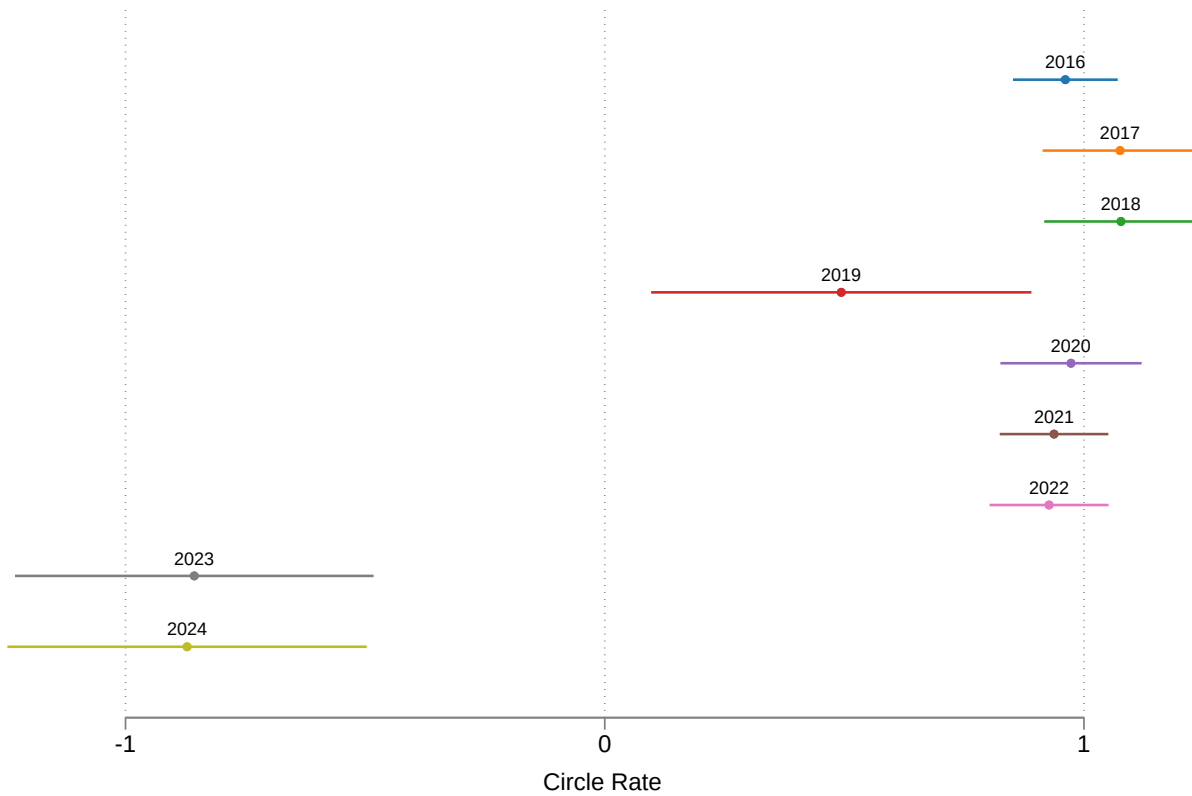
Terciles: baseline circle rate heterogeneity



95% CIs; SEs clustered by parcel. Compliance uses full-bin means.

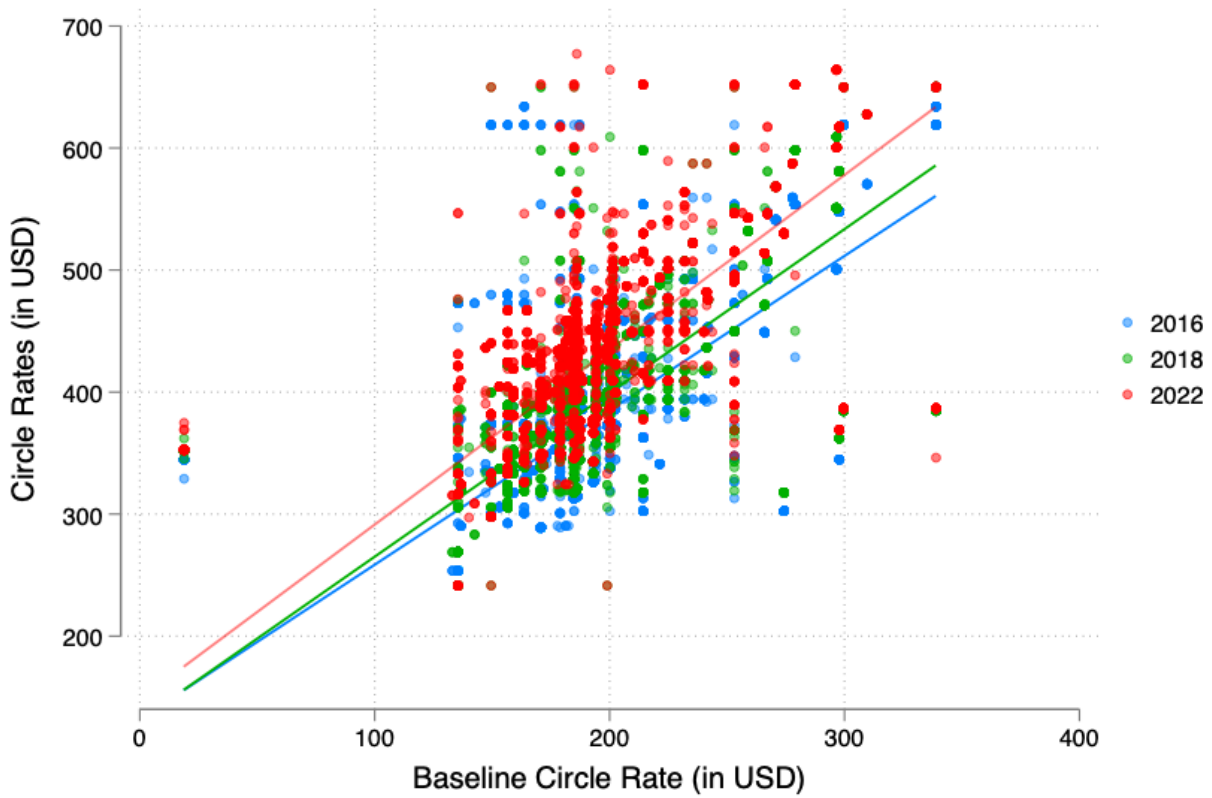
Notes: I exclude pairs with mean baseline circle rates above USD 280 and recompute the groups. Points show group elasticities with 95% confidence intervals. The left panel reports compliance and the right panel reports tax paid among payers. The sample uses matched property pairs within 0.01 miles (about 16 metres), restricted to balanced-panel, owner-occupied, non-commercial and non-institutional properties observed before 2025. Groups rank pairs by their mean log baseline circle rate, giving each distinct pair one vote when setting the cutpoints and assigning both members to the same group. Ties can leave unequal group sizes. Each specification interacts log baseline circle rate with the group indicators, allowing a separate slope in each group, and includes property-pair-by-year and ward-division fixed effects. Compliance elasticities divide each group slope by its full estimation-sample mean compliance. Tax paid is estimated in logs among positive payers; its group slope is already an elasticity. These group-specific slope models complement the main continuous interaction model. Standard errors cluster by circle area (parcel ID). Means and cutpoints are held fixed in inference. The joint Wald F-tests compare the displayed elasticities within each outcome using the full covariance matrix. Joint equality tests, compliance: $F(2, 142) = 1.457$, $p = 0.236$; tax paid among payers: $F(2, 129) = 2.428$, $p = 0.092$. Estimates and tests also appear in Appendix Table A15.

Figure A17
Dynamics: How are contemporaneous circle rates related to baseline 2011 circle rates?



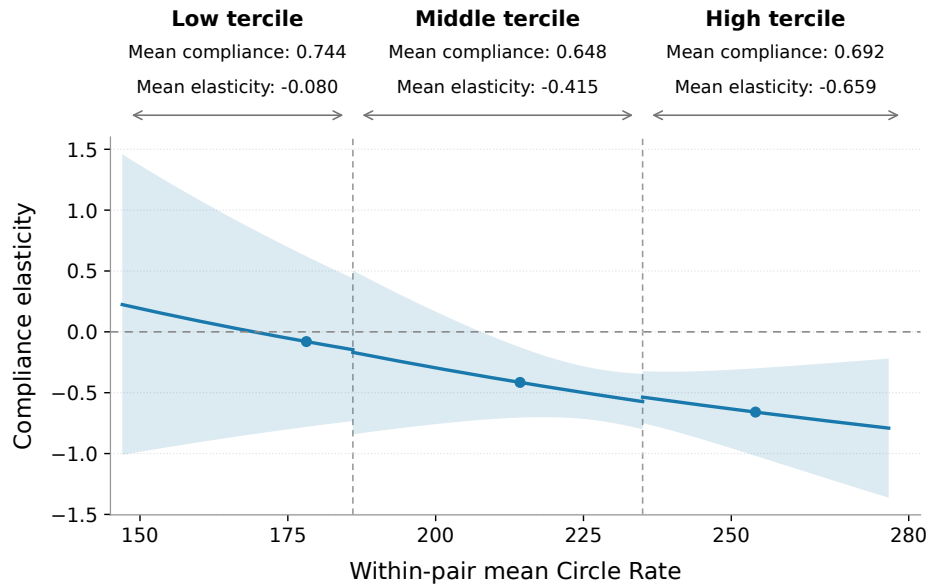
Notes: I exclude pairs with mean baseline circle rates above USD 280. This figure reports dynamic estimates of the relationship between baseline circle rates and contemporaneous circle rates using the 16-metre matched-pair sample. Points show year-specific or event-time coefficients from modified versions of the main matched-pair specification. Confidence intervals are shown at the 95% level. The figure is used to assess the evolution of the relationship between circle rates in 2011 and those from 2016–2024.

Figure A18
Baseline circle rate versus circle rates in 2016, 2018, 2022

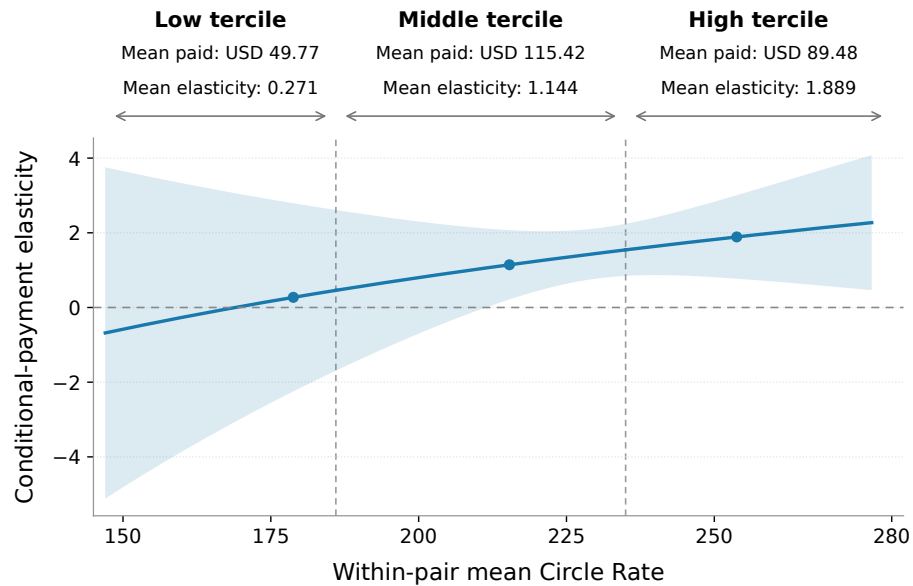


Notes: This figure plots baseline 2011 circle rates against later circle rates in selected years. Each point represents a circle-rate area. The figure is used to assess whether areas with high baseline circle rates also experienced faster subsequent growth in administratively assigned circle rates.

Figure A19
Elasticity of Compliance and Taxes Paid: heterogeneity, nonlinear moderator



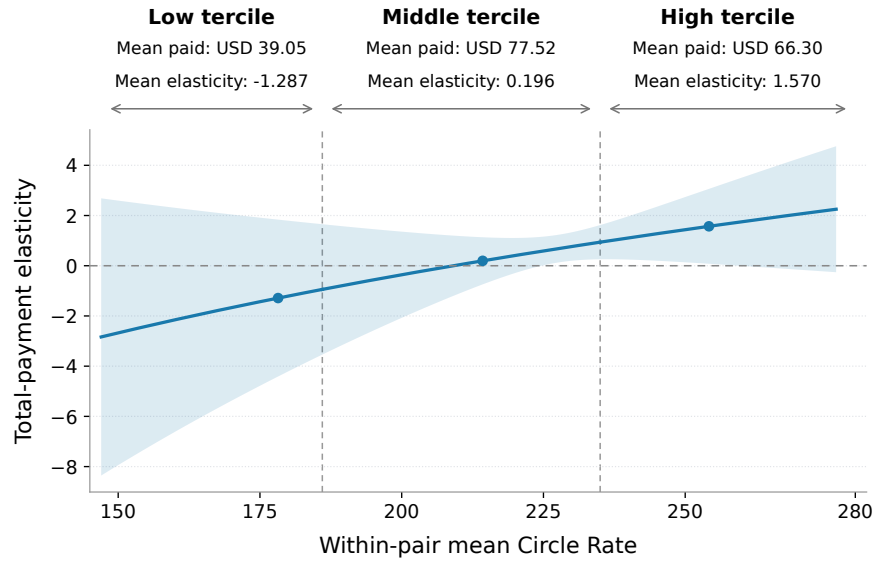
Panel A. Compliance



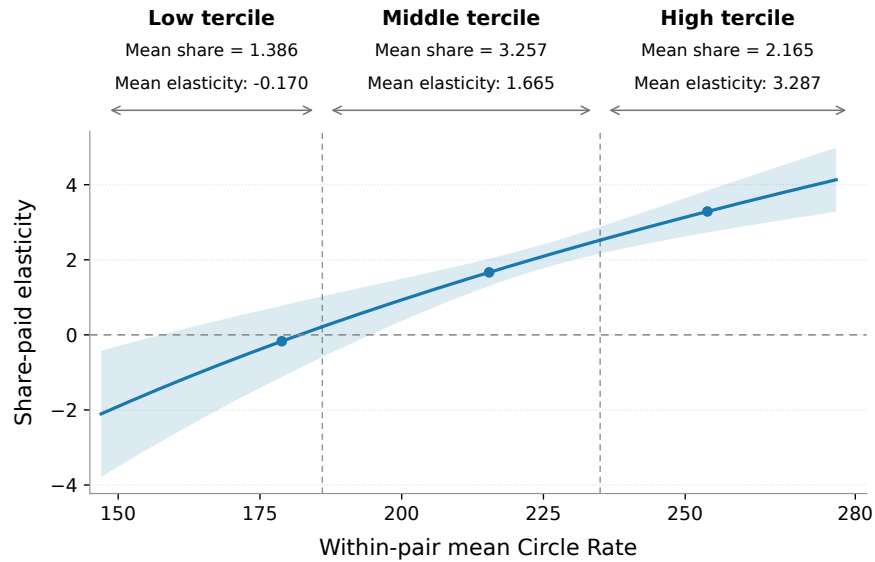
Panel B. Elasticity of Taxes Paid

Notes: Preferred distance: 0.01 miles. Entire pairs with geometric-mean baseline circle rates above USD 280 are excluded. Terciles are recomputed from retained distinct pairs, counting each pair once (cutpoints: USD 186 and USD 235), as in Table 1. The nonlinear moderator is the log of the within-pair mean log baseline circle rate, centered in this restricted sample. Panel A divides compliance slopes by the corresponding tercile mean compliance. Panel B models log payments among payers; its slopes are elasticities directly. Both horizontal axes show pair geometric-mean circle rates in USD. Dashed lines and arrows mark terciles; dots and labels report their means. Shading shows 95% confidence intervals. Joint equality tests: compliance $F(2,142)=12.663$ ($p<0.001$); conditional payments $F(1,129)=0.929$ ($p=0.337$). Models include pair-by-year and ward-division fixed effects; inference clusters by circle area and coordinate-group-by-circle area, with means and cutpoints held fixed.

Figure A20
Elasticity of taxes paid and share of taxes paid: heterogeneity (PPML)



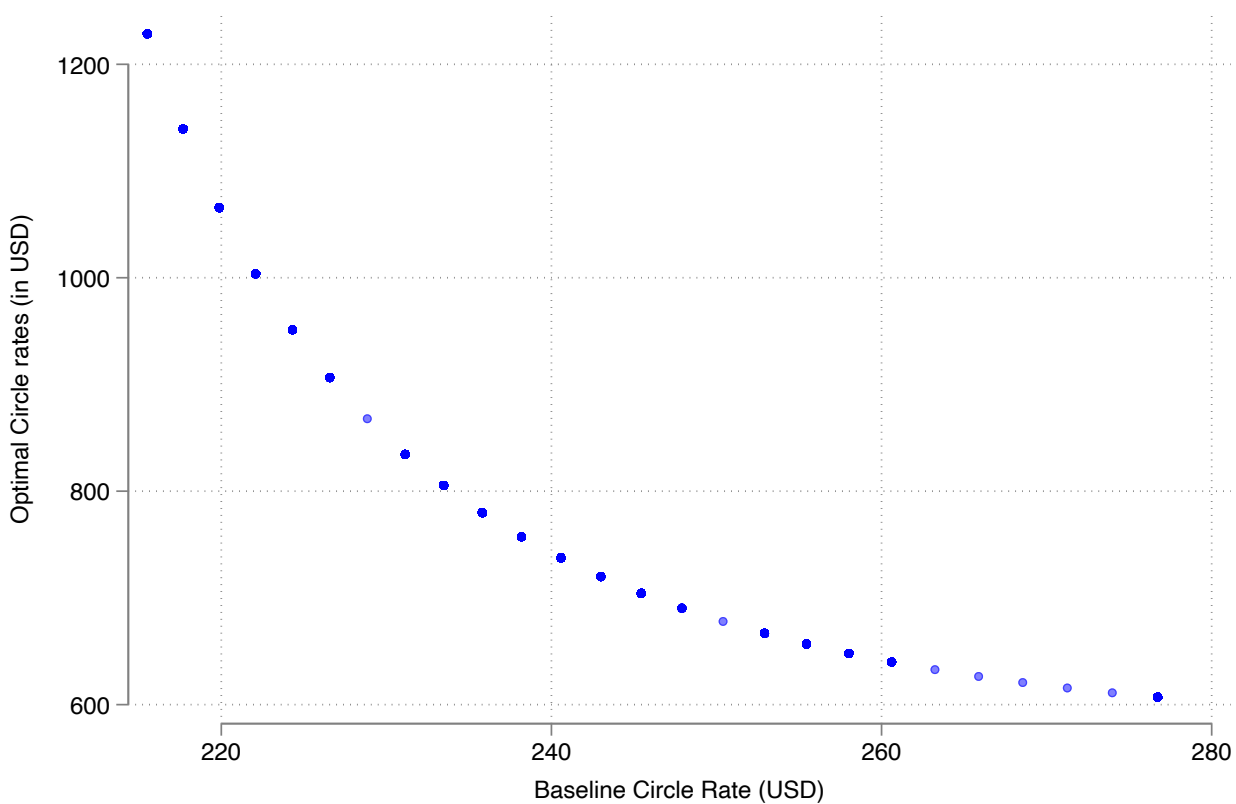
Panel A. Total payments, including zeros (PPML)



Panel B. Share of liability paid among payers (PPML)

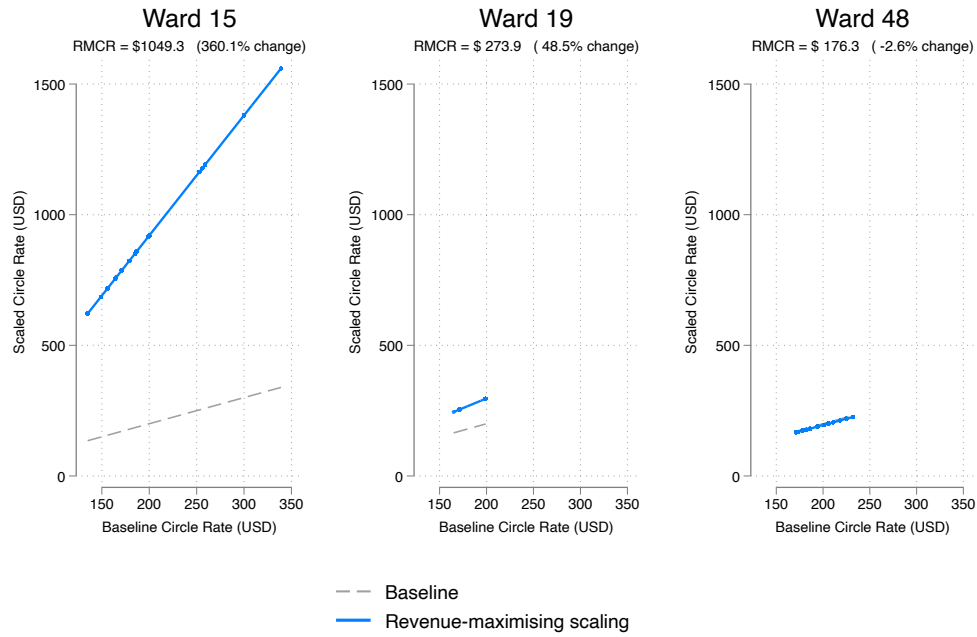
Notes: Both panels use Poisson pseudo-maximum likelihood (PPML) with outcomes in levels and a log link. The preferred distance is 0.01 miles. Entire pairs with geometric-mean baseline circle rates above USD 280 are excluded before estimation; retained distinct pairs define the terciles (cutpoints: USD 186 and USD 235), as in Table 1. Panel A uses payments including zeros; Panel B uses the share of liability paid among payers. Both interact log baseline circle rate with centered within-pair mean log baseline circle rate. Slopes are elasticities directly, without division by outcome means. Dashed lines and arrows mark terciles; dots report elasticities at tercile mean moderators, and mean outcomes are descriptive. Shading shows 95% confidence intervals. Joint tercile equality tests: total payments $F(1,142)=1.604$ ($p=0.207$); share paid $F(1,129)=25.490$ ($p<0.001$). Models include pair-by-year and ward-division fixed effects and two-way clustering by circle area and coordinate-group-by-circle area. Means and cutpoints are held fixed. PPML excludes singleton and separated observations, including all-zero fixed-effect cells; Panel A uses 17,158 observations and Panel B 9,036. Wald statistics are reported as F statistics using restriction rank and the smaller cluster count minus one.

Figure A21
Local RMCRs measured using heterogeneous elasticities



Notes: This figure plots local revenue-maximizing circle rates implied by heterogeneous elasticities across the baseline circle-rate distribution. Each point maps a baseline 2011 circle rate to the RMCR implied by the estimated local compliance and payment elasticities. The figure is an appendix exercise because local estimates at the lower end of the circle-rate distribution imply very high RMCRs, reflecting relatively weak estimated compliance responses in that part of the distribution.

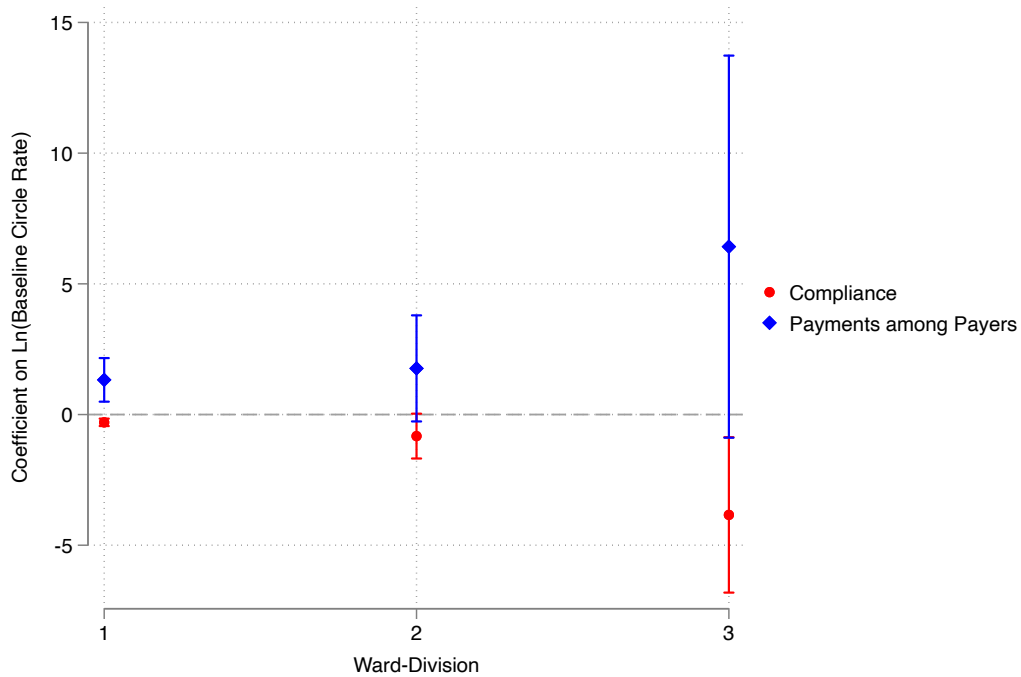
Figure A22
Ward-level circle rate schedule for revenue maximisation



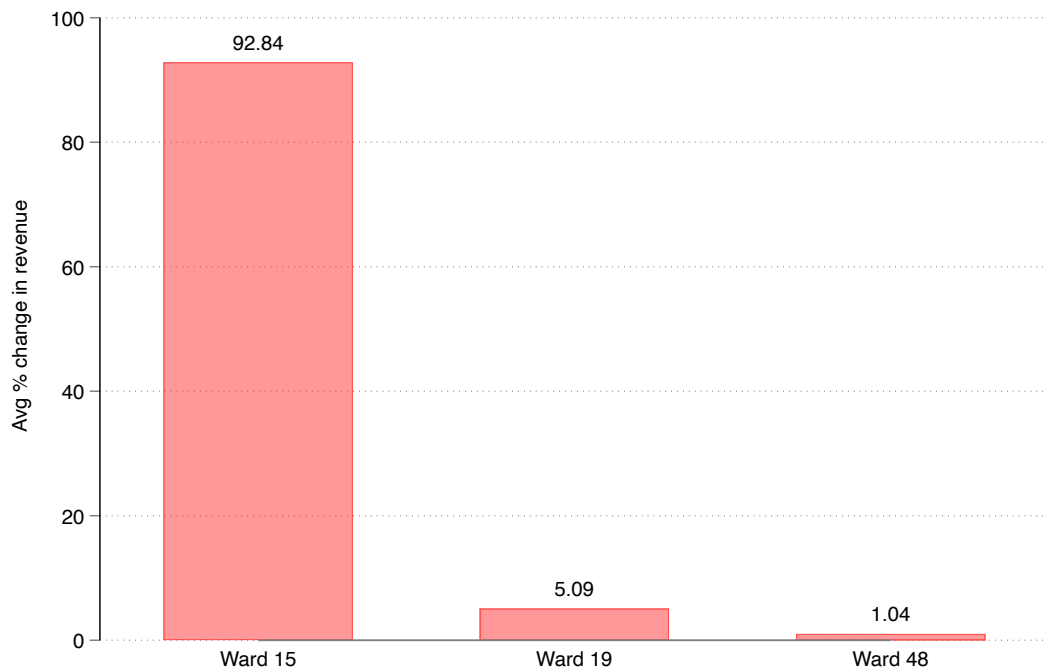
Ward-level updates using ward-specific average responses

Notes: This figure shows ward-specific RMCR schedules for selected wards using ward-specific average responses. Not all baseline circle rates are observed within every ward. I plot the corresponding revenue changes from updating to the ward-level elasticities in Figure A23.

Figure A23
Ward-specific RMCR



Panel A. Ward-level elasticities



Panel B. Ward-level changes in revenue

Notes: This figure reports ward-level heterogeneity in behavioral responses and implied revenue gains. Panel A plots ward-division-specific coefficients on log baseline circle rates, estimated separately for compliance and log tax paid among payers; whiskers show 95 percent confidence intervals. Panel B reports predicted average percentage changes in revenue from moving selected wards to their ward-specific RMCR schedules. Calculations combine the estimated extensive-margin compliance response with the intensive-margin payment response. Positive values indicate higher predicted revenue relative to the baseline schedule.