Extended Abstract of
Greener on the Other Side: Inequity and Tax Collection

Michael Carlos Best François Gerard Evan Kresch Joana Naritomi
Laura Zoratto

June 6, 2019

Abstract

This project works with the government of the city of Manaus, Brazil to improve compliance with the municipal property tax. Specifically, we aim to understand the role that inequity – similar households being treated differently by tax policy – plays in determining compliance with the tax. We do this by combining a novel survey experiment raising the salience of horizontal inequity with rich administrative data on tax liabilities and tax compliance.

1 Introduction

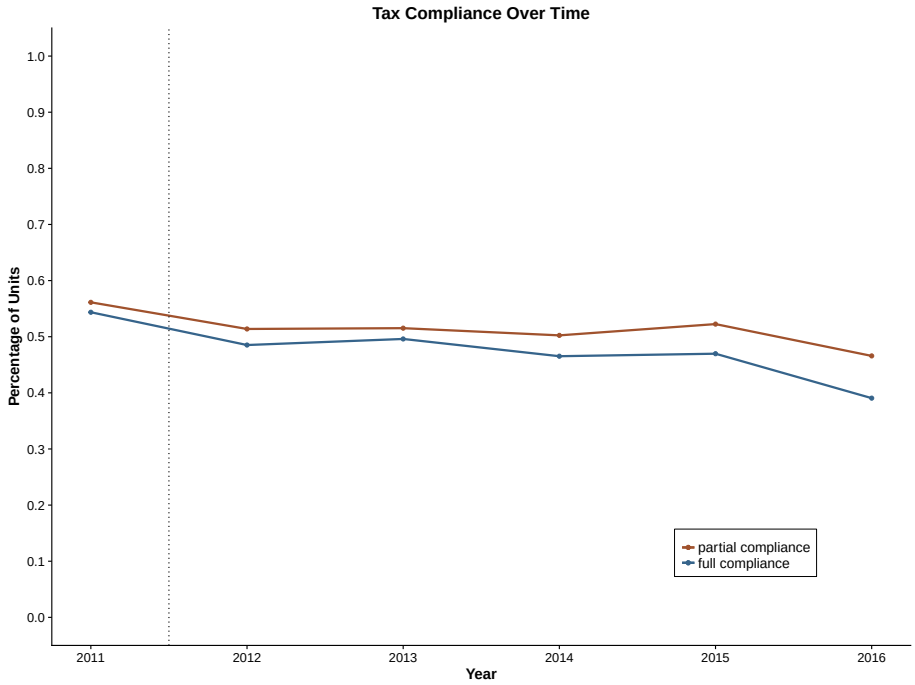
1.1 Motivation

Low state capacity in developing countries has severe implications for tax compliance and the design of tax systems. It is more challenging to assess tax liabilities correctly, especially for modern tax instruments such as personal or corporate income taxes. It is also more challenging to collect payment of assessed tax liabilities. To improve compliance, developing countries use presumptive tax bases more extensively, basing tax liabilities on variables that are easier to measure for the government and harder to misreport for taxpayers. However, presumptive taxes often generate severe horizontal inequities across taxpayers (e.g., Auerbach and Hassett, 1999). For example, the ideal tax base for a recurring property tax is the market value of a property (Poterba 1984), but under a presumptive tax system based on coarse but easy to measure proxies for market value, two properties with identical market values can face markedly different tax liabilities.

The literature has discussed the revenue efficiency implications of using presumptive tax systems, and how the associated gains may outweigh possible losses in production efficiency (e.g., Best et al., 2015). There has been less attention, however, to their fairness implications and how the perception of fairness may affect tax compliance, and the efficiency of the tax system (Besley et al., 2014). There is compelling evidence from a range of other domains that perceptions of horizontal inequity can distort behavior in important ways. Horizontal inequity in pay within firms, for instance, affects workers' job satisfaction, as well as their behaviors along various relevant margins (e.g., attendance, quit, productivity), leading to meaningful efficiency losses.

In partnership with the Finance department of the city of Manaus, Brazil, this project aims to provide some of the first compelling evidence on the importance of horizontal equity for tax compliance and the efficiency of presumptive tax systems, which are ubiquitous in developing countries. The setting for our project is the city of Manaus, Brazil, where the property tax is a key source of tax revenue, but tax delinquency is a major issue: in 2017, only 48% of taxpayers paid their property tax bills, the lowest compliance rate among all Brazilian cities. The property tax (IPTU) liability is based on proxies for home values and the formula generates clear horizontal inequities. Specifically, it depends on a scaling factor that varies across 65 tax sectors (i.e., geographical zones) of the city, with most borders set in 1983. As a result, adjacent properties along the tax sector boundaries can face very different tax liabilities despite being otherwise identical. In fact, citizens' complaints about the unfairness of the system forced the government to divide 3 tax sectors into finer sectors (with their own scaling factors) in 2017. This context is convenient for our study. First, the horizontal inequities can be very salient (neighbors in similar houses facing very different liabilities). Second, unlike in other settings, the tax liability is essentially fixed. It depends on a few easily-observable inputs, not on property tax inspectors' or homeowners' reports of

Figure 1: Overall Compliance with Property Tax



property characteristics. This allows us to focus on an underused measure of leakage – tax delinquency – that we can directly observe in administrative data, as opposed to tax evasion.

This extended abstract contains two parts. Section 2 describes a quasi-experimental analysis of administrative records studying the impact of tax rates on tax delinquency. Section 3 presents the pre-analysis plan for a survey experiment allowing us to focus on the role of fairness perceptions in determining property tax delinquency.

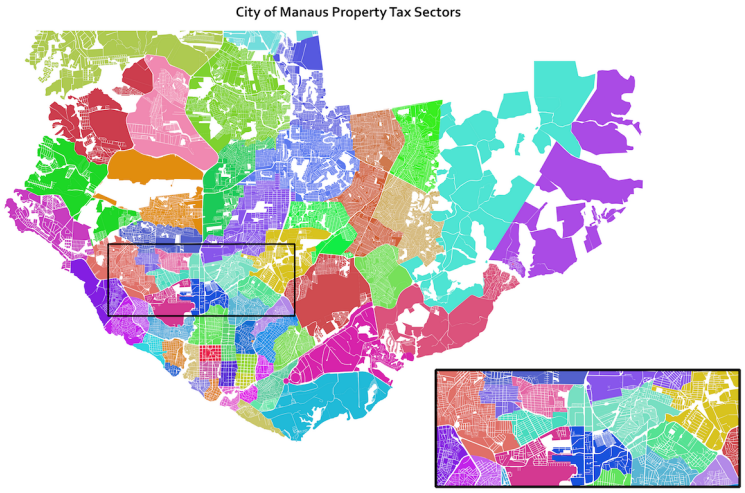
2 Tax Rates and Delinquency in a Spatial Regression Discontinuity Design

2.1 Setting and Data

Manaus is a city of 2 million people, with GDP per capita of US$10,000. The property tax is a major source of revenue for the municipal government. The tax liability is assessed by the municipal government and the government informs households of their liability in January of each year. Failure to pay bill results in fines, interest, or legal recourse. Despite this, compliance is fairly low: Average compliance is $\sim$ 50%, driven by extensive margin delinquency: households who fail to pay any of their property tax liability. Figure 1 shows the evolution of overall compliance since 2011.

A household i's property tax liability in year t T_{ist} is calculated according to the following

Figure 2: Tax Sectors of Manaus



formula

$$T_{ist} = A_t \left[(P_s \times S_{it} \times L_{it}) + H_{it} \right] \tag{1}$$

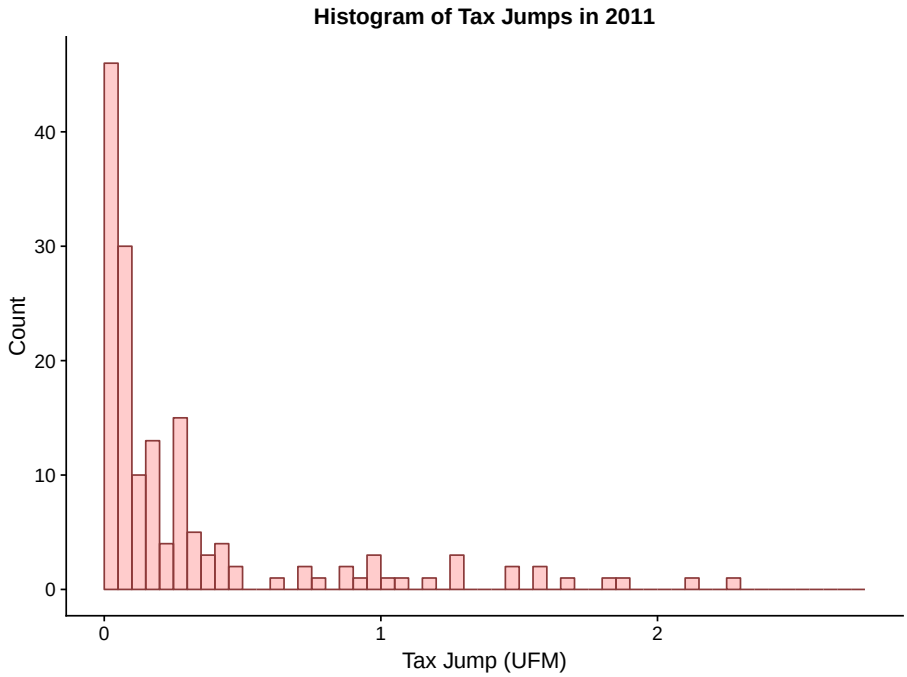
where H_{it} are characteristics of the house, S_{it} is the size of the lot, L_{it} are characteristics of the land; P_s is the price assigned to each square meter of land in sector s, and A_t is a year-specific scaling factor.

The key source of our identifying variation comes from the prices assigned to land in each sector s of the city. Manaus was divided into 62 sectors in 1983. A committee made divisions trying roughly to approximate house prices in 1983. These sectors are not used to determine the provision of any other munipal services. Figure 2 shows a map of the sectors

The existence of these tax sectors means that at the boundaries of the sectors, tax liabilities jump discretely. These tax jumps create differences in the tax treatment of neighbors who, as we argue below, are otherwise comparable. Figure 3 shows that these differences can often be large. The figure shows the distribution of tax rate jumps at all of the boundaries in the city in 2011. These jumps thus present an opportunity to employ a spatial regression discontinuity design to study the impact of the tax rates on delinquency.

Our data come from the Finance Secretariat (SEFAZ) in Manaus. The data includes the full cadaster of over 500,000 properties in Manaus, GIS data on all city blocks and the road network as well as detailed records on each property's tax liability, assessed fines and interest, payments and legal follow-up on delinquent households. We also have tax records on property transfers including their transfer value and accompanying transaction tax payment. The data spans the period from 200 to the present.

Figure 3: Jumps in Tax Rates at Sector Boundaries Can Be Large



2.2 Spatial Regression Discontinuity Design

For all adjacent sectors, we define "High" & "Low" tax sectors using the prices per square meter (P_s). We then implement the following spatial RD specification:

$$y_{ib} = \alpha + \beta HighRate_i + f(distance_i) \tag{2}$$

$$+ g\left(distance_i\right) \times HighRate_i + \delta_b + \varepsilon_{ib} \tag{3}$$

where $distance_i$ is the distance of house i to closest sector boundary with $distance_i < 0$ for household on low tax side of boundary; $HighRate_i \equiv \mathbf{1}\{distance_i > 0\}$; δ_b are boundary fixed effects, and $\varepsilon_i b$ is a residual. We present figures using a region within 300 meters of the boundaries, while for estimation we zoom in on the region within 50 meters of the boundary.

2.3 Results

We first check for manipulation of the running variable distance to the boundary: Figure 4 shows the distribution of the distance to the boundary, and shows no evidence of bunching at 0.

We also do not see any jump in the assessed house value at the tax sector boundaries. Figure 5 shows how house values change with distance to the sector boundary.

By contrast, we see large changes in the price per square meter (P), the assessed land value ($P \times S \times L$) and the tax liability T. Figures 6 – 8

Figure 4: No Manipulation of the Running Variable

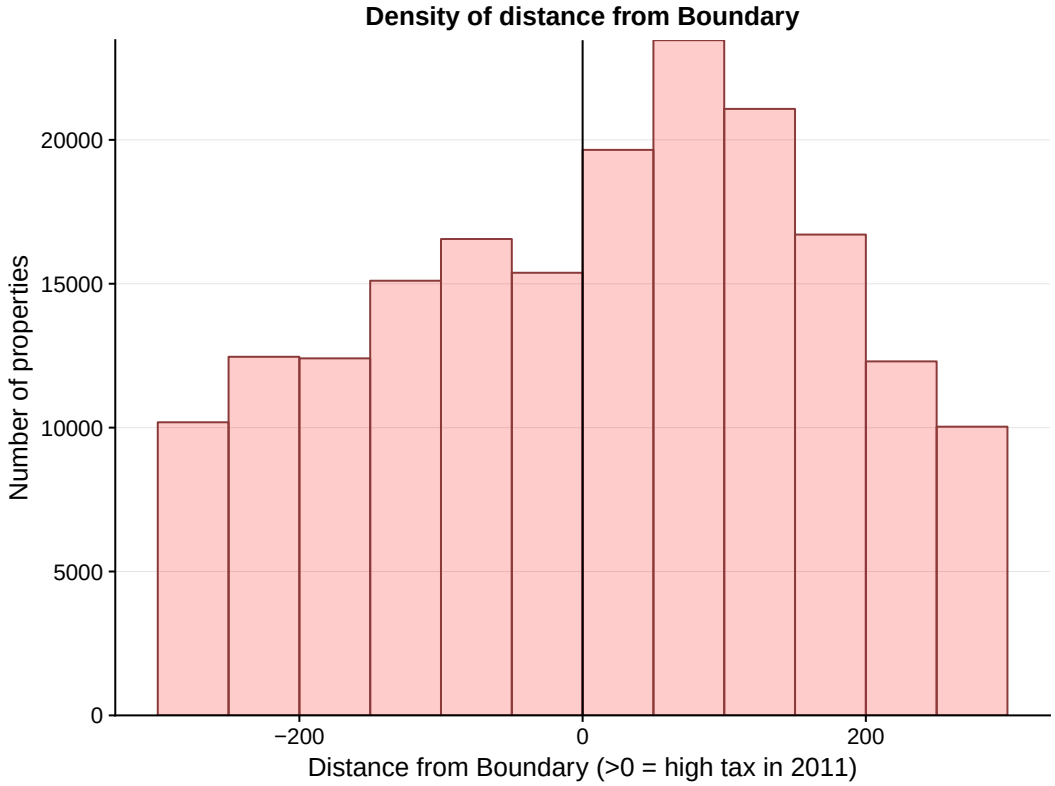


Figure 5: No Jump in House Value at the Boundary

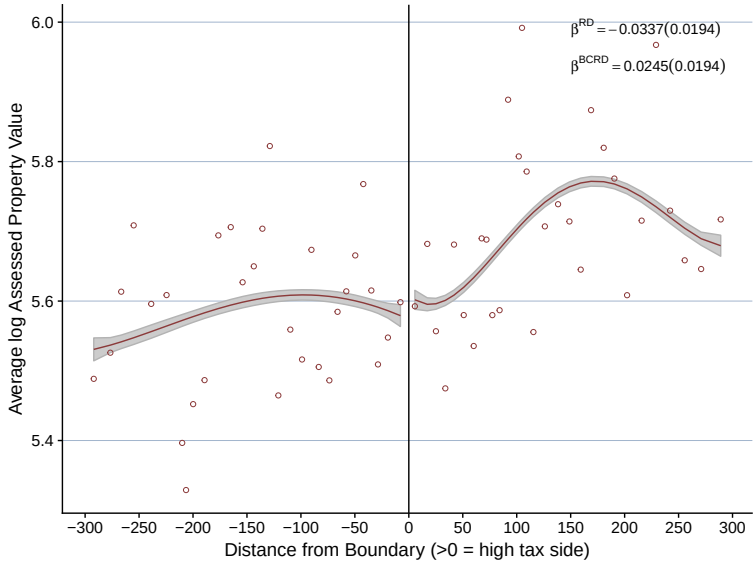


Figure 6: Price Per Square Meter Jumps at the Boundary

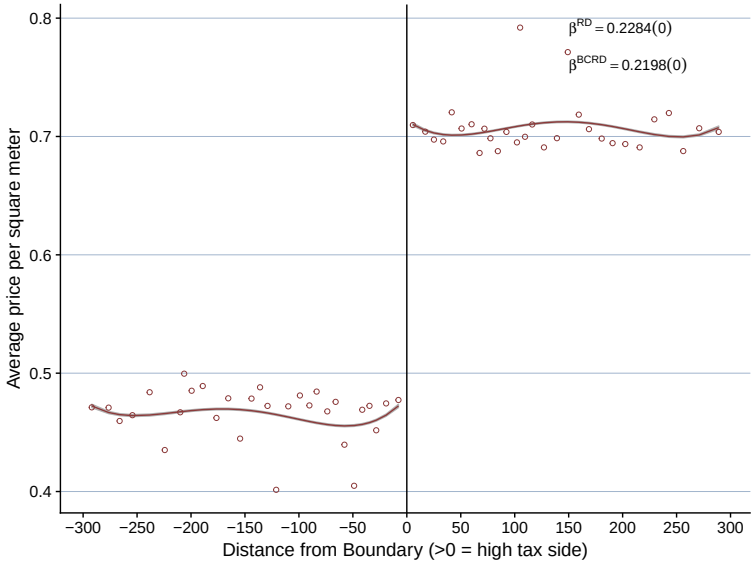


Figure 7: Land Value Jumps at the Boundary

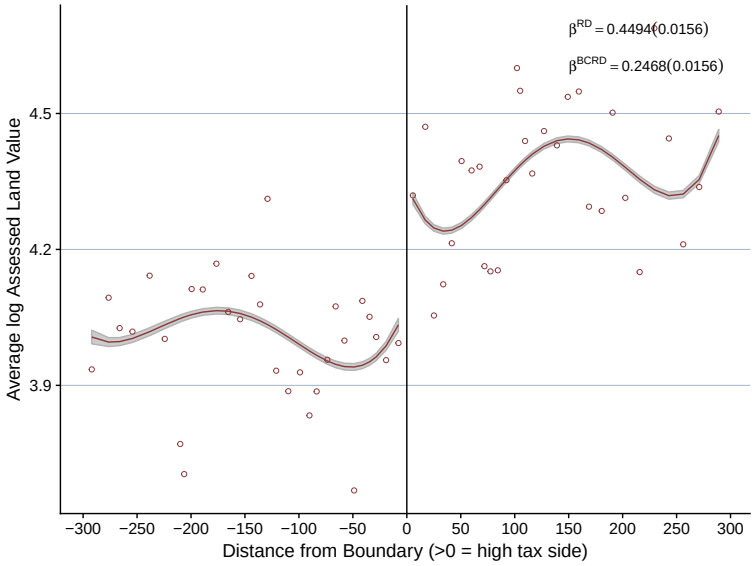
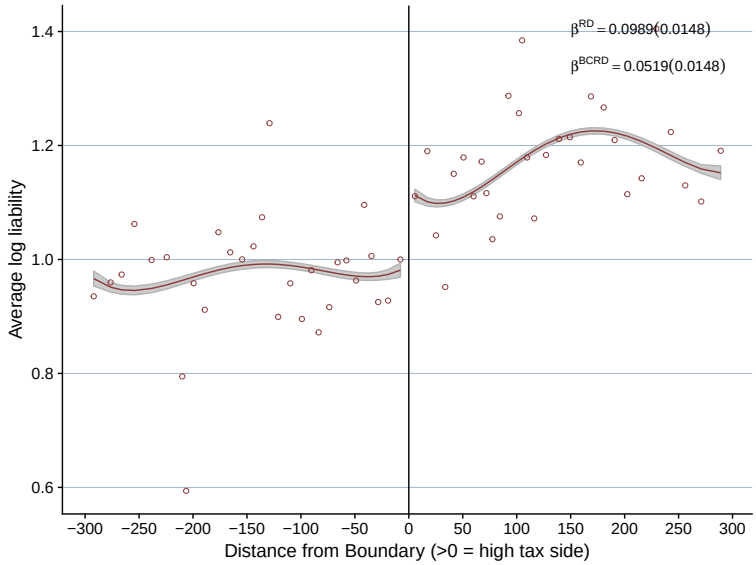


Figure 8: Tax Liability Jumps at the Boundary



Turning to outcomes, figure 9 shows that the tax jump is accompanied by a jump in delinquency at the boundaries. This is driven by extensive margin responses. Figure 10 shows that the probability of making any payment jumps down at the boundary. Nevertheless, as figure 11 shows, overall tax payment jumps up, the increase in delinquency is smaller than the increase in the tax rate, showing that the tax rates are still below the Laffer rates. Interestingly, there is no evidence that enforcement effort is targeting the high delinquency side of the boundary. Figure 12 shows that the jump in the probability of legal follow-up at the boundary is the same size as the jump in the probability of delinquency.

3 Survey Experiment Pre-analysis Plan

3.1 Research Questions

The survey experiment is designed to study the impact of two treatments. All respondents receive information about how the tax system works. Respondents in our main treatment arm then also learn that there is inequity in the system and how their own tax liability compares to their neighbors on the other side of the tax sector boundary, i.e., whether they are treated better (for those on the low tax side) or worse (for those on the high tax side) than comparable peers. The second treatment highlights instead key tools the government has available to enforce the payment of the property tax by households. This information treatment of penalties that can be applied to delinquent taxpayers will also serve as a benchmark to the fairness treatment. The survey instrument with the treatments in it can be found in our AEA RCT registry entry at https://www.socialscienceregistry.org/trials/4055

We will then measure and test for the existence of peer-comparison effects on self-reported

Figure 9: Delinquency Jumps Up at the Boundary

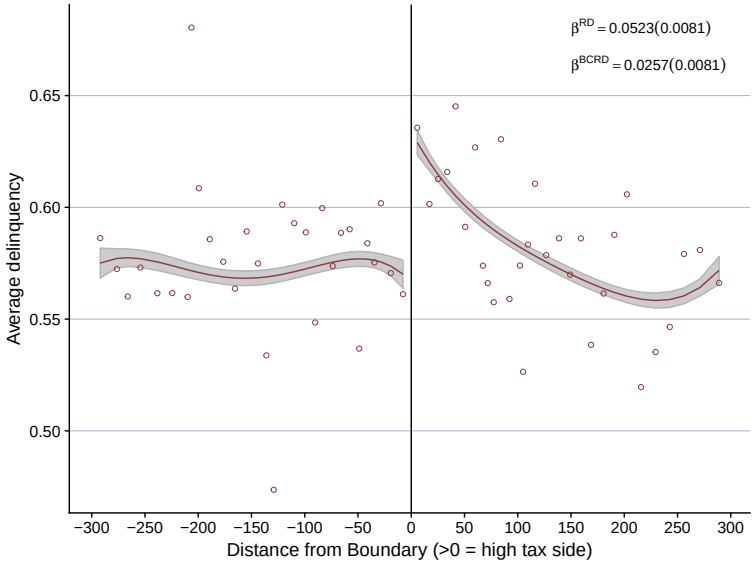


Figure 10: Extensive Margin of Payment Jumps Down at the Boundary

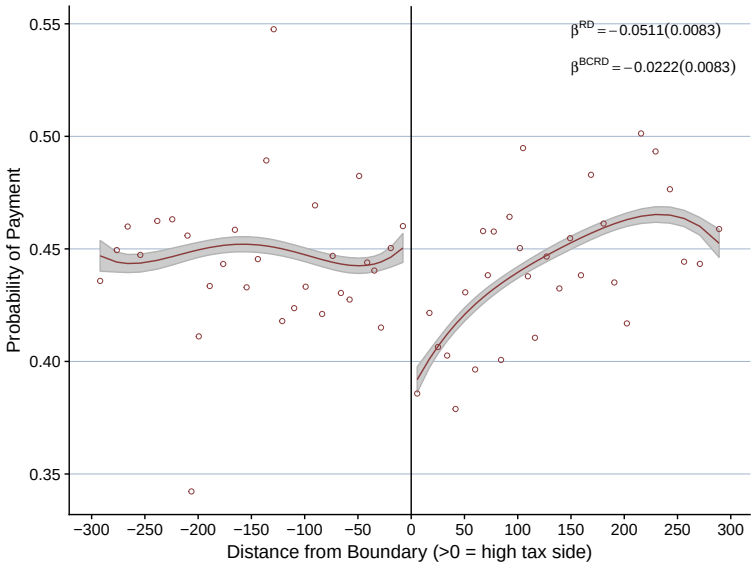


Figure 11: Extensive Margin of Payment Jumps Down at the Boundary

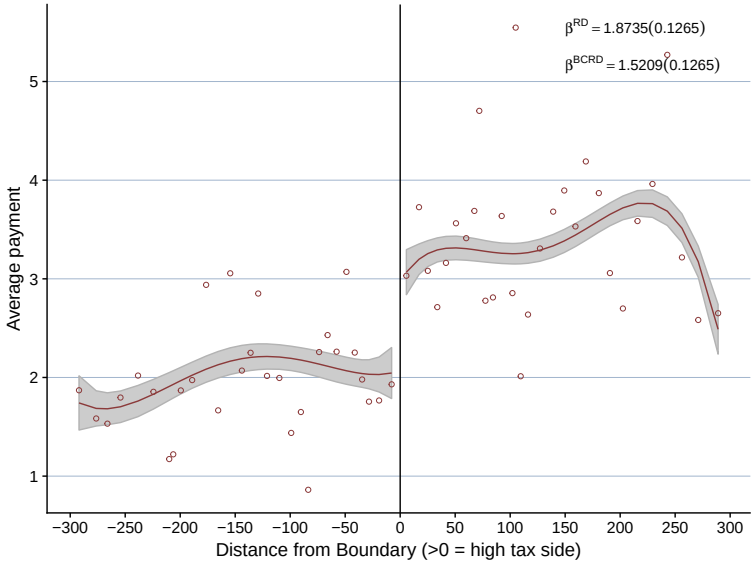
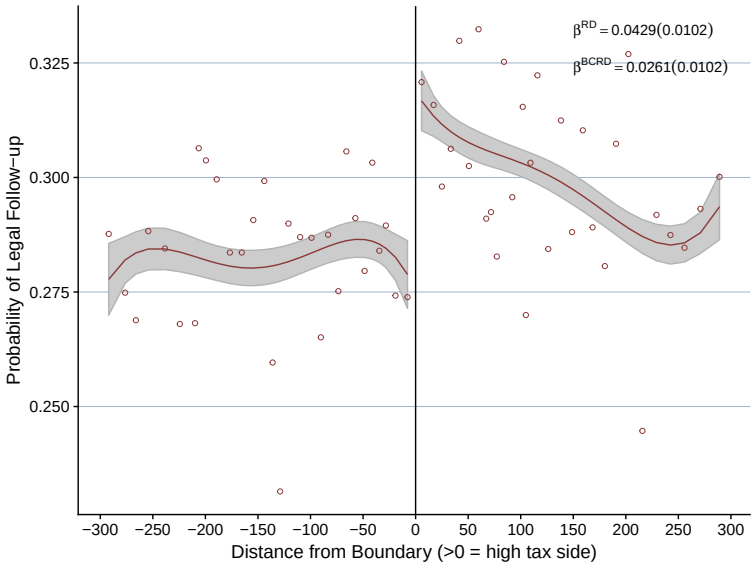


Figure 12: Extensive Margin of Payment Jumps Down at the Boundary



and actual tax compliance, and separately for advantageous equity and disadvantageous equity effects, in the taxation context. We will also test for impacts on policy preferences.

3.2 Sampling

3.2.1 Sampling Frame

There are approximately 600,000 taxable properties in Manaus. Our government partners shared their cadaster with us, anonymizing each property such that for each property, the finest level at which it could be identified is the lot the property is on. There are 456,000 lots in the data. We identified all lots on blocks that are either adjacent to the boundary of a tax sector or adjacent to a block that is adjacent to a tax boundary to form a buffer two blocks deep around the boundaries of the tax sectors. Within these blocks, we drop non-residential properties, apartment buildings (entering apartment buildings requires the superintendent's written permission, complicating surveying significantly), and blocks where the average tax change for residential properties were they to be located on the other side of their closest tax boundary is either less than 4% or less than 15 Reais. This results in a sample of 39,306 properties on 32,908 lots.

3.2.2 Assignment to Treatment

We first created 40 strata for the randomization, combining

1. 2 groups for above and below median delinquency in 2017.
2. 4 groups for quartiles of the size of the tax liability.
3. 5 groups for the size of the tax jump were residential properties located on the other side of their nearest tax boundary. For the negative jumps (people on the high tax side of the boundary) we split them into the 40% largest negative jumps (group 1), the next 40% (group 2) and the 20% smallest negative tax jumps. Similarly for the positive tax jumps (people on the low tax side of the boundary) we split them into the 40% largest positive jumps (group 3), the next 40% (group 4) and the 20% smallest positive tax jumps. The 20% smallest positive and negative tax jumps combined to create group 5.

We then randomized the lots into the 6 treatments:

1. Control survey; Short list first (in list randomization)
2. Control survey; Long list first (in list randomization)
3. Enforcement treatment survey; Short list first (in list randomization)
4. Enforcement treatment survey; Long list first (in list randomization)
5. Fairness treatment survey; Short list first (in list randomization)

6. Fairness treatment survey; Long list first (in list randomization)

with probabilities that varied with the strata depending on the tax jump group the lot was in. In the strata with small tax jumps (group 5) we randomized 50% of lots into the enforcement treatment and 50% into the control survey. In the strata with the large (positive or negative) tax jumps we randomized 20% of lots into the enforcement treatment, 40% of lots into the fairness treatment, and 40% into the control survey. These proportions were picked to maximize overall power subject to the constraint that no lots in the small tax jump group would receive the fairness treatment. Each lot was then further randomized into receiving either the short list first and then the long list, or the long list first and then the short list, for a list randomization at the end of the survey that includes two questions (each with its own list randomization; see below).

3.3 Analysis

Our main analysis of the experimental results will consist of running regressions of the form

$$y_i = \alpha + \beta_1 T_{1i} + \beta_2 T_{2i} + \mathbf{X}_i\delta + \varepsilon_i \tag{4}$$

$$\begin{aligned} y_i &= \alpha + \beta_1 T_{1i} + \beta_2 T_{2i} + \alpha_H HighTaxSide_i \\ &\quad + \beta_{1H}\left[T_{1i} * HighTaxSide_i\right] + \beta_{2H}\left[T_{2i} * HighTaxSide_i\right] + \mathbf{X}_i\delta + \varepsilon_i \end{aligned} \tag{5}$$

$$\begin{aligned} y_i &= \alpha + \beta_1 T_{1i} + \beta_2 T_{2i} + \sum_{g=1}^{4} \alpha_g \mathbf{I}\{group_i = g\} \\ &\quad + \sum_{g=1}^{4} \beta_{1g}\left[T_{1i} * \mathbf{I}\{group_i = g\}\right] + \sum_{g=1}^{4} \beta_{2g}\left[T_{2i} * \mathbf{I}\{group_i = g\}\right] + \mathbf{X}_i\delta + \varepsilon_i \end{aligned} \tag{6}$$

where y_i are our outcomes of interest for survey respondent i; T_{1i} indicates households administered treatment 1; T_{1i} indicates households administered treatment 2; $\mathbf{X}_i$ are controls (described in section 3.7 below); $HighTaxSide_i$ indicates households on the high tax side of their nearest tax boundary; $group_i$ indicates the group that the households belong to for the size of the tax jump were residential properties located on the other side of their nearest tax boundary; and ε_i is a residual clustered at the lot level.

The specification in equation (4) tests for the average effect of the two treatments compared to households in the control group. We hypothesize that the effect of the fairness treatment may depend on the size of the tax jump and on whether the household is on the low tax side (advantageous inequity) or the high tax side (disadvantageous inequity) of a tax boundary. Therefore, the specification in equation (5) tests whether the treatment effects differ for households on the high tax side of the tax jump; and the specification in equation (6) tests whether the treatment effects differ also by the size of the tax jump.

For inference we will present asymptotic standard errors clustered at the lot level, and p-values from randomization inference under the null of zero treatment effects (Young, 2018).

3.4 Primary outcomes

The primary outcomes for our main analysis can be divided in three groups:

1. Opinions about fairness of the property tax.
2. Reported intention to pay the property tax.
3. Actual payment of the property tax.

1. Opinions about fairness of the property tax. We will test the effect of the treatments on respondents' opinions regarding the fairness of the property tax system. This is measured in the survey by the respondents' level of agreement or disagreement (strongly agree, agree, neutral, disagree, strongly disagree) with the following statements:

- "The rules for calculating the IPTU liabilities for properties in Manaus are fair."
- "The amount of IPTU charged for your property is fair compared to what other properties in Manaus are charged."

2. Reported intention to pay the property tax. We will test the effect of the treatments on respondents' reported intention to pay the property tax, which is measured in two main ways in the survey. First, it is measured indirectly by the respondents' answer to the following question:

- "In your opinion, how many families with properties and property tax bills similar to yours will fully pay the property tax due this year (2019)?
  - Less than 2 families for every 10 families;
  - Between 2 and 4 families for every 10 families;
  - Between 4 and 6 families for every 10 families;
  - Between 6 and 8 families for every 10 families;
  - More than 8 families for every 10 families;"

and by the respondents' level of agreement or disagreement (strongly agree, agree, neutral, disagree, strongly disagree) with the following statements:

- "Not paying the property tax this year can be justifiable for families with a property and property tax bill similar to yours."
- "It may be justifiable for families like mine not to pay their IPTU bill this year if they are struggling to pay other bills."
- "It may be justifiable for households not to pay their property tax bill this year if their property tax bill is higher than the property tax bill of families with a similar property and neighborhood."

Second, it is measured directly by two list experiments. Each respondent is asked the number of statements that they agree with for each list. Half of respondents (randomly selected) see the first list with the 5 statements and the second list with the 4 statements, excluding the statement of interest. Half of respondents (randomly selected) see the first list with the 4 statements, excluding the statement of interest, and the second list with the 5 statements. The first list is (** indicates the statement of interest):

- "I will move to a property with a lower property tax liability."
- "I will not pay my electricity bill on time this year."
- "Property tax revenues should be used to improve the road surfaces in Manaus."
- "I will not pay my property tax bill on time this year." **
- "I have held an informal job."

The second list is:

- "I am in favor of lowering the property tax rates."
- "I will not pay my water bill on time this year."
- "I have made a purchase without asking for a VAT receipt."
- "I will not pay my property tax bill on time this year." **
- "I will protest formally to the municipality about the property tax liabilities in Manaus."

3. Actual payment of the property tax. We will test the effect of the treatments on respondents' actual payment of the property tax. For this purpose, we will merge survey data with administrative data on actual tax payments for 2019 for all properties in each lot. Non-compliance mostly takes place at the extensive margin in Manaus (paying in full vs. not paying at all). So we will construct a variable capturing the share of households that did not pay their property tax in full in 2019 at the lot level.

For this outcome we will lose power when estimating equations (4)–(6) for two reasons. First, this outcome is observed only in the administrative data, which we can merge with the survey responses only at the lot level. Second, the decision to pay taxes is made after the survey is over, potentially several weeks or even months later, so the treatment effects may have decayed. Moreover, if people discuss the survey with their neighbors, there may be spillovers of the treatments into the survey's control group.

Therefore, to complement the analysis in equations (4)–(6), for this outcome we will use a difference in differences approach. To do this, we will construct a pure control group of properties that are one or two blocks further away from the tax boundaries than the survey

respondents. These properties should be broadly comparable to our survey respondents. We can then compare the evolution of tax compliance in this pure control group to its evolution amongst our survey respondents.

3.5 Secondary outcomes

In addition to our primary outcomes, we will study impacts on secondary outcomes:

1. Opinions about penalties for non-compliance.
2. Preferences about criteria to use to set property tax liabilities.
3. Beliefs about overall level of compliance with the IPTU and its responsiveness to tax rates.

1. Opinions about penalties for non-compliance. We will test the effect of the treatments on respondents' opinions regarding the penalties for non-compliance with the property tax. This is measured in the survey by the respondents' level of agreement or disagreement (strongly agree, agree, neutral, disagree, strongly disagree) with the following statement:

- "The penalties for taxpayers who do not pay their property tax are very high."

2. Preferences about tax liability criteria. We will test the effect of the treatments on respondents' preferences regarding the criteria to use to set property tax liabilities. This is measured in the survey by the respondents' answer to the following question (respondents were asked to rank these 5 items by order of importance):

- "If you could decide the property tax calculation criteria, which ones would be more important?
  - Taxpayers who receive the same services from the municipality government must pay the same amount.
  - Taxpayers who have the same income must pay the same amount.
  - Taxpayers who are neighbors must pay the same amount.
  - Taxpayers with buildings of the same size (square meters) must pay the same amount.
  - Taxpayers with land area of the same size (square meters) must pay the same amount."

3. Beliefs about overall level of compliance with the property tax and its responsiveness to tax rates. We will test the effect of the treatments on respondents' beliefs about the overall level of compliance with the property tax and its responsiveness to tax rates. This is measured in the survey by the respondents' answer to two questions:

- "In your opinion, in general, how many families will fully pay the property tax due this year (2019)?
  - Less than 2 families for every 10 families;
  - Between 2 and 4 families for every 10 families;
  - Between 4 and 6 families for every 10 families;
  - Between 6 and 8 families for every 10 families;
  - More than 8 families for every 10 families."
- "What if the value of the property tax was cut in half? (for example, an property tax bill of 500 reais would be 250 reais)
  - Less than 2 families for every 10 families;
  - Between 2 and 4 families for every 10 families;
  - Between 4 and 6 families for every 10 families;
  - Between 6 and 8 families for every 10 families;
  - More than 8 families for every 10 families."

3.6 Non-compliance

There are three potential sources of deviation from perfect compliance with our intended experimental protocols.

First, as this is an unincentivized household survey in a large urban area, many households may refuse to participate. Second, our enumerators were given address lists and maps with the properties in which to survey households, but may have deviated from the treatment assigned to each property. Third, enumerators were paid per completed survey and may, in some cases, have rushed through the survey. We discuss each of these in turn.

3.6.1 Non-Response

Our survey is relatively short (we expected it to take around 15 minutes to administer) but our surveyors arrived unannounced, and respondents are not compensated for their participation. Moreover, Manaus is a large city and crime risk is high, so we expected many households to be unwilling to take part in the survey.

To deal with the potential sample selection this introduces we will

1. check that the refusal rate is balanced across the treatment arms;
2. use our rich covariates (described in section 3.7 below) to predict refusal and control for observable determinants of refusal;
3. use the observables to predict refusal rates and reweight the sample to ensure representativity.

3.6.2 Incorrect Assignment

The randomization was conducted by us before the surveyors were sent into the field. In the field, the enumerators had an address list with each property's lot code and address, and a map of the area they were surveying with the lots they were to visit highlighted on it. Before starting the survey, the enumerators had to type the lot code into their tablet such that the survey that lot was assigned to appeared on their screen for them to administer. Yet, the treatment actually administered to each household may deviate from the intended treatment for two reasons. First, enumerators may type the wrong lot code into the tablet. This can occur because they misread the lot code on the address list or misread the map/address and were interviewing a different household. Second, the administrative cadaster of properties the government provided us is imperfect. It has lot codes for each property in it and the government also provided us with geographic data on the lots; however, not all the lots in the cadaster appear in the geographic files. We assigned treatments based on the lots in the cadaster and then matched these lots as best as we could to the lots that appear in the geographic files the maps are built from. The enumerators were given clear instructions to use the lot code that appears on the address list not the lot code shown on the maps. Nevertheless, there may be instances in which the incorrect lot code is used.

To deal with both of these types of incorrect assignments, we will employ an IV strategy in which we instrument for the actual treatment with the assigned treatment. This is made possible because in the survey the enumerators will record the address of the property they are at, allowing us to match the survey responses to the address list, and in turn, their intended treatments.

3.6.3 Enumerator Rushing

The survey company we are working with pays their enumerators a flat fee per completed survey (subject to some basic quality requirements). This means that enumerators may have an incentive to rush through the survey, leading to a degradation of the quality of the survey administration. This is particularly important if they rush through the treatment sections of the survey. The survey data comes with time stamps for the the start- and end-times of the survey and for each section of the survey. This allows us to measure the amount of time the enumerator spent on each section. In particular, it allows us to identify outliers in terms of the speed with which the sections are administered. Since these very short surveys are unlikely to be correctly implemented surveys we will drop outliers in terms of the duration of the full survey and the duration of the treatment sections. We will also show robustness to using different trimming thresholds.

3.7 Control Variables

We will control for baseline covariates in three ways. First, we will only control for the strata related to the size of the tax jump (5 groups). Since the treatments are randomized, all covariates that predict our outcomes are balanced in expectation and so controls are not necessary for identification, only for power. However, the treatment probabilities were different in the strata containing households facing small tax jumps so we must control for these strata (Athey & Imbens, 2017). Second, we will control for the 40 strata used in the randomization based on lot-level delinquency (2), size of the tax liability (4), and size of the tax jump (5). Third, we will control for the randomization strata and also a set of observable control variables to increase statistical power. The set of potential controls we will pick from is:

1. Household Demographics (from the survey)
   (a) Respondent age (question R.1)
   (b) Respondent gender (question End.3)
   (c) Respondent education (question R.2)
   (d) Respondent sector of occupation (question R.3)
   (e) Household size (question R.4)
   (f) Household income (question R.5)
   (g) Number of bedrooms (question D.1)
   (h) Property size in square meters (question D.2)
   (i) Approximate rental value (question D.3)
   (j) Respondent views on the quality of public services (questions B.1, B.2, B.3, B.4, B.5 for water, sewage, trash collection, road, public lighting, respectively)
   (k) Respondent views on the quality of municipal governance (question B.6)
   (l) Respondent consideration of property tax when decided to buy property (question A.7)
   (m) Respondent beliefs about property tax bills for similar properties in the same neighborhood (question A.12)
   (n) Respondent beliefs about tax delinquency rates for the property tax last year (question A.13)
   (o) Respondent beliefs about the consequences of tax delinquency (question A.14)
   (p) Respondent knowledge about the criteria used to set the property tax bills in Manaus
2. Lot-level Average Property Characteristics (from the administrative data)
   (a) Completeness of registered addresses

(b) Completeness of ownership details
(c) Registration date
(d) Dates of updates to cadaster (where available)
(e) Inspection type/date (where available)
(f) Ownership type (individual, religious, non-profit) (where available)
(g) Position on street block
(h) Use of land (construction, abandoned, built-up, parking, etc) (where available)
(i) Topography of the plot (where available)
(j) Flood risk (where available)
(k) Delimitation type (fencing, wall, etc) (where available)
(l) Paving (where available)
(m) Drainage (where available)
(n) Square-footage (where available)
(o) Subsoil (where available)
(p) Terracing (where available)
(q) Type of building (temporary, house, apartment, commercial property, gas station etc) (where available)
(r) Structural construction material (concrete, masonry, wood etc) (where available)
(s) Facade construction material (where available)
(t) Roof material (where available)
(u) Wall material (where available)
(v) Building standard (luxury, medium, economy) (where available)
(w) State of conservation (where available)

3. Lot-Level Past Tax data (from the administrative data)

(a) Presence on active debtors list
(b) Size of outstanding debt (if any)
(c) Amount of fines owed
(d) Previous late payment fines/interest
(e) Size of tax liability
(f) Changes in tax liability over time
(g) Previous payments in one payment/in instalments
(h) Previous payments

We will also include all two-way interactions of these variables.

To pick which variables in this extremely large set of controls to use, we will employ appropriate methods. Exactly what the appropriate method is, is a very active area of research, and so we will use the latest methods when we perform this analysis. At the time of writing these include Belloni *et al.* (2014, 2017); Athey *et al.* (2018); Chernozhukov *et al.* (2018).

3.8 Heterogeneity

Finally, we plan to test for heterogeneity in treatment effects for the primary outcomes by

1. Approximate rental value (using question D.3)
2. Tax liability (used for stratification)
3. Tax liability relative to household income (using question R.5)
4. Tax liability relative to approximate rental value (using question D.3)
5. Tax liability relative to monthly water and electricity bills (using questions A.5 and A.6)
6. Prior tax compliance (used for stratification)
7. Respondent views on the quality of public services (aggregating answers to questions B.1, B.2, B.3, B.4, B.5)
8. Respondent views on the quality of municipal governance (question B.6)
9. Respondent consideration of property tax when decided to buy property (question A.7)
10. Respondent beliefs about property tax bills for similar properties in the same neighborhood (question A.12)
11. Respondent beliefs about tax delinquency rates for the property tax last year (question A.13)
12. Respondent beliefs about the consequences of tax delinquency (question A.14)
13. Respondent sector of occupation (question R.3)

References

ATHEY, SUSAN, & IMBENS, GUIDO. 2017. The Econometrics of Randomized Experiments. *Chap. 3, pages 73–140 of:* BANERJEE, ABHIJIT, & DUFLO, ESTHER (eds), *Handbook of Field Experiments*, vol. 1. Elsevier.

ATHEY, SUSAN, IMBENS, GUIDO W., & WAGER, STEFAN. 2018. Approximate Residual Balancing: Debiased Inference on Average Treatment Effects in High Dimensions. *Journal of the Royal Statistical Society: Series B*, **80**, 597–623.

BELLONI, ALEXANDRE, CHERNOZHUKOV, VICTOR, & HANSEN, CHRISTIAN. 2014. Inference on Treatment Effects after Selection among High-Dimensional Controls. *Review of Economic Studies*, **81**, 608–650.

BELLONI, ALEXANDRE, CHERNOZHUKOV, VICTOR, FERNANDEZ-VAL, IVAN, & HANSEN, CHRISTIAN. 2017. Program Evaluation and Causal Inference With High-Dimensional Data. *Econometrica*, **85**, 233–298.

CHERNOZHUKOV, VICTOR, CHERTVERIKOV, DENIS, DEMIRER, MERT, DUFLO, ESTHER, HANSEN, CHRISTIAN, NEWEY, WHITNEY, & ROBINS, JAMES. 2018. Double/Debiased Machine Learning for Treatment and Structural Parameters. *The Econometrics Journal*, **21**, C1–C68.