ECONOMIC DEVELOPMENT & INSTITUTIONS

EDI WORKING PAPER SERIES

MEASURING THE POTENTIAL FOR PROPERTY TAXATION IN DAKAR, SENEGAL*

Justine Knebelmann
Paris School of Economics, EHESS
Victor Pouliquen
University of Oxford
Bassirou Sarr
EHESS
Denis Cogneau
Paris School of Economics, IRD, EHESS
Marc Gurgand
Paris School of Economics, CNRS, ENS

February 2021

Oxford Policy Management
UKaid from the British people
UNIVERSITÉ DE NAMUR
PARIS SCHOOL OF ECONOMICS ÉCOLE D'ÉCONOMIE DE PARIS
ADE ANALYSIS FOR ECONOMIC DECISIONS

Abstract

Property taxes are in theory easy to enforce in their simplest form due to their tangible tax base, and are considered an equitable means to raise revenue in low-income countries. In spite of these features, African countries, where cities are growing at an unprecedented pace, are raising only 2 percent of fiscal revenue in property taxes, against around 9 percent in OECD countries, and this figure is at 0.3 percent in Senegal. Focusing on Dakar, the capital city where real estate has been buoyant over past decades, we document the extent and nature of the property tax gap. Using administrative and survey data, as well as satellite images and property valuation methods, we estimate that less than 20 percent of property owners are in the tax net, and that only 9 percent of tax potential is being collected. This weak performance is put into historical perspective using colonial archives. Finally, we compare the observed distribution of the tax burden with the theoretical one under full compliance, and find that weak enforcement leads to a tax profile that is more regressive than what is provided for in the legal framework. These results reinforce the justification for reform and modernization of the property tax management system.

About Economic Development & Institutions
Institutions matter for growth and inclusive development. But despite increasing awareness of the importance of institutions on economic outcomes, there is little evidence on how positive institutional change can be achieved. The Economic Development and Institutions – EDI – research programme aims to fill this knowledge gap by working with some of the finest economic thinkers and social scientists across the globe.

The programme was launched in 2015 and will run until 2021. It is made up of four parallel research activities: path-finding papers, institutional diagnostic, coordinated randomised control trials, and case studies. The programme is funded with UK aid from the UK government. For more information see http://edi.opml.co.uk.

UKaid
from the British people

* Corresponding author: justine.knebelmann@psemail.eu. We thank the Senegalese Tax Administration – Direction Générale des Impôts et Domaines – for excellent collaboration. We specifically thank Serigne Mabeye Fall and Ousseynou Niang. We also thank the Treasury for its collaboration. We thank Economic Development and Institutions for funding this project. This project was funded with UK Aid from the UK government. We also thank additional funders: International Growth Centre, International Centre for Tax and Development, J-PAL Governance Initiative, J-PAL DigiFi, EUR-PGSE. We thank the Centre National d'Etudes Spatiales for making high resolution satellite images accessible through their Geosud program. For excellent contributions to project activities and research assistance, we thank Papa Daouda Diene, Ndeye Rokhaya Diao, Mor Tacko Hane, Adrien Ciret, Aramata Badji, Oumar Fara Diop, Nicolas Orgeira, Samuel Allain. We thank participants of the EDI conferences, the PSE Applied Economics and CFDS seminars, the International Centre for Tax and Development congress, the African Property Tax Initiative congress, and the National Tax Association conference for useful comments. All errors are our own.

1 Introduction

Developing countries still lag behind their wealthier counterparts in terms of fiscal capacity, understood as the State's ability to successfully register taxpayers, assess tax liabilities, and collect tax revenues. Property taxes are no exception: they yield 2 percent of total tax revenues in Sub-Saharan Africa, against 8 to 10 percent in OECD countries (Franzsen & McCluskey, 2017), and the poor performance of property taxes in various contexts across low- and middle-income countries has been documented (Weigel, 2020; Khan *et al.* , 2016; Brockmeyer *et al.* , 2020).

However, this seems to contradict the theoretical characterization of this tax instrument as relatively easy to enforce under its simplest form. Indeed, the tax base for recurrent taxes on built property is both tangible and immobile – making it easier to apprehend than income or revenue streams. In this respect, among tax instruments still in use today, the property tax stands out as having an early appearance – in ancient Rome (*Ostarium*), in France in the aftermath of the 1789 Revolution (Piketty, 2020). In former French colonies in Africa, the property tax was introduced as early as the mid-nineteenth century.

As of today, property taxes in Africa are considered to offer not only a significant untapped potential ("a doubling of property tax revenues over present levels in Africa might be feasible", Franzsen & McCluskey (2017)), but also positive consequences in terms of equity and redistribution, and local accountability (Monkam & Moore, 2015; Weigel, 2020). Property taxes therefore appear as a low-hanging fruit for developing countries with weak administrative capacity.

Understanding why and how the property tax strongly under-performs in the developing world, in spite of these positive features, as well as assessing the effects of this weak enforcement on tax revenues and on tax equity therefore appears as a crucial question with strong policy implications. In this paper, we open the black box of property taxation in the region of Senegal's capital city, Dakar: we first quantify the property tax gap, and assess the implications of the weak enforcement on the progressivity of the tax. We investigate the economic and geographic determinants of property tax enforcement, and finally, we put the current weak performance into historical perspective. We rely on administrative data, on survey data and satellite images collected for the purpose of the study, and on historical archives. We complement the data analysis by thorough qualitative and observational insights on the institutional hurdles.

Senegal performs relatively well in terms of overall fiscal capacity compared to other countries of the region – its tax to GDP ratio is 15.1 percent against 14 percent on average in Sub-Saharan Africa, excluding natural resource revenue,[1] – but recurrent property taxes account only for 0.3 percent of total taxes, and 10 percent of local governments' tax revenues.[2] Yet, property taxation has existed in Dakar since the early colonial period, and the city grew precociously, being the capital of French West Africa. Following strong decentralization policies since independence[3] local governments

[1] 2019 Article IV IMF Report, figure for 2018.

[2] Authors' calculations using administrative data. Property taxes exclude taxes on firms' real estate. Local governments' revenues including all taxes and fees, excluding transfers from the central government.

[3] Decentralization reforms were enacted in 1972, 1996, and most recently through the *Acte III de la décentralisation* in 2013.

are now responsible for an increasing share of public services, supposed to be funded by local taxes and notably the property tax. Dakar's current population is 3.7 million[4], the real estate sector has been booming in the recent years, as the city grows at an unprecedented pace and the country's stability attracts investors. Based on the fiscal legislation, property tax revenues should immediately follow increases in built units, and in rental values. However this progress has not been observed, suggesting an ever-growing untapped potential.

Against this backdrop, we rigorously document the extent of the property tax gap in Dakar. The valuation roll includes entries for approximately 54,500 properties, while the estimated number of plots in the region is 374,000. The Treasury collects only 27 percent of assessed liabilities. An advantage of the setting is that thanks to the specific feature of the property tax (its visible tax base) we are able to precisely estimate the revenue potential, and to measure the tax gap on the extensive as well as the intensive margins. We do this using data from a survey conducted in a subset of areas in the region of Dakar with tax potential, comprising over 70,000 plots. We find that less than 20 percent of property owners are in the tax net, defined as receiving a tax notification and/or paying the property tax. Only 9 percent of potential revenue is collected, 75 percent is foregone due to dysfunction on the extensive margin, and 16 percent is foregone due to dysfunction on the intensive margin. Total potential revenues under full compliance would amount to XOF 9.7 billion in these areas of the capital city. In contrast, total collections for the whole of Senegal in 2019 are at XOF 3.8 billion (USD 16.5 million against USD 6.5 million).

Furthermore, we investigate whether property taxation in Dakar (as it is currently functioning) is progressive or regressive. The proportionality of property taxation in France – where the legal system resembles the one in Senegal because of colonial legacies – and in the United States has been criticized in the public finance literature for being insufficiently progressive (Piketty, 2020; Saez & Zucman, 2019). In Senegal, an abatement for owner-occupied properties makes the legal taxation profile slightly more progressive. Moreover, considering that the administration has limited resources, and does not tax all properties, it could be that it focuses efforts on the wealthiest, thus reinforcing the progressivity. If so, the tax could turn out to be even more progressive than what is provided for in the legal framework. We show that the facts contradict this hypothesis: although the share of property owners who pay the tax increases with property value, the overall effective tax rate is almost flat across the distribution of property values suggesting that there isn't a stronger fiscal pressure on higher valued properties. In fact, the gap between the effective rate and the theoretical rate is significantly wider for top deciles of property value. Furthermore, the effective tax rate faced by $payers$ is lower in the top deciles of property value. This has important consequences for the fairness of the tax.

We provide suggestive evidence on economic and geographic factors that explain variations in property tax performance. Finally, to determine whether there is a colonial legacy of poor performance of property taxation, or alternatively strong disruptions at the time of independence, we use novel evidence on property taxation in Dakar throughout the colonial and post-colonial pe-

[4]Source: Agence Nationale de la Statistique et de la Démographie, 2019.

riod, to put current fiscal capacity into perspective. We find that property tax capacity was quite significant and higher than nowadays throughout the colonial period, peaked in the 1930s and started declining before independence.

This paper speaks to the literature investigating the capacity to tax (property) in developing countries (Bahl *et al.* , 2008; Khan *et al.* , 2016; Brockmeyer *et al.* , 2020; Franzsen & McCluskey, 2017; Weigel, 2020; Balan *et al.* , 2020; Zebong *et al.* , 2017; Kopanyi, 2015). The first section provides a tax gap analysis, conceptually similar to those produced by the IMF or the IRS, for instance. Combining administrative and survey data allows not to be restricted to the universe of formalized or registered economic agents, as is often the case in the literature on taxation in developing countries relying on administrative data. Several of these previous studies provide a useful benchmark, confirming that Dakar performs specifically poorly in property taxation with respect to what could be expected based on the country's and city's characteristics. Both in Kampala, Uganda (Manwaring & Regan, 2019) and in the Punjab province, Pakistan (Khan *et al.* , 2016), the share of properties that are registered is higher than 80 percent – against 15 percent in Dakar. Kampala is comparable in size with Dakar, but the country's overall tax to GDP ratio is 5 percentage points lower than Senegal (13 percent). In the Kasai province in the DRC, Weigel (2020) documents an increase in the share of compliant owners from 0.1 to 11 percent - the final figure is not far from what we observe in Dakar, although in Dakar, the tax potential is manifold higher and the administration is less resource constrained. In Monrovia, Liberia Okunogbe (2020) documents a situation that is comparable to Dakar in terms of poor property tax capacity, in some areas the rate of registration is as low as 5 percent.

Second, this paper contributes to the literature discussing property tax progressivity and fairness. Piketty (2020) and Saez & Zucman (2019) argue that property tax systems in developed countries (France, United States) are not progressive enough.[5] The main finding of this paper is that in this developing country context (where the legal framework is similar to the one prevailing in France), the very weak enforcement creates a situation that is even $more$ regressive than what would be expected based on the tax code: the share of foregone revenue is larger in upper deciles of the property value distribution.[6] Results on the distribution of the burden of property tax in developing countries are still relatively scarce. Our findings speak to the gap in empirical evidence mentioned in Norregaard (2013), and resemble findings in de Carvalho Jr. (2015) which show that the property tax burden in Brazil is progressive when considering the share of owners who pay (which increases with property value), but regressive when considering effective taxation condi-

[5] Saez & Zucman (2019); Piketty (2020) suggest that the tax rate should increase with the wealth of the taxpayer instead of a proportional tax, and that all types of wealth and debt should be considered together

[6] An important dimension of property tax incidence is the split of the tax burden between owners and tenants – theoretical results are divided into three strands, the traditional view, the new view and the benefits view (see Oates & Fischel (2016) for a summary and Sennoga *et al.* (2007); Norregaard (2013) for applications to developing countries). In this analysis, we assume that the tax is borne by owners, and we analyze progressivity with respect to $owners'$ wealth and income. This can be rationalized by the results in Sennoga *et al.* (2007) for developing countries – where the authors show theoretically that in the context of underdeveloped property markets, the tax burden is most likely to fall on the owner. We also provide results that contradict the benefit view – according to which property taxes paid are equivalent to a fee for the utilization of local public services – in the context of Dakar.

tional on payment.[7]

Third, this paper speaks to studies that shed light on how the actual tax burden may differ from what is provided for in the legal framework under weak or heterogeneous enforcement. Bachas *et al.* (2020) show that the VAT is regressive in theory, but that if informality is taken into account it appears to be progressive instead. Brockmeyer *et al.* (2020) and Weigel (2018) show that liquidity constraints are an important factor of property tax (non)compliance - thus creating a disconnect between the tax base (property value or wealth) and the effective tax rate paid. Similarly, Balan *et al.* (2020) document that depending on the identity of tax collectors, owners may be targeted more according to their capacity to pay, rather than the value of their property. Khan *et al.* (2016) in Pakistan and Avenancio-Leon & Howard (2019) in the United States both shed light on distortions in targeting and valuation in property taxation.

Finally, this paper contributes to the strand of literature providing a quantitative analysis of fiscal capacity over the long run, more specifically in former colonies (Cogneau *et al.* , 2020a,c; Frankema & van Waijenburg, 2014; Albers *et al.* , 2020). Both the historical and the contemporaneous analysis can be related to the literature investigating determinants of property tax capacity at the district or country level (Norregaard, 2013; Awasthi *et al.* , 2020). The evidence provided in the context in Dakar helps shed light on why property tax capacity remains so low in developing countries.

Section 2 depicts the contemporary institutional setting for property taxation in Dakar, and highlights the main impediments to its functioning. Section 3 describes the data used in subsequent sections. Section 4 displays results on the property tax gap. Section 5 studies the implications of weak enforcement for the progressivity profile of the tax. In Section 6, we provide suggestive evidence on correlates of property tax enforcement. Section 7 provides a historical perspective on property tax capacity throughout the colonial and post-colonial period. Section 8 concludes.

2 Institutional setting

2.1 Country and city context

Fiscal revenues in Senegal

Tax revenues in Senegal amount to 15.1 percent of GDP – a relatively good performance with respect to Sub-Sahara Africa counterparts.[8] Total revenues accruing to local governments amount to XOF 100.9 billion (USD 171 million), XOF 57.8 billion when excluding transfers from the central government and decentralization funds – as an order of magnitude, this corresponds to 2.9 percent of national tax revenues.[9] Recurrent property and assimilated tax assessments are a small

[7] The reasons for this distortion mentioned in de Carvalho Jr. (2015) are: outdated assessment values, lack of comprehensive cadastral mapping, and political resistance at the municipality level.

[8] Source: 2019 Article IV IMF Report, figure for 2018, corresponding to XOF 2,016 billion. The main sources of tax revenue are: taxes on goods and services (48 percent of total tax revenues) taxes on income, profits and capital gains (30 percent of total), taxes on international trade (15 percent of total) (IMF, 2019).

[9] Source: Data shared by the Ministry of Finance, figure for 2018. Local tax revenues are not systematically included in total tax figures computed by the IMF.

contributor as of today. Total *assessments* amount to XOF 23 billion nationally (USD 39 million), of which XOF 20 billion in the region of Dakar.[10] However assessments are not paid in full, and total property and assimilated tax *revenues* amount to XOF 6.3 billion (USD 11 million) – only 27 percent of assessments – corresponding to 10 percent of the tax revenues of local governments.[11]

In spite of low revenue collections, property taxes are crucial in the Senegalese context for two main reasons. First, in a country where only an estimated 4 to 20 percent of working age population is in formal employment (Mbaye, 2019), for many citizens, property taxes are the sole direct tax they are subject to. 2013 census data indicates that 73 percent of households own the property they live in (Agence Nationale de la Statistique et de la Démographie, 2020). Furthermore, property taxes, as other local taxes, have received increased political attention in recent years. A set of major decentralization policies has been passed in 2013,[12] which convey more responsibilities to municipalities. To a large extent, they are supposed to fund the public goods they provide with property tax revenues. The revenues of local governments are very low and property taxes offer a large untapped potential. Hence, property taxes are nowadays high in the government's agenda.

Real estate in Dakar

Dakar is the fifth largest agglomeration in West Africa. In recent years, the real estate sector has been particularly vibrant – see excerpts from specialized outlets in Appendix A.1, Figures A.1 and A.2. The real estate sector grew by 8.9 percent in 2018, the price of one square meter of land was multiplied by 2.5 between 1994 and 2000, by 2 between 2000 and 2009, and property prices have grown by 256 percent between 1994 and 2010. Yields (around 7 to 10 percent, measured by the ratio of rent to purchasing price) are very high by international standards and property prices are superior than those prevailing in comparable cities – eg, rents in downtown Dakar are almost twice as high as in the central business district of Abidjan (Centre for Affordable Housing Finance in Africa, 2019; Global Property Guide, 2014; Agence Nationale de la Statistique et de la Démographie, 2012). This suggests an important potential for property tax revenues.

The situation is similar in other large African cities – as is the pressing need to improve the taxation of this urban wealth (Ali *et al.* , 2017; LaSalle, 2015). However some of Dakar's idiosyncratic features which contribute in explaining why wealth creation in its real estate sector is notably strong are its political stability in comparison to neighboring countries, thus attracting investors (individual property buyers but also international actors choosing Dakar as a regional capital); the geographical limitations to city sprawling since Dakar is a peninsula, population growth leading to a high demand for housing.

[10] Own calculation using administrative data. This excludes: taxes on real estate transactions, taxes on rental income, taxes on property owned by firms that are due under another tax head.

[11] The tax that contributes the most to local revenues is the local business tax. Source for revenue collections: Treasury.

[12] *Acte III de la Décentralisation, 2013* reshapes the roles and responsibilities of local governments, and at the same time requires that they assist the national tax administration in enforcing local taxation. The PACASEN *Projet d'Appui aux Communes et Agglomération du Sénégal* is a 2019-2023 project co-funded by international donors seeking to increase local tax revenues and strengthen municipalities' capabilities.

2.2 Institutional framework for property taxation

The property tax and the garbage tax (that we classify as an assimilated tax)[13] are managed at the national level, and the fiscal rules are the same throughout the country. The valuation roll is established and maintained by the national tax administration (*Direction Générale des Impôts et Domaines*, henceforth, DGID), responsible for registration and assessments. The distribution of tax notifications, the management of payments, and enforcement actions are carried out by the Treasury. At the end of the fiscal year, the Treasury remits revenues to municipalities. Although the tax administration and the Treasury are national administrations, they have multiple geographical offices within the region of Dakar, and each office is responsible for the territory it covers. The region of Dakar is divided into four cities – Dakar, Pikine, Guédiawaye and Rufisque. Each city is further divided into communes, and both levels of local government recoup a share of property tax revenue.[14]

Property owners are required to make annual declarations at the tax administration, before the end of January of each year, where they report their properties and corresponding values. Based on these declarations, on verification activities, and on additional property census activities in the field to detect non-compliant owners, the tax administration prepares the valuation roll. The value of interest (the tax base) is the rental value of the property – the value that could be obtained if the property was rented at market prices[15] – and is either declared by the owner, or assessed by the administration by comparison with similar properties, if the reported value is deemed incorrect or was not declared. The valuation roll and printed tax notifications are shared with the Treasury around March or April of each year. The Treasury offices start distributing in May.[16] Owners make payments by visiting the Treasury offices in person.

Both the property tax and the garbage tax are due by the owner of the property. The tax liability for the property tax is 5 percent of the annual rental value of the property. If the property is owner-occupied and is the main residence of the owner, there is an abatement reducing the tax base by XOF 1.5 million. If the annual rental value of an owner-occupied property is below XOF 1.5 million XOF there is no tax due under the property tax. Other exemptions apply, although to a much rarer extent. The tax liability for the garbage tax is equal to 3.6 percent of the annual rental value of the property, and there is no abatement for owner-occupied properties.[17] Total assessment in year 2019 for the region of Dakar amount to XOF 10.6 billion for the property tax, and XOF 7.7 billion for the garbage tax (this corresponds to a total of USD 29.8 million).[18]

[13] Respectively *Contribution foncière* and *Taxe d'enlèvement des ordures ménagères*. Both taxes are managed using the same valuation roll and emitted on combined tax notifications.

[14] There are eight DGID offices in the region of Dakar: five in Dakar, and one for each of the other cities. There are four Treasury offices, one for each city.

[15] This is used as a tax base irrespective of whether the property is or not for rent.

[16] These dates are the theoretical ones, in practice in 2019, notifications were emitted in August, the distribution lasted until January 2020.

[17] The rate is 3 percent outside of the region of Dakar.

[18] Ignoring taxes on vacant land, which amount only to 0.3 percent of entries on the valuation roll.

2.3 Current impediments to effective property taxation

Several major challenges impede the functioning of property taxation. First, only a very small fraction of owners actually come to the tax administration to make their annual declarations. Tax culture and voluntary compliance do not seem strong enough as of today to justify this declarative system. The current valuation roll is both largely incomplete and outdated, and everything seems to hint to a situation where the city has outgrown the administration's capacity to keep up. In the absence of voluntary declarations, maintaining up-to-date information on properties and owners is a resource intensive task, for which the relevant departments are insufficiently coordinated. The cadaster (which is a department within the tax administration) has modernized its way of identifying plots, and has an identifier for any given plot, albeit without information on property owners nor property values. These efforts have been made independently of the fiscal offices, and as a result the current valuation roll doesn't include any plot identifier nor harmonized address information. Ideally, information from land and property registries could be used to inform the valuation roll and help link plot and taxpayer identifiers. However, this information has not been integrated with the fiscal information (it is managed by a third department, the department for property and land registration).[19]

The way forward to expand and update the property tax register is for the administration to conduct fieldwork and collect up to date property information, linked with harmonized address information. However, conducting comprehensive field operations of this scale is costly, and investments in this direction have not been made to date, perhaps because they were not considered cost-effective by the tax administration. As a benchmark, we estimate the cost of collecting the relevant information on any given property to maintain an up-to-date valuation roll to be around XOF 14,000 (USD 23). If this cost is applied to the total number of plots in the region of Dakar, total estimated costs would be around XOF 5 billion. While as of today, property and garbage tax collections only amount to XOF 6.3 billion.[20] Revenues would of course increase if the valuation roll were totally updated, but this shows that such activities necessitate an upfront investment. The administration does conduct small scale localized property censuses, but rarely, and with substantial loss of information throughout the process.[21] Improvements of the valuation roll are therefore extremely marginal from one year to another. Some components of these high costs could be reduced with technological improvements. For instance, our estimated cost of obtaining the required information in the field relies on data collection conducted using tablets. Until recently however, the administration carried out these tasks on paper. The staff time required to process handwritten information and upload it in the administration's system slows down the exercise and generates

[19] Furthermore, formal property registration covers only a small share of actual property owners. This is reinforced by the fact that a very large share (42.8%) of owners in the region of Dakar acquired their property through inheritance (Agence Nationale de la Statistique et de la Démographie, 2012) – in these cases, formal paperwork to designate the new legal owner is rarely carried out.

[20] As a benchmark for the cost of collecting relevant information on properties and property owners in the field, we use the total cost of the property survey conducted for this project and presented in the subsequent section.

[21] All is done on paper and part of the field work information is never entered into the valuation roll. Furthermore, each tax office has its own processes and there are no harmonized rules to collect, clean and store this information.

large inefficiences, which also serve as a disincentive to invest in this fieldwork according to discussions with tax administration officials. Furthermore, when applying valuation techniques to verify the tax base for property taxation, the administration relies on lengthy manual inspections using measuring devices to retrieve built area. The human and technical resources required could be significantly reduced if there was a stronger reliance on high resolution satellite images. Indeed the cadastral offices within the tax administration include staff with the required skills to retrieve built areas from such images.

Another source of inefficiency is the lack of cooperation between the tax administration and the Treasury, stemming from political rivalries between the two administrations in spite of the fact that both are under the authority of the Ministry of Finance. One cause if this lack of cooperation is that the responsibility of collecting payments for most tax heads has been shifted to the tax administration (for instance income tax and taxes paid by large and medium enterprises). According to discussions with officials within the tax administration, and within the Treasury, the Treasury fears that by cooperating too much on property taxes, it might lose the collection function for this tax as well and become more obsolete as an institution. The Ministry has not pushed the two administrations to work more closely together, possibly because these taxes accrue to local governments and are a lower priority than taxes accruing to the central government. As a result there are very few interactions between the tax administration and the Treasury on the topic of property taxes, although they are at either end of the fiscal chain. More concretely, Treasury agents in charge of distribution have an important local knowledge of taxpayers, and are also aware of many inconsistencies existing on the valuation roll (owners which deceased or moved out, properties which no longer exist, imprecise addresses, multiple entries that in reality correspond to the same property). They typically compile this information on notebooks used in the distribution, however these notes are never compiled and shared with the tax administration at the end of the distribution campaigns. Therefore, year after year, the Treasury keeps receiving erroneous tax notifications, and recent information from the field is not exploited to improve the valuation roll. The Treasury also complains that they do not receive a digital version of the valuation roll, that would help them manage the distribution process and compile their feedback. According to the Treasury, the meager performance of property taxation is due to the lack of effort of the tax administration to improve the valuation roll. On the other hand, when confronted with the poor collection ratio, the tax administration accuses the Treasury of not carrying out sufficient efforts in the distribution and enforcement process.

This sheds light on what appears to us as an important hurdle to improvements in property taxation. The incentives of the tax administration are totally separated from the ones of the Treasury, since they focus only on assessment *amounts*, rather than on the quality of the information in the valuation roll, and irrespective of final amounts of *payments*. More precisely, the fiscal service department is evaluated based on the number of entries and total amounts on the valuation roll, and tax notifications that cannot be distributed by the Treasury because they are irrelevant or obsolete are never removed from this count. Therefore the fiscal service department has little incentive to

invest in the costly activity of updating the valuation roll and potentially "losing" some invalid entries. A positive shift could be operated if the incentives were set at a higher level (Ministry of Finance) and took into account the quality of information in assessments as well as actual distributions and payments that occur down the road.

3 Data

3.1 Administrative data

Cadastral data

The cadaster department gave access to its cadastral GIS dataset for the region of Dakar. GIS assistants from the research team helped digitize, extend and edit the dataset. The data includes administrative boundaries of cities, communes and cadastral sections;[22] plot boundaries; plot unique identification numbers;[23] plot geolocalization; plot area; built area for a subset of plots. There are 688 cadastral sections in the region, but up-to-date and digitized GIS data exists only for around 215 as of today. Built area measurements are from two sources: drone images shared by the cadaster, and, for areas in which these were not available, high resolution satellite images obtained for the purpose of the study.[24] The GIS assistants retrieved built area from the images.

Tax assessment data (valuation roll)

This dataset is the universe of properties that are registered on the valuation roll for the property tax and the garbage tax, for years 2015 to 2019. It includes the following variables: tax notification number, name of the property owner, tax head, tax base (annual rental value of the property), tax liability for each tax head, address information.[25] For a subset of entries (30 percent), the dataset includes the tax identification number of the property owner. For another share, the dataset includes a temporary identification number specific to property taxes (84 percent). Each entry also includes the corresponding tax administration office and the Treasury office where the taxpayer is expected to pay the tax. 54,472 properties are registered on the 2019 valuation roll for the region of Dakar – 39,026 properties are registered both for property tax and garbage tax, 12,605 for garbage tax only, and 2,840 for property tax only.[26]

[22] On average, a section includes 380 plots.

[23] *Numéro d'Identification Cadastral* or *NICAD*

[24] Accessed through the *Centre National d'Études Spatiales*, France. Resolution of drone images: 30cm. Resolution of Pleiades satellite images: 50cm.

[25] Address information doesn't follow systematic rules and cannot be linked with the plot identifiers from the cadastral data.

[26] In Appendix A.6, we highlight some inconsistencies found in the valuation roll.

Tax revenue data

There are two sources of administrative data on revenues. The first corresponds to total revenues and expenditures of local governments (cities and communes) for 2018, aggregated at the country level, obtained from the Ministry of Finance. The second corresponds to total collections for the property tax and the garbage tax for 2019, at the country level.[27]

3.2 Property owner survey

A property owner survey was conducted between September and December 2018, targeting 202 cadastral sections in the four cities of the region. The sections were selected based on their tax potential (excluding informal settlements, industrial and market areas) and existence of up-to-date cadastral data. The map in Figure 1 indicates surveyed sections, and Appendix A.2 provides details on the sampling strategy and survey protocol. Within each section, plots were drawn randomly using cadastral data, and the targeted respondent was the owner of the property. If he or she was absent for a significant period, the person in charge of matters related to the property was to be surveyed. Data was collected for 2,474 respondents and associated properties. The survey is representative of 59,505 plots – if considering all eligible plots in survey areas[28] – and 32,370 plots if taking into account the non-response rate.

The survey lasted around 40 minutes and included modules about: owner identification, socio-economic information, utilization of the property, value of the property, rents and tenants, knowledge about local taxation, personal experience with property tax and the tax administration, perception of the government and of local governance, interactions with neighborhood delegates, observable characteristics of the property. Importantly, the survey was conducted by a private survey company independent from the administration.

Throughout the analysis, we apply weights to make the sample representative of all properties in surveyed sections which would be eligible for the survey. The survey aims at being representative of properties owned by individuals in a large span of areas eligible for property tax improvement. There are three important dimensions of selection – selection of cadastral sections, definition of eligible plots within these sections, non-response. We discuss all of them thoroughly and provide arguments as to why they do not hinder the analysis in Appendix A.2.2. In Appendix A.4 we describe the computation of the key variables used throughout the analysis.

3.3 Real estate expert valuations

Four real estate experts were hired to provide valuations for a subset ($N = 441$) of surveyed properties. They observed the plots from the outside, had information on the built area, and were asked to provide a lower bound and upper bound of the rental value for each property. Sample size was

[27] Revenue collection data at the city, commune or property level is not yet available.

[28] Eligibility was defined according to the following criteria: there is a construction on the plot, the owner is an individual (and not a firm nor a public administration), the construction is maximum five floors high.

set according to budgetary constraints, and the proportions of properties within each of the four cities was maintained. The sample included all properties for which a photo had been taken during the survey,[29] and was completed by over-representing properties in wealthy neighborhoods. See Appendix A.2.3 for more details on the sampling and protocol.

3.4 Data from the Treasury distribution survey

The objective of the distribution survey was to collect information on the outcome of tax notifications throughout the 2019 distribution campaign.[30] Between August and December 2019, the Treasury agents in charge of distribution were equipped with tablets and a short questionnaire, prompting them to answer questions such as whether a given tax notification was usable or non-usable (if so, why), if usable, whether the tax notification was distributed or not, and the geographic location of the property where the tax notification was distributed. There was a low compliance among Treasury offices with the survey protocol, as a result we only recovered data for a subset of tax notifications (around 20 percent of those listed on the valuation roll). Estimations based on this data are therefore to be treated with caution. In Appendix A.5, we provide further detail on the protocol and data, and show that we can rule out a strong and systematic selection on the tax liability amount determining the inclusion of specific tax notifications in the distribution survey dataset.

4 Assessing the property tax gap

4.1 Extensive margin: too few properties in the tax net

Property tax compliance on the extensive margin relies on three steps: (i) the property must be included on the valuation roll, (ii) the owner must receive a tax notification, (iii) the owner must pay the tax. We have information on (i), partially on (ii), on (iii) at the aggregate level from administrative sources. We have information on (ii) and (iii) in the survey.[31]

Measuring the tax gap using administrative and distribution data

The total number of plots in the region of Dakar is estimated to be around 374,000, based on cadastral data.[32] In contrast, the valuation roll includes entries only for 54,472 properties in 2019. Thus, only around 15 percent of plots are on the valuation roll.[33] This is a first major margin of

[29] This was implemented at random in the survey protocol.

[30] This corresponds to the fiscal year subsequent to the property owner survey.

[31] We cannot link entries from the valuation roll with the survey at the plot level.

[32] The exact number of plots as per cadastral data is 373,844.

[33] The valuation roll includes one entry for each property, meaning that if two owners own two different apartments on the same plot, this should be associated with two entries instead of one. A comprehensive valuation roll would therefore include more entries than there are plots in the city, and 15 percent is therefore an upper bound. The 54,472 figure includes 197 entries for vacant land. In theory the valuation roll should include all plots, whether with or without built property, since there exists a tax on vacant land. Finally, a reform in 2018 changed the instrument used to tax

dysfunction of the property tax.[34]

Being included on the valuation roll does not equate with receiving the tax notification. The distribution of tax notifications is a lengthy and cumbersome process which is only partly successful. First, the tax notification outcome survey revealed that a large share of tax notifications are classified as non-usable by the Treasury agents (29.9 percent), and hence not distributed. In the majority of cases (around 60 percent), being categorized as non-usable is due to the fact that the same property actually appears two times on the valuation roll, although with slight differences in the name or address.[35] A small fraction of tax notifications are classified as usable but not distributed (1.2 percent). As a result, only around 68.8 percent of the notifications are distributed.[36] This reveals that conditional on being on the valuation roll there is a substantial probability of not receiving a tax notification. Overall we estimate that the share of properties that receive a tax notification is around 10 percent (15 percent on the valuation roll multiplied by 69 percent distribution rate).

Finally, data on tax revenues reveals that collection ratio is at 27 of total assessments in 2019, for the property and the garbage tax combined. Figure 3 contrasts assessments and revenues, overall and by tax head.[37] Enforcement is slightly superior for the property tax than for the garbage tax, but remains overall strikingly low.

Measuring the tax gap using survey data

This results in a situation where the fact of receiving a tax notification for the property tax, and of paying this tax, is an exception rather than the rule. The property owner survey confirms these orders of magnitude. Figure 1 shows a map of the region where the surveyed sections are classified based on the share of respondents in the tax net in 2018. We define being in the tax net as having paid and/or having received a tax notification for 2018. Yellow sections are the ones where the share in the tax net was 0 percent (21 percent of sections). The share in tax net was 25 percent or above in only 30 percent of sections (dark red on the map).

Table 1 displays descriptive statistics from the survey shedding light on the tax gap on the extensive margin at the level of the property owner. 12.4 percent received a tax notification, and 16.8

property owned by firms, as a result firms were excluded from this valuation roll. Firms were still included in the 2017 valuation roll, in that year, the roll included 1,964 firms.

[34] As a benchmark, comparable data for the capital city of Uganda, Kampala, which is similar in population size to Dakar, reveals that out of an estimated 350,000 properties, 300,000 (86 percent) are enlisted on the administration's valuation roll, after donor-supported efforts carried out between 2014 and 2018 (Manwaring & Regan, 2019).

[35] The other main motive for being classified as non-usable is *Imprecise or unknown address*.

[36] These figures are to be treated with caution since they are estimated on the subset of tax notifications that were included in the distribution survey. This percentage relies on the strong assumption that the rate of distribution is the same for tax notifications included in the distribution survey, and those excluded for which we do not know the outcome. It is likely that non-distribution is under-declared by Treasury agents, leading to a higher rate of non-distribution in the subset of notifications excluded from the distribution survey. This would mean that the 68.8 percent is an upper bound of the distribution rate.

[37] Figure 3 shows collection ratios at the national level. We do not have the breakdown of revenues by region. The national collection ratio is 27.4 percent. If we assume that all revenues are collected in the region of Dakar, the collection ratio in the region would be 31 percent. If we assume that the percentage of collected revenue that is in Dakar is the same as the percentage of assessed liabilities that are in Dakar (87 percent), the estimated collection ratio in the region of Dakar is 27.5 percent.

paid the property tax in the year of the survey.[38] When combining both, we find that 19.7 percent of property owners are in the tax net.[39] Figure A.13 in Appendix shows the reasons cited by respondents when asked why they did not pay the tax. 86 percent report that it is because they did not receive a tax notification, 11 percent that they did not know they were subject to this tax (multiple responses were possible). Having no money to pay, or refusing to pay, were rarely selected.
All the variables used in in this section are defined in Appendix A.4.

4.2 Intensive margin and total foregone revenue

The tax gap on the intensive margin is the difference between theoretical tax liabilities based on market property values, and actual tax payments. These differences may have several sources. First, due to the declarative system, the normal assessment process is one in which owners self-declare the value of their property at the administration's office. The administration lacks the means to verify declared values at a large scale, creating the potential for under-valuation by owners. Second, when the administration does carry out property censuses, field agents may have insufficient information to correctly assess property values, and are also subject to declarations made by owners or other inhabitants met during the fieldwork. Third, although owners have the legal obligation to make declarations annually, this is extremely rare, therefore the vast majority of entries on the valuation roll are maintained across time, in spite of rapidly increasing market real estate values. Third, discrepancies on the intensive margin could also be due to negotiations and informal arrangements between taxpayers and tax officials, although survey data suggests this is not widespread.[40] Fourth, taxpayers may not pay their liability in full.

Comparing values on the valuation roll with property values declared in the survey

It is not possible to restrict the assessment data to the geographical sample corresponding to the survey, therefore the comparison of values is only done at the aggregate level, and is to be treated as suggestive since they do not correspond to the same pool of neighborhoods. The distributions of property values in the assessment data versus in the survey are plotted in Figure A.3. The right

[38] How can the share paying be larger than the share receiving a tax notification? First, to be conservative, we code as having paid observations where the respondent declares paying *"less than a year ago"*, in addition to *"for this year"*. Recollection problems might induce some discrepancies regarding dates, for example some respondents being counted although they had in fact paid for 2017. Furthermore, based on qualitative work in the field and with the Treasury offices, we know that some taxpayers who are already on the valuation roll might pay the tax without waiting for the tax notification – either because they do not expect the delivery to be successful but know that their name is on the valuation roll, or because the Treasury office contacts them asking them to come and pay, without bothering to distribute the notification. 45 percent of the $payers$ indicate that they did not receive the tax notification in 2018.

[39] The share of properties in the tax net as measured in the survey is higher than in the previous paragraph using administrative data. This can be explained by the fact that the survey only covered a set of cadastral sections within the region, which had been identified as having fiscal potential.

[40] The perception of corruption is higher than the level of personal experience of informal arrangements with the tax administration. Among respondents who ever paid the property tax, 1.3 percent report that they ever made an informal payment to a tax agent, against 0.6 percent among those who never paid. Overall, 16 percent of respondents believe that tax assessments in general may depend on informal arrangements between owners and tax agents. Results from a list experiment were inconclusive to shed light on the extent of corruption in this setting.

panel is restricted to entries on the valuation roll that are registered for the property tax, and survey observations where the respondent declares paying the tax. We observe that the distribution of market values as measured in the survey is significantly shifted to the right compared to the values on the valuation roll. While in the survey, 15.5 percent of properties are *below* the abatement threshold, this is the case of 41 percent of the properties listed on the valuation roll.This diagnostic is very much shared by the tax administration, which often refers to values on the valuation roll as *"obsolete"*.

Comparing theoretical and actual tax payments based on survey data

We can compute theoretical tax liability for each property in the survey sample, using property values, and the owner-occupied status of the property:

$$Theoretical\ tax_i = 0.05 * (AnnualPropertyValue_i - 1,500,000 * H_i)$$

were $AnnualPropertyValue_i$ is the property value variable measured in the survey and H_i is a dummy equal to one if the property is the main residence of the owner. Property values are declared, and when missing, imputed using predictions based on a hedonic regression model.[41] This amount can be compared with the tax amount paid, as declared by the respondent. Results are shown in Table 1. The average amount paid by payers is XOF 16,518 while the average theoretical tax for the same properties is XOF 271,116 (USD 28 against USD 461). The actual tax paid is lower than the theoretical one in 61 percent of cases.[42]

Total theoretical tax revenues amount to XOF 10.1 billion (USD 17 million, weighted survey sample). Yet, as per survey responses, total tax payments amount to 9.4 percent of this potential.[43] In Figure 4, we provide a breakdown of the property tax gap, into extensive margin – foregone revenue due to the exclusion of some taxpayers (74.5 percent of total potential revenue or 82 percent of foregone revenue) and intensive margin – due to tax amounts that are lower that theoretical liabilities for taxpayers who do pay (16.2 percent of total potential revenue, or 18 percent of foregone revenue).

4.3 Empirical insights on the factors leading to a large tax gap

A candidate determinant of property tax performance often mentioned in the literature is the existence of clearly defined property rights, and of formal real estate and rental markets (Awasthi *et al.* , 2020). Table A.3 displays some descriptive statistics from the survey on property ownership and the rental market. This novel data suggests that the real estate market is indeed plagued by informality and a lack of interactions with the administration, but that significant segments of it

[41] See Appendix A.4 for details.

[42] The average theoretical tax liability for the whole sample is XOF 177,106. Figure A.4 shows the distribution of the difference $Theoretical\ tax - Actual\ tax$ conditional on $Actual\ tax > 0$.

[43] The result is similar if we consider unweighted survey data, potential revenues are at XOF 400 million for the 2,474 properties, and paid tax amounts at XOF 39.5 million.

correspond to an active market with transactions between previously unconnected individuals.[44] Only 35 percent of owners have a formal property title,[45] but another 53 percent have a weak property title (such as a temporary title, an occupation permit). 20 percent of owners inherited the property, a situation in which the paperwork to transmit legal ownership is rarely carried out. This suggests however that overall, more than 80 percent of owners declare having an ownership document of some type, and 66 percent purchased the property – indicating that the administration could potentially leverage these mechanisms to collect information on owners and rents. Additionally, the rental market appears to be vibrant, and functioning as an anonymous market: 30 percent of properties are at least partly for rent, and in these cases, the tenant and the owner know each other (family or friends) in only 6.6 percent of cases. In 50 percent of cases the rent was set as recently as in the past three years.

Another factor that plays a substantial role in the success of property taxation is the existence and credibility of enforcement actions (Okunogbe, 2020). The survey data presented in Table A.2 depicts a situation in which enforcement actions are extremely rare - among owners that are known by the tax administration, defined as having ever received any type of tax notification, only 5 percent have ever been subject to an enforcement action. In terms of $perceived$ enforcement, it seems that respondents are mostly unaware of the situation of property tax enforcement in general in their city (77 percent declare not knowing whether most people pay this tax or not); but also over-estimate the likelihood of enforcement actions targeted at non-compliers (49 percent consider that such actions are either very likely or likely). This is consistent with a situation in which the majority of owners are not aware of their property tax obligation, and enforcement actions are quasi non-existent.

5 Implications of weak enforcement for the progressivity profile of the tax

The previous section has shown that tax enforcement is very low. In this section, we examine whether this weak enforcement (i) makes the tax more progressive than what is provided for in the legislation (ii) maintains the same progressivity profile as what is provided for in the legislation, or (iii) increases the regressivity of the tax.

The tax rate is proportional to the value of each property, and the abatement threshold for owner-occupied properties is relatively low with respect to current market values in the region of Dakar. As such, one could argue that on paper, the taxation profile is rather regressive, in line with arguments developed by Piketty (2020) or Saez & Zucman (2019) for property taxes in general.[46] However, since the administration is resource-constrained, it could be focusing its efforts on properties of higher value to maximize revenue collections. In other contexts, Khan *et al.* (2016), Weigel

[44] As opposed to a situation where all transactions occur within the family or social network

[45] *Titre foncier*.

[46] In spite of being set at a relatively low level, the existence of the abatement does make the theoretical taxation profile more progressive than in France for example.

(2018) and Okunogbe (2020) observed that under limited administrative capacity, wealthier owners indeed faced a higher fiscal pressure either due to voluntary compliance or targeted enforcement efforts. Furthermore, discussions with tax administration and Treasury agents indicate that in some stages of the taxation process, their $stated\ intention$ is to focus on higher potential taxpayers.

5.1 Loose connection between average property value and tax enforcement in a given neighborhood

We first conduct an analysis at the section level. The maps in Figures 1 and 2 show the share of properties in the tax net and the average property value by cadastral section.[47] If tax enforcement was significantly stronger in high real estate value neighborhoods, the two maps would overlay quite well. It is visually possible to see that this is not the case.

We test this correlation at the section level in a regression where the dependent variable is a dummy equal to one when there is at least one property paying the tax in a given section (Table 2, column 1), and the share of properties paying the tax (columns 2 and 3). We find no significant correlation between average property value and tax enforcement on the extensive margin (column 1). On the intensive margin, when average property value increases by 1 percent, the share of properties in the tax net increases by 0.12 percentage points or 0.5 percent (column 2). A doubling of average property value in the area would increase the share in tax net by 12 percentage points (50 percent). Column 3 shows the correlation on the intensive margin conditional on at least one property being in the tax net, the results have a similar order of magnitude.[48] This shows that there is slight a correlation between tax compliance and average property value at the section level, but its magnitude is limited, and more importantly, the correlation of average property value with the likelihood of at least one taxpaying property owner in a given section is not significant.

5.2 Fiscal pressure by decile of property value

We rank properties based on (deciles of) property value to assess the progressivity of the tax – by looking at whether observed fiscal pressure varies across deciles and how it differs from theoretical fiscal pressure.

In Figure 5, we plot the share of owners theoretically subject to the tax (red line), and the share that actually paid the tax (blue line) for each decile of property value.[49] The theoretical share of payers is lower than 100 percent in deciles one and two, because of the abatement threshold on owner-occupied properties. The share paying increases across deciles - from 7 percent in the first decile, to 35 percent in the tenth decile.

[47] In Figure 2 sections are divided into quartiles of the median property value.

[48] The covariates other than property value will be studied in Section 6, but the results on our predictor of interest are unchanged when they are excluded.

[49] This section focuses on the property tax per se and excludes the garbage tax.

Figure 6 allows a more comprehensive analysis by including the intensive margin. The grey horizontal bars (right-hand side y-axis) show the share paying by decile. The grey dashed line shows the theoretical tax rate that should be observed in each decile, computed as the ratio of total theoretical tax amount over total property value in any given decile. It increases, from 1.9 percent in the first decile to 4.7 percent in the tenth decile. This increase is due to the fact that the abatement corresponds to a larger share of total property value in lower deciles. This grey dashed line is to be compared with the green dashed line that represents the $observed$ effective tax rate by decile. The effective rate is almost flat across deciles, and never exceeds 0.5 percent. It very slightly increases with property value, averaging at 0.4 percent in deciles eight to ten, against 0.2 in deciles one to three. Because the increase is minimal compared to the increase in the theoretical rate, this means that owners in top deciles are significantly further away from compliance: in the first decile, the observed rate is 1.6 percentage points lower than the theoretical one (0.3 percent against 1.9 percent); while in the tenth decile, the observed rate is 4.2 percentage points lower than the theoretical rate (0.47 percent against 4.7 percent).

Still in Figure 6, we restrict to $payers$, and compare the theoretical effective tax rate by decile (orange line) with the observed effective tax rate (red line). The observed rates range from 1.6 percent to 5.7 percent, similar to the range of the theoretical rates. However, what is striking is the X-pattern resulting from the red and orange lines: for deciles one to six, the observed rate is $higher$ than the theoretical rate, while for deciles seven to eight, the opposite is true. The observed effective tax rate is particularly high in the first decile, increases modestly from deciles two to six, and decreases from deciles seven to ten. This suggests a strongly regressive profile, conditional on payment, since in lower deciles taxpayers are paying more than they should, while in higher deciles taxpayers are paying less than they should. Taxpayers in the first decile faced a 5.7 percent tax rate, while those in the tenth decile faced a 1.7 percent tax rate.

In Figure 7, we show how this situation translates in terms of revenue implications. The grey dashed line (right-hand Y-axis) shows amounts of foregone revenue, computed for each decile as $theoretical\ tax\ liability - tax\ amount\ paid$. These strongly increase with property value. They amount to XOF 3.8 billion for the tenth decile, and XOF 95 million for the ninth decile. Foregone tax revenues are thus both very large, and highly concentrated in the highest deciles.

The blue line shows the share of tax revenue that is foregone, for each decile, computed as $theoretical\ tax\ liability - tax\ amount\ paid$ divided by $theoretical\ tax\ liability$. The orange line shows the same ratio, but restricted to payers. A value of zero suggests that taxpayers are paying exactly the theoretical amount, a value of fifty that they are paying half of the theoretical amount, and negative values mean that they are paying more than the theoretical amount. Variations in the blue line are hard to observe visually, but reveal that the share of lost revenue varies between 84 and 96 percent – it averages 89 percent in deciles one to three, and 91 percent in deciles eight to ten. The orange line shows that payers in decile one are paying $+284$ percent (almost four times) the theoretical liability, payers in deciles two are paying around $+80$ percent, payers in deciles three to six are paying a liability close to the theoretical liability, and payers in deciles seven to ten are

paying 20 to 65 percent less than the theoretical liability, the figure is increasing with deciles.
We perform robustness checks by using alternative definitions of the property value variable – using in turn, exclusively declared values, exclusively predicted property values based on a hedonic regression model, and finally assigning random residuals to predicted values – as well as by excluding owners of multiple properties. These tests are described in Appendix A.7. We find that the main conclusions hold: the effective taxation profile is more regressive than what is provided for in the legal framework.

5.3 No targeting on property values in the distribution process

The set of owners paying the tax results from the combination of (i) targeting by the tax administration (registration and assessment) and (ii) targeting by the Treasury (distribution and revenue collection). The tax notification distribution survey provides suggestive evidence on step (ii), the correlation between tax liability and distribution of a tax notification, conditional on being on the valuation roll. Through qualitative work, the directors of the different Treasury offices stated that their strategy was to focus on tax notifications with the highest liabilities. It was also clear that enforcement and prosecution actions, albeit rare, were focused on taxpayers with higher liabilities. In spite of this, we do not find a correlation between tax liability and probability of being distributed. Figure A.12 plots tax notification outcome by decile of tax liability, and Table A.24 show results of a simple regression of a dummy indicating whether the tax notification is distributed on tax liability. We observe that the share of non-usable notifications is slightly larger in deciles seven to ten than in deciles one to six, and in the regression, the coefficient on tax liability amount is negative and significant (column 3), negative and non-significant when we control by Treasury office (column 4).[50] This suggests that there is no strategic targeting of higher valued properties in the distribution process.

5.4 Benefit view of property taxation?

Under the *benefit view* of property tax, this tax is equivalent to a fee paid by owners (or occupants) for local public services (Oates & Fischel, 2016). We find that the current state of property taxation in Dakar is very far away from such a situation, by contrasting fiscal pressure with a standardized local public service index in Figure 8. The grey bars display the overall share of tax liability being paid in each decile of property value, and the red line shows the effective tax rate conditional on paying. None are increasing with decile of property value (the effective tax rate is slightly decreasing). On the contrary, the standardized local public service index monotonically increases across deciles of property value. The index combines information on street lights, the existence of a sewage system, the presence of piped water, the existence of a garbage removal service – see Appendix A.4 for details on its computation. This sheds light on a situation where property owners who benefit from better local public services are far from being the ones paying the largest

[50] The distribution survey data should be treated with caution, see Appendix A.5 and footnote 36.

share of their tax liabilities.

5.5 Simulations under revenue-maximizing targeting

We compare the observed fiscal situation – the set of properties for which the tax is paid, *observed targeting* – with two alternative targeting strategies the tax administration could perform, while keeping the number of payers constant within each of the four cities. The alternative targeting scenarios are respectively *revenue maximizing targeting*, and *random targeting*, under the same resources constraints, i.e. without increasing the number of payers.

Results are shown in Figure 9. In the left panel, the blue line represents the share of properties for which the tax is paid in each decile of property value, in the observed targeting. The red line represents revenue-maximizing targeting: under this scenario it would make most sense for the tax administration to focus solely on properties in decile seven, eight, nine and ten. The green line represents random targeting, as such, the share of properties paying is more or less constant across deciles. Although observed targeting is slightly increasing in property value, it appears much closer to random targeting than the revenue-maximizing targeting. This is confirmed when we consider expected tax revenue under the different scenarios – in the right panel of Figure 9. Tax revenue under observed targeting is almost equal though very slightly inferior to what is found with the random targeting;[51] this amount is almost doubled if property values were updated keeping observed targeting; and this amount is more than quadrupled if targeting was done in a revenue-maximizing way.[52]

We perform a robustness check in which we conduct the same analysis, but accounting for a cost of distribution of each tax notification, based on the distance between the property and the Treasury office. The results are very similar, they are described in Appendix A.7.

Overall, this section has shown that although the share of property owners who pay the tax increases with property value, the effective tax rate is almost flat across the distribution of property values suggesting that there isn't a stronger fiscal pressure on higher valued properties. In fact, the gap between the effective rate and the theoretical rate is significantly wider for top deciles of property value. Furthermore, the effective tax rate faced by $payers$ is lower in the top deciles of property value. This results in a situation where the tax is more regressive in its current implementation than on paper. The administration would collect four times as much revenue with the same number of payers if these were concentrated in the upper deciles of the property value distribution.

[51] The bar labeled *Observed payments* show total tax amounts as declared in the survey, with missing tax amounts for respondents who declare paying the tax set to zero. *Observed targeting* show total tax amounts for all respondents who declare paying, where tax amount is replaced by the theoretical amount when it is missing.

[52] We are not making any assumptions on compliance here since we assume that properties in the tax net would pay their tax liability in full.

6 Economic and geographic determinants of being in the tax net

6.1 Income and wealth

In this section we try to take into account property owners' capacity to pay in a wider sense, rather than restricting the progressivity analysis to a ranking based on property value (Norregaard, 2013; de Carvalho Jr., 2015; Balan *et al.* , 2020). We restrict to cases where the respondent is either the owner or a family member of the owner (95.5 percent of total cases) for the income variables to be relevant.

Graphical analysis

We replicate the progressivity analysis of fiscal pressure by decile, but ranking properties by owners' household income (respectively, by an overall wealth and income index ranging from 1 to 10) instead of by property value.[53] Results are displayed in Figures A.7 and A.8.

The share of owners paying increases with income, from 10 percent in the first group against 25 percent in the fourth group (grey bars in Figure A.7). The observed tax rate is 0.2 in income groups one and two, 0.3 in the third income group, and 0.6 percent in the fourth group. The effective tax rate, conditional on paying, decreases across groups one to three, but it increases again and is the highest in the fourth income group.

Figure A.8 displays more irregularities when ranking is based on both income and wealth. The share paying increases with the income and wealth index. Conditional on paying, the effective tax rate is substantially lower than the theoretical rate for values of the income index between 5.5 and 9.

Overall, it seems that the finding that the taxation profile is more regressive than it should be based on the legal framework is preserved when ranking by income or by income and wealth, but the regressivity is more mild than when focusing on property values only. The top income (and income and wealth) groups seem to be facing higher fiscal pressure than the upper middle groups.

Regression analysis

We analyze how the probability of being in the tax net, and tax amounts paid, correlate with income and wealth characteristics by estimating the following regression:

$$Tax_{is} = \beta_0 + \beta_1 Ln(PropertyVal) + \beta_2 Multiple_{is} + \beta_3 IncomeGroup_{is} + \beta_4 Rented_{is}$$
$$+\beta_5 N\_RegularIncome_{is} + \beta_6 EmpStatus + S_s + \epsilon_{is}$$

where Tax_{is} is a tax outcome of owner i in section s. $Ln(PropertyVal)$ is the logarithm of annual

[53] Appendix section A.4 provides details on how these income and wealth variables are computed.

rental value of the property, $Multiple$ is a dummy equal to one if the owner possesses another property in the region. $IncomeGroup$ indicates the category of household income, $Rented$ is a dummy equal to one if at least part of the property is being rented out, $N\_RegularIncome$ indicates the number of household members who earn a regular income (including the owner), $EmpStatus$ indicates the employment status of the respondent (formal employment, informal employment, no employment or retired). Variables $Ln(PropertyVal)$ and $Multiple$ are proxies for wealth, while variables $IncomeGroup$, $Rented$, $N\_RegularIncome$ and $EmpStatus$ are rather measures of income and liquidity. S_s is a section fixed effect. We run the analysis both with and without section fixed effects.

Results are displayed in Table 3. In column 1, the dependent variable is the $TaxNet$ dummy. A doubling of property value is associated with a 5 percentage point (25 percent) increase in the probability of being in the tax net. Possessing another property is strongly associated with the probability of being in the tax net (larger by 13 percentage points or 65 percent). Being non-employed rather than in a formal employment is significantly and negatively associated with being in the tax net. There is a negative association between the number of people in the household with regular income and the probability of being in the tax net. In column 2, we control for section fixed effects. We find the same order of magnitude for the coefficient on property value. Within section, the significance of the multiple owner dummy disappears, but the one of non-employed remains. The signs on being retired and informal employment are also negative although not significant. Furthermore, the coefficients on income group become significant within section – being in income group three (respectively, four) is associated with a 6.6 (resp., 12.7) percentage points higher probability of being in the tax net compared to income group one. *Within area*, it seems that owners with higher capacity to pay are more likely to be in the tax net. These results show that although the property tax is a wealth tax, property value (the tax base) is only a moderate predictor of being in the tax net, but variables capturing income and capacity to pay play a role.

In columns 3 and 4, the dependent variable is the inverse hyperbolic sine transformation of tax amount paid, $asinh(tax\ amount)$. A one percent increase in property value increases tax payment by 0.24 percent (0.20 percent within section). Therefore the correlation between tax payment and the tax base exists, but is much weaker than what it would be under perfect enforcement (the coefficient should be close to one). There are also positive and significant coefficients on income group, and the rented dummy. When introducing section fixed effects (column 4), the income group variable remains significant (32.5 percent higher tax amount for income group four compared to income group one), so does employment status (23 percent lower tax amount for non-employed), and the negative coefficient on retired also becomes significant (24.5 percent lower tax amount). This reveals that even when controlling for property value – the tax base – actual tax payments vary with proxies of income, of capacity to pay, and employment status. This holds both within and across sections.

In columns 5 and 6, we restrict the sample to *payers*. A striking finding is that the coefficient on

property value is no longer significant.[54]

To assess how much of the variation in tax compliance income and wealth explain, we consider the predictive power of the different models. For the models with the binary outcome ($TaxNet$), we use the Youden index, computed as:

$$Index = Sensitivity + Specificity - 1$$

Sensitivity and Specificity are computed as:

$$Sensitivity = \frac{Truly\ predicted\ positives}{All\ positives}$$

and

$$Specificity = \frac{Truly\ predicted\ negatives}{All\ negatives}$$

The index takes values between 0 and 1, 1 being a perfect predictive model. We use two cut-offs for the predicted outcome, the standard 0.5 cut-off and a 0.3 cut-off, which performs better on this sample. Sensitivity, Specificity and the value of the Youden index are displayed in the bottom panel of Table 3. Without section fixed effects, we obtain a relatively low Index of 0.21. It increases to 0.52 when section fixed effects are included (columns 1 and 2). Similarly, the R-squared for the model with the continuous dependent variable increases from 0.09 to 0.27 when we add section fixed effects (columns 3 and 4). Finally, property value and other economic variables only explain 0.29 of variation in tax amount paid conditional on paying (columns 5 and 6).

We run some robustness checks, described in Appendix section A.7.3, by using alternative definitions for the property value variable, as well as by running a logistic regression for the binary outcome $TaxNet$. Results remain qualitatively similar.

Overall, we find that income and wealth variables correlate with tax payment, even when controlling for the tax base, although under perfect enforcement property value should be a perfect predictor of tax payment. But wealth and income variables alone (as we can measure them) only capture a very small share of the variation in tax outcomes. Specifically, they fail to capture a large share of the geographical variation across sections.

6.2 Urban geography

The previous results show that there are strong geographic determinants that are not captured in economic variables. We explore this further by introducing geographic variables that attempt to capture what could be in the "black box" of the section fixed effects. These variables are *Presence of street name*, as a proxy for the quality and visibility of addressing, *Distance to closest Treasury*, and *Distance to closest Tax office*.[55]

[54] As a sanity check, in Appendix Figure A.5, we verify that there is a positive correlation between property value and tax amount paid.

[55] See Appendix Section A.4 for details on computation of these variables.

In Table 2, we first estimate a section level regression with these geographic variables (where we also control for average property value). The variable for street name is the share of properties in the section with a street name, and the distance variables are averaged by section. We find that each kilometer further away from the closest Treasury office significantly decreases the probability of having one property in the tax net by 4.6 percentage points or 5.8 percent. It is notable that this variable is significant while property value is not. The share with street name and distance to tax office are not significant, and none of the geographic variables are significant on the intensive margin (columns 2 and 3).

In Table A.6, we estimate individual level regressions with the geographic variables. In column 1 the only covariates are section fixed effects. The Youden index is $higher$ than in column 1 of Table 3 (0.43 against 0.21): using sections alone has a higher predictive power than introducing all economic variables, and only slightly lower predictive power than when both are combined. However, street name, distance to the Treasury and distance to a Tax office do not come out as significant at the individual level, whether we include them with or without section fixed effects (columns 2 and 3). When geographic variables are included alone without section fixed effects the predictive power of the model is very low (columns 2 and 5).

Overall, these geographical variables fail to fully capture the influence of section fixed effects. In Appendix section A.7.4, we further contrast economic and geographic determinants, for two different outcomes: ever receiving a property tax notification versus ever receiving an income tax notification. The results confirm the strong role played by geography for *property* taxation in particular.

6.3 Socioeconomic characteristics

Finally, we examine the predictive power of a wide array of socioeconomic variables in explaining the fact of being in the tax net using a classification tree (CART, borrowed from machine learning techniques.[56] As for the previous regressions, this analysis is purely descriptive and no causal interpretation is possible. The full set of included covariates is presented in descriptive statistics in Table A.4, and includes variables on visible property characteristics, personal ownership characteristics, individual sociodemographic characteristics, and tax related variables. Results are shown in terms of variable importance in Figure A.14.[57] The left panel displays results from the CART when section fixed effects are excluded, and the right panel shows results with section fixed effects. In the former, the distance to Treasury office comes out as the most important variable, followed by property value, and with lower variable importance: age, employment status, household income, and finally some visible property characteristics (presence of tiles, number of floors). In the

[56] Contrary to linear or logistic regressions, no assumption is made on the functional form of the error term. Variables are used to split the sample in a way that minimizes node impurity, i.e., the extent to which the resulting subset is far away from including only one type of outcome value ($TaxNet = 1$ or $TaxNet = 0$). Furthermore, the tree is pruned – meaning that we impose a limit on the number of splits and leaves, to facilitate interpretation and generalization – using the complexity parameter $cp = 0.001$.

[57] Variable importance of variable var is the mean decrease in node impurity from splits based on var.

model with section fixed effects, the section variable largely dominates all others. It is followed by distance to treasury, trust in tax administration, and number of floors, while property value is only fifth.

7 Historical perspective: early but short-lived property tax capacity

The previous section has shown that it is difficult to pin down factors explaining the strong variation in tax enforcement across neighborhoods. Taking a step back, another approach to interrogating the factors of property tax enforcement is to examine whether such capacity ever existed in Dakar, if so whether disruptions occurred at the time of independence, or if this tax was introduced relatively recently and never suited to the context. We show that property tax capacity was relatively high during the early colonial period, and started to fade out before independence.

The property tax was introduced very early in colonial Senegal – as soon as the 1850s – and its modalities have barely evolved since its first existence.[58] It resembles the property tax created in the metropole (France) in the aftermath of the French Revolution.[59] To put the current performance of the property tax into perspective, we investigate the evolution of property tax capacity in the region of Dakar over the long term, throughout colonial and postcolonial times. We find that property tax performance was important and increased strongly in the early colonial period, before declining in the 1940s, and is nowadays substantially lower than in the first half of the twentieth century, as shown in Table 4. As a measure of property tax capacity, we compute the ratio of property tax revenues over total population – property tax revenue per capita – the ratio of property taxes to total local taxes, and to total national taxes.[60] The ratio increased from XOF 6,530 in 1897 to XOF 10,006 in 1931 (USD 18). As such, this means that in Dakar, the administration was collecting from property taxes the same amount as what all tax heads were yielding, per capita, on average across the whole of French West Africa (XOF 10,700 in 1931, Cogneau *et al.* (2020a)). But this performance started fading out in the late colonial period: down to XOF 2,975 in 1942 – while overall in French West Africa tax capacity continued to increase. Progress since independence has been extremely mild, since in 2018, property taxes per capita amount to XOF 1,650 (USD 3). Strikingly, this is six times lower than in year 1931.[61] In 1931, property taxes amounted to around a third of local taxes, while this figure is at 10 percent in 2018. The decline is even sharper when considering the ratio of property taxes to total taxes, that declined from 1.9 to 0.3 percent between these two dates.

[58] See Appendix A.8 for details on the history of property taxation in Senegal.

[59] In France, property tax management has shifted to a more localized system, while in Senegal it has remained at the national level.

[60] See Appendix A.8 for details on the historical data sources, and methodological choices. Results are in converted to 2019 XOF using conversion and deflator from previous work on taxation in former French colonies in West Africa (Cogneau *et al.* , 2020a).

[61] A more detailed evolution of property tax revenue per capita is plotted in Figure A.19: the continuous black line shows the ratio of property tax revenue per capita, and the dashed line shows total tax revenues assimilated to property taxes – i.e. property tax per se plus taxes due by the same taxpayers on the same tax base.

As candidate historical factors leading to the fading out of property tax performance after the 1930s, we identify the following: (i) particularly massive city growth starting in the early 1900s, which potentially rapidly outgrew the administration's capacity to keep cadastral and property ownership information up-to-date (see Becker & Martin (1981); Bouche (1978); Seck (1961), and Figure A.20);[62] (ii) since independence, cleavages between the cadaster and the tax administration, two essential bodies of which the cooperation is key for an effective property taxation;[63] (iii) strong cleavages between the tax administration and the Treasury; (iv) complicated property and land ownership structures; (iv) fuzziness in the post-colonial legislation regarding the share of local revenues supposed to accrue to each layer of government, notably cities versus communes (which are districts within cities), potentially leading to reduced incentives for each layer of government to invest in this taxation capacity (Ndiaye, 2020).

8 Conclusion

This paper provides a thorough analysis of property taxation in an African capital city, Dakar. The tax potential is tremendous in the context of a largely developed real estate sector, and the tax gap is significant: as an upper bound, we estimate that 15 percent of plots in the region are registered on the valuation roll, and that the tax is paid for 10 percent of properties. The survey sample leads us to estimate that 9 percent of tax potential in being paid, in these areas considered by the administration as having tax potential.

The property tax was first implemented in Dakar in the 1890s, and by the 1930s, property tax revenue per capita was at the same level as average revenue per capita from *all* taxes in French West Africa. This capacity started declining in the late colonial period, and in 2019, tax revenue per capita is six time lower than in 1931. As such, it is not the case that this tax instrument does not work well because it was recently introduced, nor is it the case that it ceased to yield important revenues only after independence. The most reasonable scenario is one in which the size of the city outgrew the administration's capacity to maintain an up-to-date cadaster and valuation roll, and in which the divide between actors involved in the fiscal chain as well as their different incentives did not lead to efforts in the right direction.

We show that under limited resources, the two administrations involved in the fiscal chain are *not* concentrating efforts on properties of higher value. Thus, the property tax is far from being more progressive in its implementation than 'on paper'. In fact, the gap between the effective rate and the theoretical rate significantly increases with property value: in the first decile, the observed rate is 1.6 percentage points lower than the theoretical one, against 4.2 percentage points in the tenth

[62] Dakar became the capital of French West Africa in 1902, the same year the construction of its port began. It was a major economic and political hub throughout the colonial period.

[63] More precisely, there has been strong discontinuities in the institutional design, with the cadaster moving in and out of the tax administration over time. Until 1963, the two are separated. Then in 1963, the cadaster and land registry offices are nested within the tax administration. However, in 1979, these two departments are separated once again from the tax administration and nested within the Ministry of Urbanization and Habitat. Only to be moved back to the tax administration in 1980 (Direction Générale des Impôts et Domaines du Sénégal, 2013).

decile. The administration would collect four times as much revenue with the same number of payers if these were concentrated in the upper deciles of the property value distribution. When analyzing correlates of property tax enforcement, we find that there is a very strong local geographic determinant (heterogeneity by cadastral section) that is not subsumed by property value, income and wealth characteristics of the neighborhood; nor by distance to tax offices and quality of addressing. There is room for further investigation of the drivers of this geographic heterogeneity in property tax enforcement. Finally, although the property tax is a wealth tax, property value (the tax base) is only a moderate predictor of tax outcomes, while variables capturing income and capacity to pay play a role, suggesting that the existing tax is closer to a hybrid tax than a pure wealth tax.

This paper does not allow to precisely pinpoint the reasons why the observed tax profile is more regressive than what is provided for in the legal framework. The analysis on the economic and geographic determinants of being in the tax net suggests that there is no systematic targeting based on income, employment status, ethic group, type of property, that could fully explain the observed patterns. A hypothesis is that the owners that are registered and the associated property values are the result of a very slow and long-run $sedimentary$ process over decades – because updates of the valuation roll are so rare and incremental. Therefore, possibly a few decades ago some areas have been been subject to a better quality registration process by the administration, and thus are on the valuation roll with relatively high tax liabilities, while in the meantime, their relative attractiveness and real estate values have declined. On the other hand, some modern wealthy neighborhoods have developed more recently, with high market values, while the registration and assessment process has not been carried out, leading to under-payments for these property owners. The qualitative and data work does not lead us to suggest that the regressivity observed is the result of an intention by the administration to put relatively more efforts in lower segments of the property value distribution.

Property taxes display positive features and a high untapped potential, which make them an interesting tax instrument to invest in for administrations in the developing world. In rapidly growing cities like Dakar, where real estate wealth is increasing at the same time as local governments face larger responsibilities and challenges, overcoming the hurdles to effective property taxation seems like a *no-regret* policy choice. This paper shows that with the data the administration already possesses, and additional data that can easily be collected in the field about visible tax bases, the scope for improvement is huge. However, updating and modernizing the property valuation roll is costly and net benefits may only materialize after the very first years.

References

AGENCE NATIONALE DE LA STATISTIQUE ET DE LA DÉMOGRAPHIE. 2012. Etude monographique sur les services immobiliers du logement a dakar.

AGENCE NATIONALE DE LA STATISTIQUE ET DE LA DÉMOGRAPHIE. 2020. Condition d'habitat des mémages: éléments d'éclairage en période de pandémie liée au COVID-19. Repère Statistique de l'Environnement d'Évolution du COVID-19 au Sénégal - Note de Synthèse Numéro 5.

ALBERS, THILO, JERVEN, MORTEN, & SUESSE, MARVIN. 2020. The Fiscal State in Africa: Evidence from a century of growth. African Economic History Working Paper n55.

ALI, DANIEL, DEININGER, KLAUS, & WILD, MICHAEL. 2018. Using Satellite Imagery to Revolutionize Creation of Tax Maps and Local Revenue Collection. **8437**. Wold Bank Policy Research Working Paper.

ALI, MERIMA, FJELDSTAD, ODD-HELGE, & GOODFELLOW, TOM. 2017. Taxing the urban boom: property taxation in Africa. CMI Insight.

AVENANCIO-LEON, CARLOS FERNANDO, & HOWARD, TROUP. 2019. The assessment gap: racial inequalities in property taxation. *Working paper*.

AWASTHI, RAJUL, LE, TUAN MINH, & YOU, CHENLI. 2020. Determinants of Property Tax Revenue: Lessons from Empirical Analysis. World Bank Policy Research Working Paper 9399.

BACHAS, PIERRE, BROCKMEYER, ANNE, & JENSEN, ANDERS. 2020. Informality, Consumption Taxes and Redistribution. *Working Paper*.

BAHL, ROY, MARTINEZ-VAZQUEZ, JORGE, & YOUNGMAN, JOAN. 2008. *Making the Property Tax Work Experiences in Developing and Transitional Countries*. Lincoln Institute of Land Policy.

BALAN, PABLO, BERGERON, AUGUSTIN, TOUREK, GARBIEL, & WEIGEL, JONATHAN. 2020. Local Elites as State Capacity: How City Chiefs use Local Information to increase Tax Compliance in the DR Congo. mimeo.

BECKER, C., & MARTIN, V. 1981. Le Seéneégal centre-ouest et son évolution deémographique. *Annales de démographie historique*, 367–386.

BOUCHE, DENISE. 1978. Dakar pendant la deuxième Guerre mondiale. Problèmes de surpeuplement. *Revue francçaise d'histoire d'outre-mer*, 423–438.

BROCKMEYER, ANNE, ESTEFAN, ALEJANDRO, SERRATO, JUAN CARLOS SUÁREZ, & RAMIREZ, KARINA. 2020. Taxing Property in Developing Countries: Theory and Evidence from Mexico. *Working Paper*.

CENTRE FOR AFFORDABLE HOUSING FINANCE IN AFRICA. 2019. 2019 Housing Finance Yearbook: Senegal profile.

COGNEAU, DENIS, DUPRAZ, YANNICK, & MESPLÉ-SOMPS, SANDRINE. 2020a. Fiscal Capacity and Dualism in Colonial States: The French Empire 1830-1962. *Mimeo.*

COGNEAU, DENIS, DUPRAZ, YANNICK, & MESPLÉ-SOMPS, SANDRINE. 2020b. Fiscal Capacity and Dualism in Colonial States: The French Empire 1830-1962. Data Appendix. *Mimeo.*

COGNEAU, DENIS, DUPRAZ, YANNICK, KNEBELMANN, JUSTINE, & MESPLÉ-SOMPS, SANDRINE. 2020c. Taxation in the Former French Empire from Colonial Times to Present. *Work in progress*.

COLONIE DU SÉNÉGAL. 1897. Budget du service local, colonie du Sénégal. 1897.

COLONIE DU SÉNÉGAL. 1904. Budget du service local, colonie du Sénégal. 1904.

COLONIE DU SÉNÉGAL. 1931. Budget du service local, colonie du Sénégal. 1931.

COLONIE DU SÉNÉGAL. COMMUNE DE DAKAR. 1947. Budget Primitif. Exercice 1947.

DAVIS, PEADAR, MCCLUSKEY, WILLIAM, GRISSOM, TERRY, & MCCORD, MICHAEL. 2012. An em-

pirical analysis of simplified valuation approaches for residential property tax purposes. *Property Management*, **30**(3), 232–254.

DE CARVALHO JR., PEDRO HUMBERTO BRUNO. 2015. Distributive aspects of real estate property and its taxation among brazilian families. Institute for Applied Economic Research Discussion Paper.

DIRECTION GÉNÉRALE DES IMPÔTS ET DOMAINES DU SÉNÉGAL. 2013. Aperç historique de l'évolution de l'organisation des structures de la direction générale des impôts et des domaines. Rapport.

FISH, PAUL. 2018. Practical Guidance Note: Training Manual for Implementing Property Tax Reform with a Points-Based Valuation. ICTD African Tax Administration Paper 2.

FRANCE. MINISTÈRE DE LA MARINE ET DES COLONIES. 1887. Colonie du Sénégal. Service Local. Compte définitif des recettes et des dépenses de l'exercice 1883.

FRANCE. MINISTÈRE DE LA MARINE ET DES COLONIES. 1897. Colonie du Sénégal. Service Local. Compte définitif des recettes et des dépenses de l'exercice 1895.

FRANCE. MINISTÈRE DE LA MARINE ET DES COLONIES. 1923. Colonie du Sénégal. Budget des territoires d'administration directe. Compte définitif des recettes et des dépenses de l'exercice 1920.

FRANKEMA, EWOUT, & VAN WAIJENBURG, MARLOUS. 2014. Metropolitan Blueprints of Colonial Taxation? Lessons from Fiscal Capacity Building in British and French Africa, 1880-1940. *Journal of African History*, **55-3**, 371–400.

FRANZSEN, RIEL, & MCCLUSKEY, WILLIAM. 2017. *Property Tax in Africa: Status, Challenges and Prospects*. Lincoln Institute of Land Policy, Cambridge, Massachusetts.

GLOBAL PROPERTY GUIDE. 2014. Senegal online Property Market Overview.

GOUVERNEMENT GÉNÉRAL DE L'AFRIQUE OCCIDENTALE FRANÇAISE. 1925. Budget de la circonscription de Dakar et deépendances (annexe au budget général). Exercice. 1925.

GOUVERNEMENT GÉNÉRAL DE L'AFRIQUE OCCIDENTALE FRANÇAISE. 1931. Colonie du Sénégal. Régime Fiscal.

GOUVERNEMENT GÉNÉRAL DE L'AFRIQUE OCCIDENTALE FRANÇAISE. 1937. Budget de la circonscription de Dakar et deépendances (annexe au budget général). Exercice. 1937.

GOUVERNEMENT GÉNÉRAL DE L'AFRIQUE OCCIDENTALE FRANÇAISE. 1942. Budget de la circonscription de Dakar et deépendances (annexe au budget général). Exercice. 1942.

GOUVERNEMENT GÉNÉRAL DE L'AFRIQUE OCCIDENTALE FRANÇAISE. 1946. Colonie du Sénégal. Régime Fiscal.

GUAN, JIAN, ZURADA, JOZEF, & LEVITAN, ALAN. 2011. A comparison of regression and artificial intelligence methods in a mass appraisal context. *Journal of Real Estate Research*, **33**(3), 349–387.

IMF. 2019. Senegal Article IV Report.

KHAN, ADNAN Q., KHWAJA, ASIM I., & OLKEN, BENJAMIN A. 2016. Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors. *The Quarterly Journal of Economics*, **131**(1), 219–271.

KOPANYI, MIHALY. 2015. Local Revenue Reform of Kampala Capital City Authority. **C43306UGA1**. Working Paper.

LASALLE, JONES LANG. 2015. Emerging Beyond the Frontier: An Overview of Sub-Saharan Africa's Real Estate Capital Markets. Jones Lang LaSalle (JLL).

MANWARING, PRIYA, & REGAN, TANNER. 2019. Property taxes in Kampala. International Growth Center Policy Note 43448.

MBAYE, AHMADOU ALY. 2019. Supporting small informal businesses to improve the quality of jobs in Africa. Brookings Africa Growth Initiative Policy Brief.

MCCLUSKEY, WILLIAM, MCCORD, M., DAVIS, P., HARAN, M., & MCILHATTON, D. 2013. Prediction accuracy in mass appraisal: a comparison of modern approaches. *Journal of Property Research*, **30**(4), 239–265.

MONKAM, NARA, & MOORE, MICK. 2015. How Property Tax would benefit Africa. Africa Research Institute Counterpoints.

MOORE, WAYNE. 2005. Performance Comparison of Automated Valuation Models. *Journal of Property Tax Assessment and Administration*, **3**(1), 43–60.

NDIAYE, CHEIKH. 2020. *Finances et Comptabiliteé des Collectivités Territoriales au Sénégal*. L'Harmattan Sénégal.

NORREGAARD, JOHN. 2013. Taxing Immovable Property, Revenue Potential and Implementation Challenges. IMF Working Paper.

OATES, WALLACE, & FISCHEL, WILLIAM. 2016. Are local property taxes regressive, progressive, or what? *National Tax Journal*, **69 (2)**, 415–434.

OKUNOGBE, OYEBOLA. 2020. State Capacity and Taxation: the Role of Detection and Enforcement Capacity on Property Tax Compliance in Liberia. mimeo.

PIKETTY, THOMAS. 2001. *Les hauts revenus en France au XXe siècle. Inégalités et redistributions (1901-1998).* Paris, Grasset.

PIKETTY, THOMAS. 2020. *Capital and Ideology*. Harvard University Press.

SAEZ, EMMANUEL, & ZUCMAN, GABRIEL. 2019. Progressive Wealth Taxation. *BPEA Conference Drafts*.

SAHEL AND WEST AFRICA CLUB. 2018. Africapolis Dataset.

SECK, ASSANE. 1961. Dakar. *Cahiers d'outre-mer*, **56**, 372–392.

SENNOGA, EDWARD, SJOQUIST, DAVID, & WALLACE, SALLY. 2007. Incidence and Economic Impacts of Property Taxes in Developing and Transitional Countries. Andrew Young School of Policy Studies Research Paper.

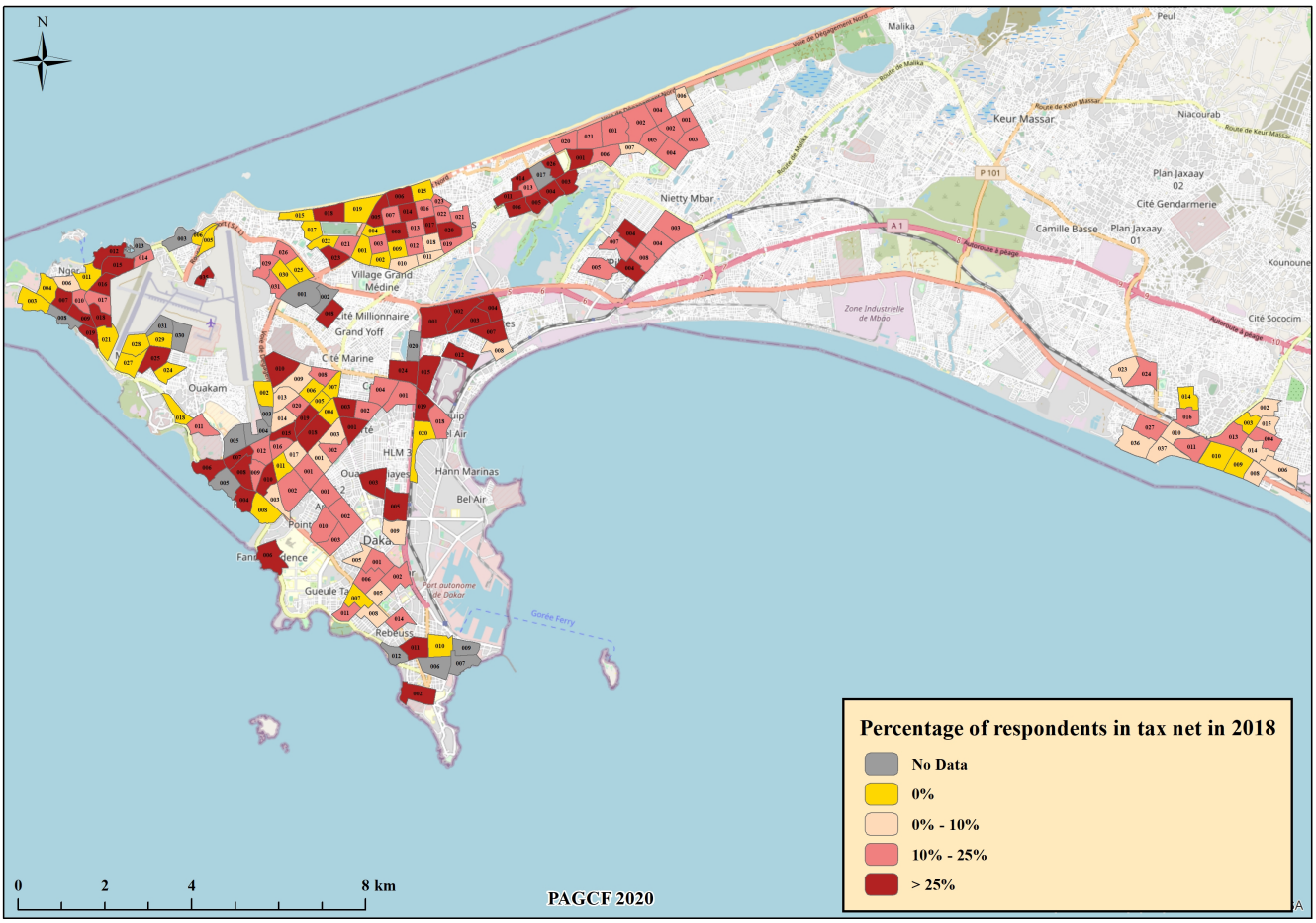
WEIGEL, JONATHAN. 2018. The taxman cometh: A virtuous cycle of compliance and state legitimacy in the D.R. Congo. *Working Paper*.

WEIGEL, JONATHAN. 2020. The Participation Dividend Of Taxation: How Citizens In Congo Engage More With The State When It Tries To Tax Them. *The Quarterly Journal of Economics*.

ZEBONG, NYAH, FISH, PAUL, & PRICHARD, WILSON. 2017. Valuation for Property Tax Purposes. ICTD Summary Brief 10.

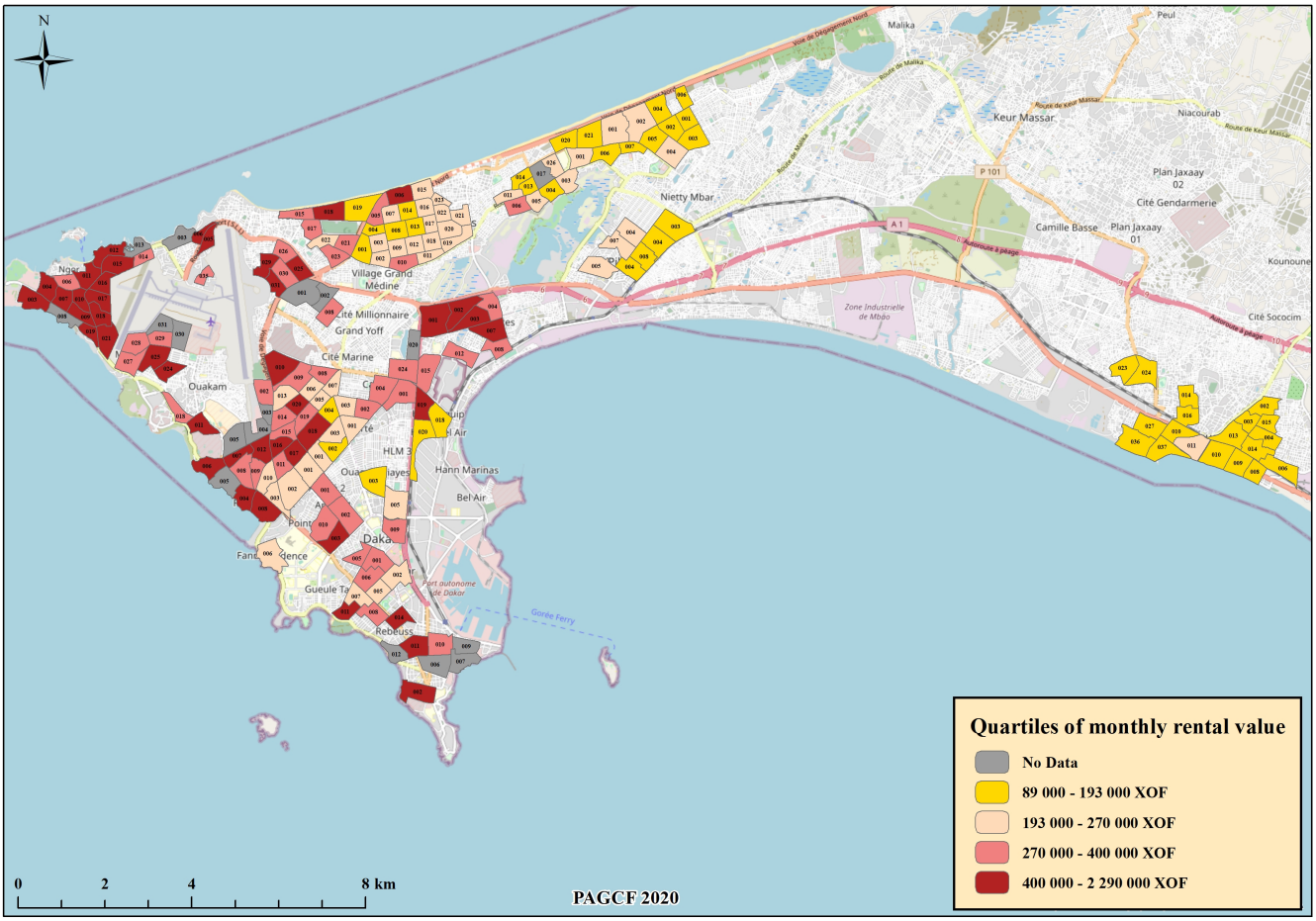
Figures

FIGURE 1
PROPERTY TAX COMPLIANCE BY CADASTRAL SECTION IN THE REGION OF DAKAR



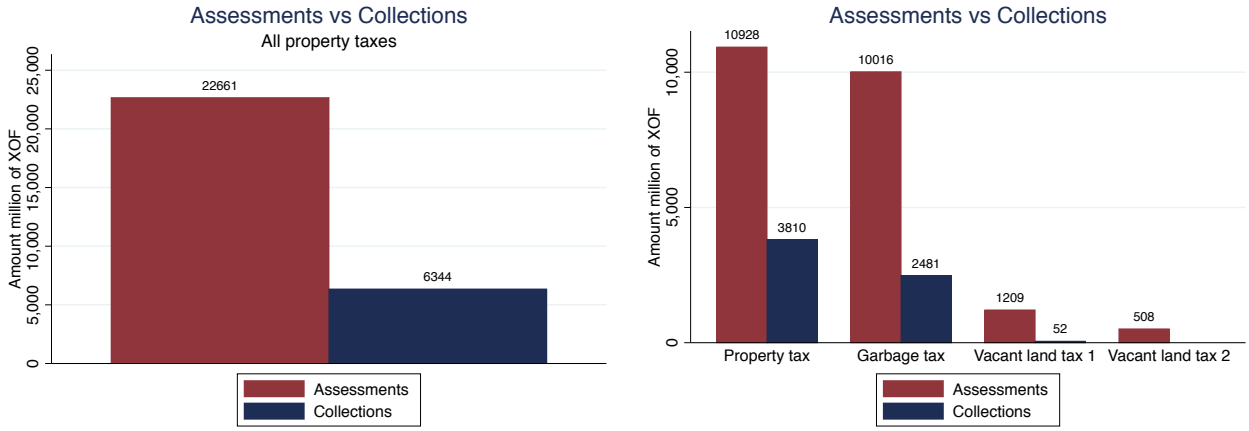
Notes: This map represents the region of Dakar and the cadastral sections covered in the property owner survey. Sections are divided into four groups based on the share of respondents who are in the property tax net in 2018 (meaning that they paid the tax and/or received a tax notification). The share of owners in the tax net is computed based on self-declared questions: *Did you pay the property tax for this property in 2018?*, *Did you receive a tax notification related to the property tax for this property in 2018? Did you ever pay the property tax for this property ? If so, when was the last time you did so?* The sample size is 2, 474 across 181 sections. Source: property owner survey 2018, DGID cadastral data.

FIGURE 2
MEDIAN REAL ESTATE VALUES BY CADASTRAL SECTION IN THE REGION OF DAKAR



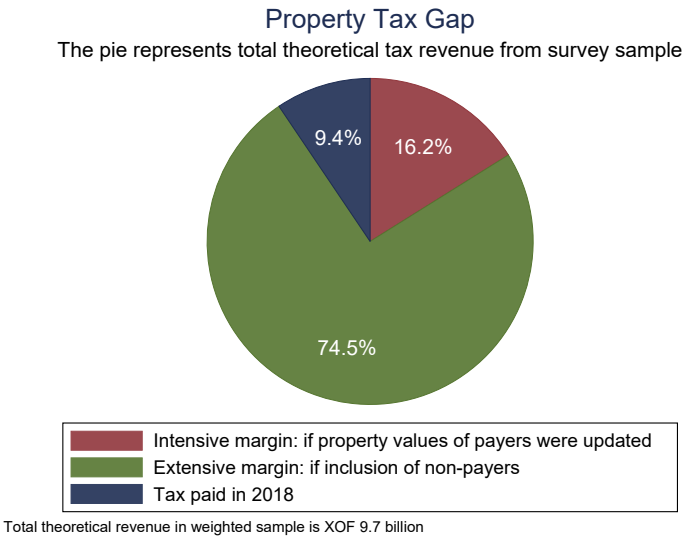
Notes: This map represents the region of Dakar and the cadastral sections covered in the property owner survey. Sections are divided into quartiles of median property value. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. The unit of observation is a property, irrespective of size. The first quartile includes sections where the median monthly rental value is below XOF $193,000$ (USD 350). The fourth quartile includes sections where the median monthly rental value is above XOF $400,000$(USD 725). The sample size is $2,443$ across 181 sections. Source: property owner survey 2018, DGID cadastral data.

FIGURE 3
ASSESSMENTS VERSUS REVENUE COLLECTIONS



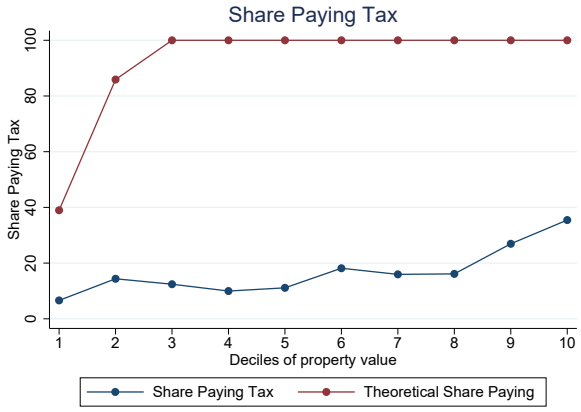
Notes: This Figure shows total amounts of assessments vs revenues, aggregated at the country level, for 2019. Source: assessment data from the tax administration, data on the revenues of local governments from the Treasury (2019).

FIGURE 4
BREAKDOWN OF THE PROPERTY TAX GAP



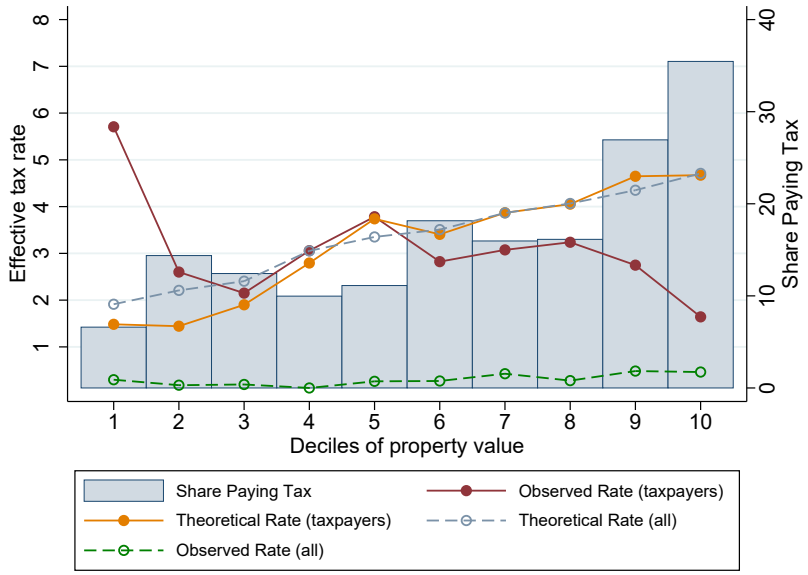
Notes: This pie chart represents a breakdown of the property tax gap as measured from the property owner survey. The pie area corresponds to total theoretical tax revenues for the whole weighted sample of properties (XOF billion 9.7). The blue slice represents the share actually paid. The red slice represents the intensive margin, the share of revenue that would be collected if respondents who paid in 2018 had paid their theoretical tax liability in full, as calculated based on the value of their property. The green slice represents the extensive margin, the share of revenue that would be collected if respondents who did not pay in 2018 paid their theoretical tax liability. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted to be representative of all properties eligible for the survey.

FIGURE 5
SHARE OF OWNERS PAYING PROPERTY TAX



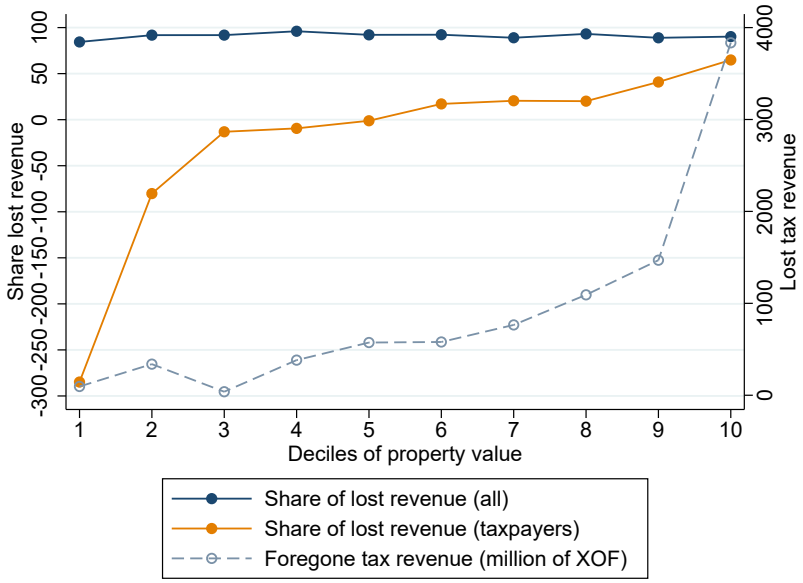
Notes: This figure plots the share of owners who paid the property tax in 2018, for each decile of property value. The share of payers is computed based on self-declared questions: *Did you pay the property tax for this property in 2018?* and *Did you ever pay the property tax for this property ? If so, when was the last time you did so?*. The theoretical share of payers is computed based on property value and fiscal modalities of the property tax, notably, an abatement for home-occupied properties. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. For details on the data and computations, see Appendix A.4. The sample size is $2,443$ (observations are weighted). Source: property owner survey 2018.

FIGURE 6
FISCAL PRESSURE BY DECILE OF PROPERTY VALUE



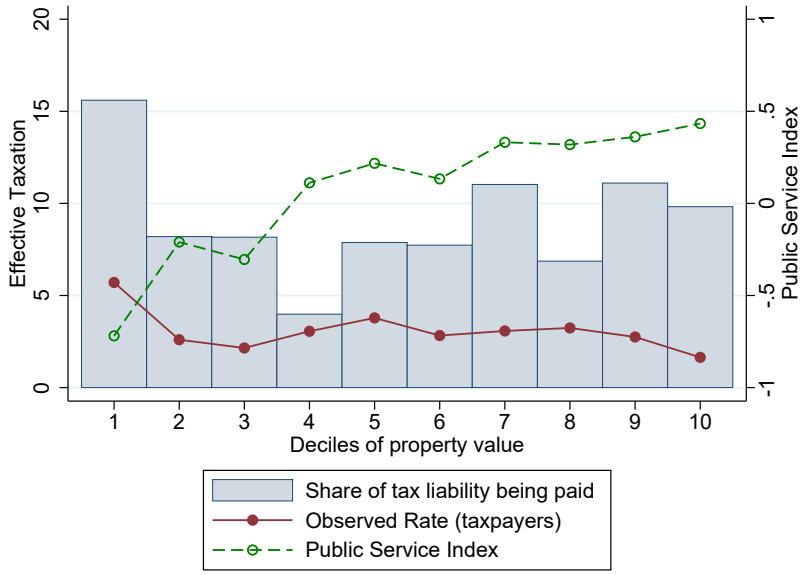
Notes: This graph shows different measures of fiscal pressure for each decile of property value. The grey bars indicate the share of owners in each decile that paid the property tax in 2018 (right-hand y-axis). The green dashed line represents the average effective tax rate observed in each decile - computed as the ratio of total tax amounts over total property value. The grey dashed line shows the theoretical tax rate, if all owners paid their liability in full. The orange and red lines show respectively the observed effective tax rate, and the theoretical rate, restricting the sample to owners who do pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

FIGURE 7
FOREGONE TAX REVENUE BY DECILE OF PROPERTY VALUE



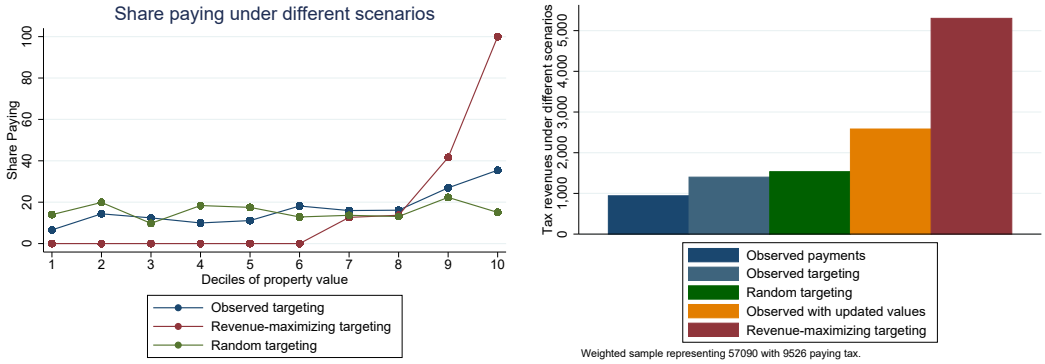
Notes: This graph shows property tax revenue that is lost due to weak enforcement for each decile of property value. Lost revenue in a given decile is computed as total theoretical tax liability minus total tax payments. The grey dashed line shows amounts of lost revenue in XOF million (right-hand y-axis). The blue line shows foregone revenue as a share of theoretical revenue within each bin. The orange line shows foregone revenue as a share of theoretical revenue, restricting the sample to owners who pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted to be representative of all properties eligible for the survey.

FIGURE 8
EFFECTIVE TAX RATE AND LOCAL PUBLIC SERVICE



Notes: This graphs shows fiscal pressure and availability of local public services by decile of property value. The grey bars show the share of theoretical tax liability being paid, within each decile. The red line shows the effective tax rate conditional on paying. The green dashed line shows the average standardized local public service score, computed as the sum of scores on existence of street lights, type of sewage system, type of water supply, existence of a garbage removal service. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. Local public services are declared and at the property level. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

FIGURE 9
SIMULATING A REVENUE-MAXIMIZING TARGETING



Notes: These figures compare different scenarios of targeting by the tax administration, keeping constant the number of payers. The left panel plots the share of payers by decile of property value under each scenario. The right panel shows total revenues under the different scenarios in XOF millions. *Observed payments* shows total tax revenues as declared in the survey. *Observed targeting* is the targeting scenario measured in the survey, including all respondents who declare paying the tax in 2018. For observations where tax amount paid is missing, it is imputed using theoretical tax. In the *Random targeting* scenario, the payer status is randomly assigned. In the scenario labeled *Observed with updated values* the payers are those counted in the survey, but all tax amounts are switched to theoretical tax liabilities. Under the *Revenue-maximizing targeting* scenario, the payer status is assigned to properties of highest value. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

Tables

TABLE 1
PROPERTY TAX GAP MEASURED IN SURVEY

N# Surveyed	2,474
Respondent and Property Characteristics	
% Respondent is owner	66.1
% Sole owner	80.6
% Respondent lives in surveyed property	92.8
% Rented (at least partly)	30.3
% Commercial (at least partly)	23.6
% Respondent gave rental value	69.1
% Declared or imputed rental value	99.1
Average monthly rental value (XOF)	369,643
Median monthly rental value (XOF)	250,000
% Knows property tax	59.2
Property Tax Gap	
% Ever received tax notification	38.2
% Ever paid property tax	43.1
% Received 2018 tax notification	12.4
% Paid 2018 property tax	16.8
% In tax net 2018	19.7
Average tax paid (XOF, payers only)	16,499
Average theoretical tax amount (XOF, payers only)	271,116
Average theoretical tax amount (XOF, all)	177,106
Total tax paid (million XOF)	949
Total theoretical tax (million XOF)	10,111
Tax amount per capita (owners, XOF)	16,478
Tax amount per capita (full household, XOF)	1,833

Notes: This table displays descriptive statistics from the property owner survey. For details on the survey sampling and protocol, see Appendix A.2. For details on the data and computations, see Appendix A.4. *% Respondent gave rental value* is the share of observations for which there is a declared property value. *% Declared or imputed rental value* is the share of observations for which there is a property value once we add expert values and predicted values, when declared value is missing. *Monthly rental value* is the market value for the whole property, whether it is partly rented or not. *In tax net* is defined as having paid, or having received a tax notification. *Theoretical tax amount* is computed using property values and fiscal modalities of the property tax. *Total tax paid* and *Total theoretical paid* are computed for the weighted sample. Source: property owner survey 2018, observations are weighted.

TABLE 2
DETERMINANTS OF TAX ENFORCEMENT AT THE CADASTRAL SECTION LEVEL

Dependent Variable	At least one in Tax Net	Share in Tax Net	Share in Tax Net
Ln(Average Property Value)	0.060	0.117***	0.141***
	(0.059)	(0.034)	(0.038)
Share with street name	0.159	0.057	-0.026
	(0.102)	(0.059)	(0.062)
Distance to closest Treasury	-0.046***	-0.006	0.006
	(0.016)	(0.009)	(0.010)
Distance to closest Tax office	0.042	0.018	0.014
	(0.030)	(0.017)	(0.018)
N	181	181	143
R2	0.07	0.06	0.16
Mean of dep.	0.79	0.22	0.28

Notes: This table shows results from an OLS regression at the cadastral section level. In column (1) the dependent variable is a dummy equal to one if at least one respondent in the section is in the tax net in 2018. Being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared. In columns (2) and (3) the dependent variable is the share of respondents in the section which are in the tax net. In column (3) the sample is restricted to sections where at least one respondent is in the tax net. *Property value* (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. *Share with street name* is a proxy for cadastral quality and is computed as the share of surveyed properties for which a street name was entered. *Distance to closest Treasury* is the average number of kilometers between properties in a given section and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the average number of kilometers between properties in a given section and the closest Tax office, where agents who conduct registration and assessment activities work from. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018.

TABLE 3
ECONOMIC DETERMINANTS OF FISCAL PRESSURE AT THE INDIVIDUAL LEVEL

	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
Dependent Variable (0,1)	TaxNet	TaxNet	TaxAmnt	TaxAmnt	TaxAmnt	TaxAmnt
Ln(PropertyValue)	0.071**	0.062***	0.235***	0.199***	0.165	-0.031
	(0.031)	(0.020)	(0.060)	(0.059)	(0.160)	(0.487)
Multiple owner	0.131**	0.051	0.113	0.006	0.816**	0.353
	(0.059)	(0.042)	(0.155)	(0.206)	(0.315)	(0.799)
Income Group 2	0.014	0.041	0.026	0.107	0.332	0.451
	(0.036)	(0.033)	(0.082)	(0.097)	(0.339)	(0.886)
Income Group 3	0.049	0.066*	0.032	0.083	0.264	0.490
	(0.032)	(0.039)	(0.088)	(0.118)	(0.303)	(0.636)
Income Group 4	0.052	0.127***	0.259**	0.326**	0.755**	1.087
	(0.056)	(0.048)	(0.108)	(0.127)	(0.345)	(0.676)
Rented	0.036	0.043	0.119*	0.105	0.377*	0.400
	(0.044)	(0.035)	(0.068)	(0.074)	(0.197)	(0.583)
N with regular income	-0.022**	-0.015	-0.029	-0.028	-0.025	0.020
	(0.011)	(0.011)	(0.024)	(0.025)	(0.070)	(0.161)
Informal	-0.067	-0.043	0.157	0.124	0.169	-0.678
	(0.097)	(0.086)	(0.223)	(0.209)	(0.333)	(0.595)
Non Employed	-0.118**	-0.086**	-0.204**	-0.231**	0.422	-0.363
	(0.059)	(0.039)	(0.093)	(0.099)	(0.258)	(0.604)
Retired	-0.055	-0.044	-0.119	-0.245*	0.035	-0.795
	(0.058)	(0.040)	(0.096)	(0.127)	(0.240)	(0.841)
Section FE	No	Yes	No	Yes	No	Yes
N	1108	1108	1106	1106	118	118
Adj R2	0.07	0.30	0.09	0.27	0.29	0.29
Mean of dep.	0.20	0.20	0.32	0.32	2.68	2.68
0.5 Cutoff						
% Correct	0.83	0.86				
Sensitivity	0.01	0.33				
Specificity	1.00	0.97				
Index	0.01	0.30				
0.3 Cutoff						
% Correct	0.78	0.85				
Sensitivity	0.33	0.62				
Specificity	0.88	0.90				
Index	0.21	0.52				

Notes: This table shows results from OLS regressions analyzing economic correlates of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). In columns (3) to (6), the dependent variable is $asinh(TaxAmnt)$, the inverse hyperbolic sine transformation of tax amount paid. I control for section fixed effects in columns (2), (4) and (6). In columns (5) and (6), the sample is restricted to owners who pay the tax. *Property value* (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE 4
PROPERTY TAX CAPACITY IN THE COLONIAL AND POST-COLONIAL PERIOD

	1897	1931	1942	2018
Property tax revenue per capita in Dakar (XOF)	6,530	10,006	2,975	1,650
Property tax revenue as a share of local tax revenue		29.4%	22.8%	10.4%
Property tax revenue as a share of total tax revenue	1.8%	1.9%	3.2%	0.3%

Notes: This table displays figures on property tax capacity over time. Property tax revenue includes property tax per se and assimilated taxes (*contribution mobilière* in the colonial period and garbage tax in the post-colonial period). Property tax revenue per capita is in 2019 XOF. Local taxes are total local taxes in the region of Dakar. Total taxes are total national tax revenues. For more detail on historical data sources and variables, see Appendix A.8. For year 2018, total property tax revenue is computed by applying to 2018 total assessments the collection ratio observed in 2019 (27 percent). I use 2018 because I do not have the figure for total local tax revenue in 2019. Property tax revenue per capita in 2019 is XOF 1,700.

A Appendix

A.1 Real estate sector in Dakar

Dakar has been growing rapidly over the past years, and the real estate sector has been particularly vibrant, which suggests an important potential for tax revenues. Since the tax base for the recurrent property tax is the market rental value of the property (and thus is not fixed over time, it should increase when market rents increase), soaring rents mechanically correspond to an increasing tax base. Furthermore, these also constitute the tax base for taxes on rental income (that we do not study in this paper).

Figure A.1 displays headlines in the press stressing the strong increase in property prices and rents in recent years. Figure A.2 displays information from an intelligence website about worldwide property investment (Global Property Guide), showing that Dakar is positively depicted for the buoyancy of its real estate by international standards. Below, we list some excerpts from reports on Dakar's real estate market.

Centre for Affordable Housing Finance in Africa (2019) provides some background information on the whole of Senegal. The housing sector grew by 5 percentage points between 2017 and 2018, and "the surge in housing prices is the result of a long and unprecedented process of speculation combined with the scarcity of land and the high cost of land and building materials". Some factors of this "huge expansion" are: "development programmes, including the construction of highways and roads, the creation of new urban areas and the massive production of housing. Similarly, population growth and the reputation of a stable democratic country provide a favourable environment for investors". The report also mentions a strong housing deficit (as demand exceeds supply) estimated at 300,000 housing units in 2018 for the whole of Senegal.

Focusing on the capital city, the investing guide platform Global Property Guide (2014) mentions: "Dakar's residential rents have risen continuously over the past two decades, and now often rival rents seen in large European cities"; "Rents in Dakar's downtown Plateau district are almost twice those in the central business district of Abidjan, Ivory Coast's commercial capital"; "Yields in Dakar are high"; to meet demand, the stock of housing units should increase by 10 percent every year.

Finally, the national Statistical agency takes stock of the rental market in a 2012 report: "The average monthly rental value in the region of Dakar increased by 115.8 percent between 1994 and 1999, then by 38.6 percent between 2000 and 2009"; "the price per square metre of land increased by almost 2.5 times between 1994 and 2000, and then by more than double between 2000 and 2009"; "Annual turnover from housing rental activities (...) amount to 32.9 percent of real estate activities and 2 percent of services in the economy." (Agence Nationale de la Statistique et de la Démographie, 2012).

A.2 Baseline survey

A.2.1 Sample selection for the property owner survey

The baseline survey was conducted between September and December 2018. 202 cadastral sections were selected for the survey. They were designated by the tax administration as eligible for property tax modernization – and thus also relevant to investigate the hurdles to property taxation – based on the following criteria: up-to-date cadastral data, not corresponding to informal settlements, not corresponding to traditional villages (these are specific areas within the city of Dakar).[64]

The population of interest for the survey were individual property owners, excluding properties owned by firms or administrations. Using cadastral data, 20 plots to be surveyed were randomly drawn in each of the 202 eligible sections. 20 to 40 replacement plots were additionally drawn in each section. If three visits to a targeted plot were unsuccessful, enumerators were to visit a replacement plot. Eight sections were dropped from the sample after the survey because the field work revealed that they were almost exclusively industrial areas. For each visited plot, enumerators first had to assess whether the plot was eligible for the survey or not. Eligibility was defined according to the following criteria: there is a construction on the plot, the owner is an individual (and not a firm nor a public administration), the construction is maximum 5 floors high.[65] If the plot turned out to be ineligible for the survey, the enumerators were meant to select another property from the replacement list within the same section.

The targeted respondent was the owner of the property on the plot. If there were several properties, owned by different owners, on the same plot, surveyors were supposed to survey at least one owner.[66] If the owner was temporarily absent, the enumerators were supposed to come again another day (up to three times), and/or to plan a meeting with the owner. If the owner did not live on the property, but elsewhere in Dakar, enumerators were supposed to obtain his or her contact details and conduct the interview in a location suitable to the respondent. If the owner was absent for a longer period – living outside of Dakar, or for a long trip – enumerators were required to find the person who was in charge of administrative issues pertaining to the property. These respondents were most often a relative of the actual owner.

The neighborhood delegates of the corresponding areas were also surveyed, using a dedicated questionnaire. In total, 184 delegates were surveyed. This data is not used in the current version of the paper.

The survey was carried out by a private survey company. The data has not and will not be shared with the DGID. Accordingly, respondents were informed that their answers would not be communicated to any public administration and would be used anonymously for research purposes.

[64] The survey was designed as the baseline survey for an ongoing randomized experiment testing a property tax modernization program.

[65] Higher buildings were excluded because finding the owner in these cases is particularly challenging. These cases represented only 4.7 percent of ineligible properties. We also know from external sources that buildings represent a minor share of the real estate in Dakar: according the 2013 census, only 3.7 percent of inhabitants of Dakar live in a building with multiple floors or an apartment block, versus a house (Agence Nationale de la Statistique et de la Démographie, 2020). Vacant land or a building that is under construction represented 50 percent of ineligible cases. A firm or administration as the owner represented an additional 15 percent.

[66] There were more than one owner on the plot for 23 percent of cases, however the majority of these were cases were a single property unit was shared among family members. There were multiple owners with a division in property units on the plot only in 1.2 percent of the cases.

A.2.2 Representativeness

In this section we discuss the characteristics of the survey that hinder its representativeness, and provide arguments as to why we consider that the data is nonetheless insightful to learn about property taxation of properties owned by individuals.

The first dimension of selection is the restriction to eligible cadastral sections. As such, we do not make any claim that the survey is representative of the whole region of Dakar. Rather, it should be thought of as aiming to represent those areas within the region that have tax potential in terms of existing real estate, and also in which the administration is willing and capable of intervening to improve property taxation. In the city of Dakar, most of the urban area is included in the sample (146 of surveyed sections out of 194). The reasons of exclusion of the sections that were excluded are mainly: market areas, informal settlements, traditional villages, lack of digitized cadastral data. In the cities of Pikine and Guédiawaye, the reasons of exclusion of excluded sections are mostly: lack of tax potential (low value areas, informal settlements), and absence of up-to-date cadastral information. In the city of Rufisque the main reason of exclusion of excluded sections are low density density and few built properties (many areas are mostly vacant land, or quasi-rural). The sections that were selected corresponds to the part of Rufisque that is urban.

The second margin of selection into the survey is the definition for eligibility of plots. 8,251 properties were visited, out of which 5,353 (65 percent) were eligible. The exclusion of vacant land and buildings under construction (50 percent of ineligible cases) does not appear to be very problematic for the study, since these properties are not subject to the tax on built property.[67] Similarly, since we focus in this paper on the taxation of property owned by individuals, and that data on this tax head can be separated in the administrative data from firm taxation, exclusion from the survey of properties owned by firms or administrations does not seem to be too problematic.[68] The only margin of non-eligibility that is problematic for the study is the criteria based on the number of floors – exclusion of buildings of five floors or more. These properties are mechanically associated with higher tax potential. Their compliance rate could differ from lower units since they are more visible. We cannot test these features, but it is reassuring that they only correspond to a small minority of units in the region, and a small percentage of visited plots (4.7 of ineligible plots). As such, it seems unlikely that their inclusion would dramatically change the results of the analysis.

The response rate of the survey is 46 percent (2,474 surveys conducted out of 5,353 eligible plots). Non-response is due to two factors. First, cases where the respondent was not found – this is the strongest driver of non-response, a respondent was found in only 48 percent of eligible plots. Second, cases where the respondent was found but refused to answer. These were rare (95 percent of acceptance conditional on finding the respondent). It is likely that respondents who are more difficult to find have different characteristics than those easier to interact with. Typically, owners that are harder to find are more likely: to own the property without living in it; to be employed and have longer working hours; possibly, to have strict rules according to which their tenants or house workers are not allowed to communicate their contact details to visitors; to live in compounds where it isn't possible to enter or find someone to talk to without having been formally let in. To assess the extent to which this is the case, one comparison we can carry out is to benchmark the survey data against a representative sample of city dwellers. We use a national census data from 2002 (data is from the national statistical agency, ANSD, accessed through IPUMS).[69] The census

[67] Including vacant land would probably reveal a larger tax gap, since taxes on vacant land exist, but their functioning appears to be even more flawed than for the tax on built property (as per the extremely low number of entries for vacant land on the valuation roll).

[68] If the study was extended to properties owned by firms, it is possible that the tax gap would be smaller.

[69] The census and the survey are sixteen years apart, which means a lot of social and economic factors may have

sample includes 29,369 households in the region of Dakar of equal weight and is representative of the region. A variable on home-ownership allows to identify households occupied by the property owner. Comparative statistics are displayed in Table A.5. The share of owner-occupied properties in the census is 45 percent,[70] while it is 65 percent in the survey sample. This confirms that the survey over-represents owner-occupied properties. We also compare the survey sample with the census data on a subset of socio-demographic characteristics. We restrict both the survey and census data to owner-occupied properties, and for census variables, we compute socio-demographic statistics on household head.[71] Average age (53.4 years old) is close to the one of our survey sample (55.6). We further restrict the survey sample to respondents are the legal owner themselves (66 percent of respondents, excluding for instance spouse of the owner). In this case, these owners are significantly older than the average household head in owner-occupied properties from the census. Other statistics also reveal some differences: the share of retired is high and similar in both samples, but the survey displays a larger share of non-employed (vs employed) than in the census. We observe significantly higher educational attainment in the survey sample (27 percent with higher education, against 6 percent). Furthermore, looking at property characteristics, the share of properties with no piped water is lower in the survey, trash collection, sewage, electricity are higher. It is not clear whether these differences should be interpreted as a selection issue: they could reasonably be attributed to the time span between the two data collections - the generation of owners in the census could be the generation preceding the one from the survey sample. As such it will be important to update these figures with data from the 2013 census when it is available.
Overall, if we consider that non-response is not significantly associated to property and owner characteristics, we can consider that the survey is representative of the 59,505 eligible plots within these sections. More conservatively, the sample is representative of the 32,370 properties for which the respondent would have responded (accounting for non-response rate by section).
Enumerators introduced themselves to respondents as working for an independent survey company, and presented the survey as a research project on local governance and citizens' perception of local authorities. They made clear that responses would only be used for research purposes and not be communicated to any public administration.[72] The issue of taxation was mentioned as one of the topics among others that would be addressed, but not highlighted as the main motivation for the survey. Indeed, only a subset of the questions were directly about taxation, and they were at the end of the survy, to avoid possible contamination of other replies. For these reasons, and thanks to observational work carried out during the survey, it does not seem likely that respondents would intentionally manipulate their responses by fear of having their replies transmited to the tax administration. It is also not clear why respondents would see interest in manipulating the property values or rents they declare. To check whether the owners of properties with higher values – that might potentially be most worried about tax enforcement – are more likely not to provide answers to questions on property values, we conduct a " balance check" in Table A.1, where we compare expert values and predicted property values based on the valuation formula, across cases where the owner declared versus did not declare the rental value. Expert values are 21 percent higher for properties with no owner-declared value (row 1) and predicted values 7 percent higher for properties with no owner-declared value (row 2), but in none of these two cases is the difference significant at the 5 percent level.

evolved over time. A census was conducted in 2013 but the data was not accessed yet.

[70] In the 2013 census, the figure is 41.3 percent.

[71] We do not have information as to whether the household head is at the same time the legal owner of the property, but we assume this is a widespread situation.

[72] In line with our IRB guidelines, a consent message was read out to respondents, giving them the opportunity to refuse to take the survey or to drop out at any point.

A.2.3 Sample selection for the expert valuation survey

Subsequently to the baseline survey, a small-scale real estate expert survey was conducted. The objective was to obtain more objective rental values, in addition to the ones declared by the owners. A subsample of properties ($N = 441$) were drawn from the baseline sample to be (externally) visited by real estate experts.[73] Four experts were hired and were required to provide lower bound and higher bound estimates of the monthly market rental value of each property. They were equipped with the precise address information of the property (cadastral identifier and maps), photos when available, and built area as measured from satellite images.

The properties to be assessed were selected as follows. Among all surveyed properties ($N = 2,474$) we selected all for which a photo was taken during the property owner survey ($N = 222$, this feature was implemented randomly in the survey). This minimizes potential confusion on which property is to be assessed.

We wanted the breakdown by city[74] to be the same in the total survey sample and in the expert valuation sample, so we drew properties in a way to maintain the same proportions.

We also decided to slightly over-represent properties that were located in residential areas, meaning wealthier areas, because enumerators' perception was that it was more difficult to obtain responses to this question in these areas.[75]

[73] The restriction to a small number of properties is due to budget and logistical constraints.

[74] The four cities are Dakar, Rufisque, Guédiawaye, Pikine.

[75] We test this in section A.4 and find that conditional on being surveyed, the difference in (externally measured) rental values for properties where the owner declares the value, and properties where the owner did not reply to this question, is small and not significant.

A.3 Property valuation formula

Using the data on property characteristics and rental values collected in the survey, we calibrate a property valuation formula based on hedonic regressions. This formula has a policy purpose since it is integrated in the property tax modernization program for which this paper is a preliminary study. Second, we use the predicted values based on the formula to impute missing property values in the analysis throughout this paper.

A.3.1 Selecting the specification for the hedonic regression

The baseline specification can be written as:

$$Ln(Annual\ Rental\ Value) = Constant + Ln(Total\ Built\ Area) + Sector + \beta X + \epsilon$$

where X is the vector of observable characteristics.

Table A.19 shows the property characteristics used in the model. Built areas was measured using high resolution satellite and drone images, through manual image processing by GIS experts. Total built area is computed by combining built area and the number of floors observed during the survey. As a location variable in the hedonic regressions, we use "Sectors" – they are geographical areas defined in a 2010 Senegalese decree pertaining to rent regulation, and were already used by the Valuation department in the estimation of property values. GIS experts coded up the delimitation of these sectors to be able to relate each property to its sector. In the legislation, the price of a square meter of land is associated to each sector. However, we choose an agnostic approach, and simply use the sector classification as a categorical variable, without imposing a ranking between the different sectors $ex\ ante$.

To choose the correct hedonic regression specifications, we relied on previous work, notably Franzsen & McCluskey (2017); Davis *et al.* (2012); McCluskey *et al.* (2013); Guan *et al.* (2011); Moore (2005); Ali *et al.* (2018); Fish (2018) and also interacted with practitioners and experts (among which, William McCluskey, Riël Franzen, Mihaly Kopanyi, Pedear Davis, Paul Fish).[76] We use cross-validation, first to compare different specifications and pick the preferred one, and second to obtain final coefficients that are not too dependent on the specific sample. We also compare results when the formula is calibrated using expert values instead of survey values.

Data cleaning and corrections

We reclassify some observations into the different modalities of observable characteristics, when some modalities had a very small number of observations (eg, bringing "zinc" and "metal" into a unique category for fence type), and when the enumerator listed "Other" but added information which allowed to associate the response to one of the pre-existing categories. Note that for all categorical variables, we set the reference category in the regressions to be the most frequent or

[76] The property valuation work stems from our collaboration with the administration to develop a modernized property tax management system. As such, the methodology follows some incentives from the policy perspective. First, the variables collected were retained through collaborative work with the Cadaster and Valuation department of the DGID, as well as with the ICTD's Africa Property Tax Initiative. Second, it was decided to keep only characteristics visible from the outside, because of the reluctance of owners to let DGID field agents in their houses in subsequent field work. Finally, in terms of statistical methods, we prioritized transparency for the administration, and did not explore more sophisticated specifications involving for example machine learning algorithms or geographical weighted regressions used in property valuation research. Indeed, these approaches exist in exploratory papers, but not yet in the "real work" of administrations, and we preferred, at least in a first phase, to provide the DGID with methods that had already been implemented elsewhere.

standard category.[77]
1,613 observations from the baseline have an owner-declared a rental value. We restrict to properties for which there is a unique owner on the plot, or, if there are multiple owners, cases of undivided property.[78] One extreme outliers was dropped. We use expert values to correct more subtle outliers: for the subsample of properties for which we have both owner-declared and expert values (301 observations), we investigate cases where the expert value is at least five times higher than the survey value. Using i) pictures when available ; ii) average price per room in the neighborhood, we replace the rental value by the expert value when deemed appropriate. 29 observations are replaced. We further restrict the sample to properties for which a built area was measured, leading to 1,555 observations, and drop observations for which any of the other observable characteristics is missing (7 observations). Finally, we exclude outliers from the regression analysis to increase the precision of the results: we detect observations for which the rental value is above (respectively below) the 95th (resp. 5th) percentile in the location (*Commune*), and drop them if they are not above (below) the 95th (5th) percentile in terms of built area. This leads to dropping 174 "High" and 44 "Low" observations. The final sample comprises 1,469 observations.

Selecting the type of model using cross-validation
We perform cross-fold regressions to compare the performance of the following specifications: i) OLS ; ii) OLS with interactions ; iii) Lasso ; iv) Lasso with interactions ; v) Elastic net regression ; vi) Elastic net regression with interactions. In the models with interactions, we create a categorical variable to classify Sectors into high, standard or low value, based on average price per square meter in the sector. This variable is interacted with all other characteristics. We perform cross-fold regressions with 100 iterations for each model.[79] We compute out-of-sample performance statistics (R2 and RMSE), that are averaged over the 100 iterations of each specification. These performance statistics are displayed in Table A.20 Overall, the R2 values range from 54 to 56 percent, and the RMSE values range from 0.51 to 0.55. The main finding is that the different specifications do not yield important differences in their accuracy. Interactions do not lead to significant gains in model performance, but they do add complexity to communicate the formula to the administration, and for the utilization of the formula within the property tax management software. Therefore we select the elastic net regression without interactions because it is the most general specification (since it includes the properties of the LASSO).
We then carry out the same exercise on the sample of expert values to assess whether model accuracy is stronger using this sample. We find that overall there is no significant gain in model performance compared to the survey sample: R2 values range from 47 percent to 52 percent, and the RMSE from 0.53 to 0.57. We therefore choose to calibrate the formula on the survey sample.

Final steps for the Elastic net specification
Once the Elastic net specification is retained, we carry out the following steps. we run a cross-fold elastic net regression, with 100 iterations. For each explanatory variable, we compute the median of the 100 estimated coefficients. These values are our estimated coefficients, *Model 0*. Following McCluskey *et al.* (2013), we perform a final exclusion of outliers to increase model precision. We drop observations for which the residual from the model *(0)* is more than 3 standard deviations

[77] This is conceptually similar to the approach in the points based method (Fish, 2018), when coefficients indicate deviation from the standard or mean property.
[78] Indeed, built area and other property characteristics were picked up for the whole construction on the plot.
[79] In the cross-validation approach, data is randomly split into k sub-samples. In each iteration, the model is fit on k-1 groups, and the resulting parameters are used for prediction in the subsample that was left out.

away from the mean of residuals (11 observations are dropped). We then re-run the cross-fold elastic net regressions with 100 iterations, and compute the median of the different estimates for each parameter. This yields the final coefficients, *Model 1*. They are displayed in Table A.22. Coefficients that have a value are zero are the ones which were not retained in the elastic net regression.[80]
Table A.21 shows the performance indicators associated to the retained specification. The model's performance appears lower than what is found in papers exploring the issue in cities of developed countries (Davis *et al.* , 2012; McCluskey *et al.* , 2013; Guan *et al.* , 2011), but comparable to results from other African cities (Ali *et al.* , 2018). However, it is important to note that in the existing literature, cross-validation methods are not used systematically, which could lead to potential model over-fitting in the results described in these studies. Table A.23 shows a case study to illustrate the functioning of the formula: it displays first the photo of a property and its observable characteristics, and second, the computation of the predicted property value using the estimated coefficients.

A.3.2 Analysis of residuals

In Figure A.15, left panel, we plot the distribution of the residuals from the retained specification, where residuals are computed as $r = Ln(Value) - \widehat{Ln(Value)}$ with $Value$ the observed values and $\widehat{Ln(Value)}$ the fitted values. The blue line plots the normal distribution for comparison purposes. The right panel compares the quantiles of the distribution of these residuals with those of a normal distribution. Overall, we conclude that it is reasonable to assume that the residuals are normally distributed.
In Figure A.16, left panel, we plot fitted values $\widehat{Ln(Value)}$ over actual values $Ln(Value)$, for the sample of properties for which both are available. Although overall the the observations are centered around the identity line (in red), confirming the relevance of the chosen specification for the hedonic model, it does appear that on average the model tends to slightly over-estimate property values in the lower part of the distribution, and under-estimate values in the upper part of the distribution. In the right panel, we plot residuals over actual property value. The slightly increasing relationship is confirmed. This could suggest that the functional form does not perfectly fit the data, and that more sophisticated relationships could be relevant if we were willing to introduce more complexity in the model (eg, with interaction terms and non-linearities).[81]
This confirms the relevance of conducting a robustness check where residuals are randomly assigned to the fitted values, for these observations for which we use predicted property values in the progressivity analysis. It is reassuring to find that the patterns we find in the progressivity analysis are not modified when residuals are randomly assigned.

A.3.3 Computing predicted values with a correction term

To compute predicted property value $\widehat{Value}$ based on fitted values from the valuation formula, a correction term needs to be applied to $exp(\widehat{Ln(Value)})$.[82] The corrected predicted values can be

[80] Interpretation: an additional 1 percent in built area is associated with an additional 0.31 percent in rental value on average; an additional floor is associated with a 21 percent increase in rental value on average; mixed usage is associated with a 15 percent higher rental value compared to residential usage only; there is no significant difference between the absence of a fence and a metal fence, but the rental value is 5.6 percent higher when the fence is a wall; a fence in a "Bad" state is associated with a 6.4 percent lower rental value on average.

[81] We still find this gradient in the residuals when using the specifications with interaction terms with the categories of sector values, and introducing quadratic terms for numerical values.

[82] See Woolridge (2012) *Introductory Econometrics: A Modern Approach 5th edition*, Chapter 6 Section 4.

written as $\widehat{Value} = \alpha_c \cdot exp(\widehat{Ln(Value)})$ where α_c is a correction term. If it is assumed that the error term in the hedonic regression is normally distributed, it can be shown that predicted values can be computed with $\alpha_1 = exp(\frac{\hat{\sigma^2}}{2})$ and thus $\widehat{Value} = exp(\frac{\hat{\sigma^2}}{2})exp(\widehat{Ln(Value)})$ where $\hat{\sigma^2}$ is the estimator of the variance of the error term. There are two options to predict $\widehat{Value}$ without the normality assumption. In the first, $\alpha_2 = n^{-1}\sum_{i=1}^{n} exp(\hat{u}_i)$ where $\hat{u}_i$ is the OLS residual $ln(y_i) - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - ... - \hat{\beta}_k x_{ik}$. In the second, $\alpha_3 = (\sum_{i=1}^{n} \hat{m}_i^2)^{-1}(\sum_{i=1}^{n} \hat{m}_i y_i)$ where $\hat{m}_i = exp(\widehat{Ln(Value_i)})$. We compute predicted values based on all three correction methods, and Figure A.17 shows the overlaid distributions of predicted values using each correction term. On the sample used for the analysis, we obtain the following values for the correction terms: $\alpha_1 = 1.12$; $\alpha_2 = 1.27$; $\alpha_3 = 1.44$. The results displayed in the previous section allow to make the assumption that the error term is normally distributed (based on the distribution of residuals). For this reason, we apply the first correction in the main analysis.
Figure A.18 plots residuals over actual values, when fitted values are exponentiated using correction 1 (normality assumption), for the subset of observations for which we have both the predicted and the observed value.

A.4 Computation of key variables used in the analysis

This section defines the main variables used throughout the analysis and explains how they are computed.

Property value
The survey questions on property value were as follows. 1) Is part of this property rented out ? 2) If so, what are the total rents received each month for this property ? 3) If not rented out, or if rented partly, regarding the rooms which are not for rent: what do you think would be the rent you could ask for, at market prices ? Enumerators were meant to encourage respondents to think about similar properties in their area. For each reply, if no response was made after a few minutes, enumerators were to provide respondents with brackets of monthly rental values. Therefore the responses include both direct numerical answers, and answers in the bracket format. When brackets were used, to generate a numerical value, we use the mean value between the bracket bounds. We only keep observations for which we are able to obtain a value for the whole property – e.g. excluding cases where the owner declared the value of the rent for a one room, but not for the other parts of the property. This yields 1,613 observations.[83] To the 1,613 owner-declared values, we add 140 values from the real estate expert survey, for properties where the owner-declared value is missing, bringing the total number of observations to 1,753. We then add 690 predicted values based on the property valuation formula (see Section A.3), for cases where property value is neither declared by the owner nor estimated by an expert.

Paid the tax
We classify a respondent as having paid the property tax in 2018 if: 1) He or she replies yes to the question, *Did you pay the property tax for this property for year 2018 ?* 2) If he or she replies No to this question, but Yes when asked *Did you ever pay the property tax for this property?*, and *Less than a year ago* to the subsequent question, *When was the last time this happened?*. My aim here is to be conservative in estimating non-compliance.

Tax net
We classify a respondent as being in the tax net if he or she paid the property tax in 2018, and/or he or she received a tax notification for the property tax in 2018. We classify a respondent as having received a tax notification in 2018 if he or she replies Yes to the question, *Did you receive a tax notification for property tax for this property in 2018?*. Note that this question was not asked to respondents who previously affirmed that they had never received a tax notification for any type of tax.

Tax amount paid
If the respondent replied that he or she paid the tax in 2018, the enumerator asked *What was the amount of this tax payment?* Here again, if the respondent did not provide an immediate answer or seemed reluctant, the enumerator could suggest brackets of tax amounts for the respondent to select one. Hence the final variable includes both direct numerical answers, and answers as a selected bracket. When brackets were used, to generate a numerical value, we use the mean value between the bracket bounds. As a sanity check, in Appendix Figure A.5, we verify that there is a

[83] In Table A.1, we test for a possible bias in missing declared property values, by comparing expert values and predicted property values based on the valuation formula, across cases where the owner declared versus did not declare the rental value. Expert values are 21 percent higher for properties with no owner-declared value (row (1)) and predicted values 7 percent higher for properties with no owner-declared value (row(2)), but in none of the cases is the difference significant at the 5 percent level.

positive correlation between property value and tax amount paid.

Theoretical tax amount
Following the tax code, the theoretical tax liability is computed as:
$Theoretical\ tax_i = 0.05 * (AnnualPropertyValue_i - 1,500,000 * H_i)$ were H_i is a dummy equal to one if the property is the main residence of the owner (tax rate of 5 percent, abatement for owner-occupied properties).

Household income
The question was asked as *Taking into account all sources of income(s), of all household members including you, how much income does your household receive or make every month ?*. Respondents could select one answer among: *Less than XOF 35,000, 35,001-100,000, 100,001-200,000, 200,001-300,000, 300,001-400,000, 400,001-500,000, 500,001-700,000, 700,001- 1 million, 1 million - 2 million, over 2 million* . We created four groups of similar size: Income group 1: less than XOF 100,000; Income group 2: XOF 100 to 200,000; Income group 3: XOF 200 to 400,000; XOF 400,000 and above.

Total wealth and income score
The wealth and income index is computed as the average of the raw income category (out of the ten brackets listed above) and total wealth decile. To construct total wealth decile, we use property value, and to account for other properties owned, we add half of their property value to owners who declare owning another property in the region (the 0.5 factor is arbitrary, since we do not have information on the value of other properties). Total wealth is thus computed as $TW_i = PropertyValue_i + N\_Other\_Prop_i \cdot 0.5 \cdot PropertyValue_i$ where $N\_Other\_Prop_i$ is the number of other properties owned, in addition to the property being surveyed. We compute the total wealth and income score as $WI\_Score_i = \frac{1}{2} IncomeScore_i + TW\_Dec_i$ where $IncomeScore$ is a 1-10 income score based on replies to the question on total household income, and TW_Dec is the decile of total wealth.

Standardized local public service score
The raw local public service score takes values (0,4) and is computed for property i as:
$PublicServiceScore_i = Light_i + Sewage_i + Water_i + Garbage_i$ where $Light_i$ is a dummy indicating whether the respondent declares benefiting from a functional street light, $Sewage_i$ is a dummy indicating whether the respondent declares that the property is connected to a sewage system, $Water_i$ is a dummy indicating whether the respondent declares benefiting from piped water, and $Garbage_i$ is equal to one if the respondent declares benefiting from a garbage removal service and zero otherwise. We then standardize this score and use the z-score in the analysis.

(Share with) Street name
We compute a street name dummy equal to one when the enumerator entered the street name for a given property, that she or he was supposed to pick up from street signs of map-reading. For the section-level analysis, this is the share of surveyed properties within the section for which the dummy is equal to one. This variable a proxy for the quality of the cadaster and the visibility of street nomenclature in a given area. A caveat is that the distribution of enumerators across space is not random (it stems from decisions by the survey company), and that in some instances it could be that a street name exists bu that the enumerators did not write it down.

Distance to Treasury and Tax office
Distance to closest Treasury is the kilometer distance between any given property and the closest

Treasury office (where tax agents who distribute tax notifications work from). *Distance to closest Tax office* is the kilometer distance between any given property and the closest tax office (where agents in charge of assessments and discovering non-declarants work from).

Weights

The weight of property i in section j is computed as: $w_{ij} = \frac{ShEl_j \cdot N_j}{R_j}$ where $ShEl_j$ is the share of plots in the section classified as eligible by enumerators (out of all properties visited, whether surveyed or not surveyed), N_j is the total number of plots in section j, R_j is the total number of properties surveyed in section j. If there are multiple owners on the plot (only one was surveyed as per the survey protocol), the weight of property i is further divided by the number of owners.

A.5 Data from the Treasury and distribution survey

During the 2019 distribution of tax notifications, the project team set up a collaboration protocol with the Treasury to collect data on the outcome of tax notifications. Four Treasury offices are involved in the distribution of property tax notifications in the region of Dakar, one in each city.[84]. The distribution agents are full time employees of the Treasury, and conduct other activities at other times during the fiscal year. Regarding property taxation, their task is conducted in two phases. In phase one, in the office, they sort tax notifications between notifications deemed usable and those considered non-usable. In phase two, they are split into geographical groups and go in the field with the usable tax notifications, that they distribute in the relevant properties. They do not collect payments in the field, taxpayers are required to physically visit the Treasury office to pay their liability.

These agents were equipped with tablets, on which the project team had developed the distribution questionnaire, including pre-loaded information. The process was as follows: the agent enters the (visible) tax notification number, and selects whether the tax notification is usable or non-usable (based in their judgment and local knowledge). If non-usable: the agent selects the reason. This is done in the office. Questionnaires for usable tax notifications are filled in the field. The agent selects the commune and cadastral section in which the property is located and allows the tablet to record the precise geolocalization.[85] Finally, the agent indicates whether the tax notification was indeed distributed, and if not, for which reason.

Because of coordination problems with the Treasury, the survey did not start as early as expected. Also, compliance with the protocol by the Treasury agents was lower than we had hoped. As a result, there is a large share of tax notifications for which we did not collect information on their outcome (80 percent). We do not know whether most of these were distributed or not.

To assess whether there is a systematic bias explaining which tax notifications are included in the distribution survey versus the ones that were distributed before the setting up of the protocol, the main verification we can carry out with the data at hand is to assess whether the probability of being included in the survey varies with the amount of tax liability. Results are displayed in Figure A.12 (left panel – the blue bar indicates the proportion of tax notifications which are not in our dataset for every decile of tax liability from the assessment dataset) and Table A.24. There is a slightly higher probability of being included in the survey for tax notifications with higher tax liabilities (Table A.24 shows that with a doubling of the tax liability the probability to be in the outcome dataset is 2 percentage points or 11 percent higher), but the magnitude is small and the heterogeneity by decile of tax liability is limited.

Among tax notifications that were recorded in the survey, the main reasons for being considered non-usable are: double entry (60 percent of non-usable) – the entry refers to the same property and taxpayer as another entry, widespread because of the absence of unique identifiers; the owner moved out or deceased, or the property has been destroyed; the address is not precise enough.

[84] *Recette Perception Municipale de Dakar, Perception de Rufisque, Perception de Guédiawaye, Perception de Pikine*

[85] The objective of this step is to associate a precise address to the tax notifications, since the address information on the valuation roll and thus on the tax notifications is extremely incomplete.

A.6 Analysis of administrative assessment data

The analysis of the property tax valuation roll (assessment data) reveals important inconsistencies. Although the tax identification number is in theory mandatory, only 30 percent of the entries display one.[86] 84 percent of the entries are however associated to a temporary identification number.[87]

On the valuation roll, the distinction is made between properties that are being rented out and others (this is purely informative since it does not make a difference on the tax rate). We observe that 92 percent of properties in the *assessment* data are at least partly in rent. Although in our survey, this percentage is only of 30 percent, and according to 2013 national census data 54 percent of households in Dakar are tenants (Agence Nationale de la Statistique et de la Démographie, 2020). This reveals that rented properties are largely over-represented in the assessment data.

This could be because the perceived risk of evading is higher for taxpayers in these cases, or because these cases are more visible and attract more attention. It could also be because owners consider that tenants might provide rental information to the administration in the context of their own tax obligations – although this is likely only if the tenant conducts a commercial activity on the property.

We find that 26 percent of properties in the assessment data are registered for the garbage tax, but not for the property tax. This is legitimate in cases where the property is the main residence of the owner, and when property value is below the abatement threshold. But we observe in the data that this does not explain the large majority of these cases. Exemptions from the property tax are also legitimate under specific conditions, for instance, a recent construction. However, 26 percent seems to be a large percentage explained by these cases (the information on exemptions is absent from the valuation roll).

Finally, the tax notification outcome survey with the Treasury shed light on the large numbers of double entries: multiple entries that actually correspond to the same property and taxpayer. Based on the distribution survey data, it is estimated that this involves around 18 percent of entries.

[86] *NINEA, Numéro d'Identification National des Entreprises et des Associations* – initially only meant for firms, it was later extended to individuals.

[87] *Compte contribuable*

A.7 Robustness checks

A.7.1 Fiscal pressure by decile of property value: alternative computations of the property value variable

We replicate Figure 6, which plots different measures of fiscal pressure by decile of property value, using alternative computations of the property value variable, to assess whether the main conclusions hold. The four alternative computations that we test are:

1. Property value is predicted using the valuation formula, and the sample is restricted to cases for which declared value is missing ($N = 690$);
2. Property value is predicted using the valuation formula, for the whole sample ($N = 2,384$);
3. Property value is declared ($N = 1,753$). The main sample throughout the paper is composed of (1) + (3);
4. Similar to the main sample, but we add a residual to predicted property values. The residual is drawn randomly from the distribution of observed residuals. The augmented predicted value is then exponentiated using the same correction method as in the main analysis;
5. The sample is restricted to cases where the owner does not own multiple properties. This allows to avoid confusion where the respondent would declare tax paid for several properties.

Results are Figure A.9 (left panel, version (1) of the property value variable, version (2) in the right panel), A.10 (version (3) in the left panel, and (4) in the right panel), and Figure A.11. In Figure A.9, the left panel is difficult to interpret because the number of observations is much smaller. We still observe that overall the share of payers increases with decile of property value (although less regularly), that the observed average rate is relatively flat except an increase in the last decile, and payers in deciles seven to ten are subject to an effective rate lower than the theoretical one. The right panel is similar to the one in our main sample. The average rate is almost flat except for deciles nine and ten, and the effective tax rate for payers is higher than theoretical rate in decile one, close to theoretical rate for deciles two to six, and lower than the theoretical rate in deciles seven to ten. In Figure A.10, the tax profiles obtained are also quite similar to what we observe in Figure 6 and the main conclusions still hold. Excluding multiple owners in Figure A.11 also yields the same patterns. However, one interesting finding from these robustness checks is that when using predicted values only, we find lower effective rates for payers in the lowest deciles than we do in the version with declared values. This could suggest that some payers under-estimate their property value: they are counted in higher deciles in the ranking based on predicted values, than in the ranking based on declared values.

A.7.2 Simulations under revenue-maximizing targeting: accounting for cost of distribution

We replicate the simulation analysis described in Section 5, but trying to account for the administrative cost of taxing a given property, which we approximate based on the distance between the property and the nearest Treasury office. Results are show in Figure A.6. What is kept constant in this version is $\sum_i \mathbb{1}Tax\_Net_i \cdot C_i$ with $C_i = 2 \cdot dist_i \cdot uc$. $dist_i$ is the number of kilometers between the property and the closest Treasury office. uc is the unit cost of driving one kilometer that we estimate at XOF 100 (we consider a vehicle which consumes 0.13 L per kilometer, and the price of fuel in Dakar is approximately 775 per Liter). 2 is a multiplication factor (the rationale is to mimic

the journey back and forth between the property and the Treasury office). The revenue maximizing scenario is one in which the administration targets properties such as to maximize revenue $R = \sum_i Tax_i - Tax\_Net_i \cdot 2 \cdot dist_i \cdot uc$. The results in the left panel of A.6 show that the share in tax net in the tenth decile would 75 percent instead of 85 percent in the scenario ignoring costs, and the share in tax net in sixth decile is one percent instead of zero percent. Other than these slight differences the main conclusion still holds: the current situation is closer to random targeting than to strategic targeting based on property values and tax potential.

A.7.3 Income and wealth as determinants of being in the tax net

First, we replicate the property level regression on correlates of being in the tax net, using a Logit model instead of an OLS. Results are shown in A.7. They are qualitatively similar to the main results shown in Table 3. The explanatory power of the model is also within the same order of magnitude – the Youden index is 0.21 without the section fixed effects in column 1, and 0.46 with section fixed effects in column 2. The coefficient on property value indicates that odd ratio of being in tax net increases by 0.5 percent (resp., 0.6 percent within section) when property value increases by 1 percent.

Second, we replicate both the section level and the property level regressions (Tables 2 and 3), for each alternative computation of the property value variable. Results are shown in Appendix Tables A.11 to A.18.

At the section level, the four robustness results are very similar to the main results: on the extensive margin, a significant and negative effect of the distance to the Treasury, on the intensive margin an influence of average property value, and the size of the coefficient remains of similar magnitudes. The correlation with average property value is strongest when using valuation predictions for the whole sample (A.13). A difference is that when using alternative computation (1), there is a significant and positive correlation with distance to tax office (A.11).

At the property level, when using alternative computation (2) (predicted values for the whole sample) in Table A.14, property value is a slightly stronger predictor of being in the tax net – a doubling of value increases the probability by 10 percentage points overall and 12 percentage points within section. The correlation between property value and tax amount is also slightly stronger - a one percent increase in the former is associated with a 0.37 percent (0.38 with section fixed effect) in the latter, and the correlation is significant and close to one conditional on tax payment (while this was not significant when introducing section fixed effects in the main analysis). The explanatory power for tax amount conditional on paying is also higher than in the main regression (in other columns, the explanatory power is similar to what was found previously).

When using alternative computation (3) (declared values only) in Table A.16, the coefficient on property value is closer to what we observe in the main sample, through columns (1) to (4), and it is also non-significant in columns (5) and (6). The same variables come out as significant and the predictive power of the model is of the same order of magnitude as in the main analysis.

Finally using alternative computation (4) (declared values, when they are missing, predicted values augmented by a randomly drawn residual) in Table A.18, the coefficients on property value in columns (1) and (2) are smaller than in the main analysis, and the association with tax amount is also weaker (a one percent increase in value associated with 0.22 increase in tax amount, 0.16 with section fixed effects). The association is also non significant in columns (5) and (6). Other than that the same patterns are found with respect to the covariates.

A.7.4 Property tax vs Income Tax

We compare the relative influence of economic and geographic variables on the probability of ever receiving a property tax notification, and on the probability of ever receiving an income tax notification.[88] Results are in Tables A.9 and A.10. An important caveat is that for formal employees, the income tax is often withheld from their pay bill, meaning that they might not receive tax notifications.

In Table A.9 (economic covariates), we find that when controlling for section fixed effects (columns 2 and 4) income group is significant for the property tax but not for the income tax, while employment status is significant for both. The predictive power is weak in both cases (0.27 youden index for property and 0.02 for income).

The Youden index increases sharply when adding section fixed effects for the property tax (from 0.02 to 0.27) while it does not get much higher for the income tax (0 to 0.02). This is confirmed in Table A.10 : in a model with sections only, in columns 1 and 3, the predictive power is significantly higher for the property tax than for the income tax. And the geographic variable for cadastral quality $StreetName$ is positively correlated with the fact of having ever received a property tax notification.

This confirms that urban geography and cadastral developments plays a special role in explaining variations in enforcement of the property tax.

A.8 Property taxes throughout the colonial and post-colonial periods

A.8.1 Overview of tax instruments

***Impôt foncier* (Property tax)**

History. First appears in the colony of Senegal in 1856. Initially referred to as *Impôt locatif*, then *Contribution foncière sur les propriétés bâties et non bâties* (Property tax on built property and vacant land) since 1921. The latter is the current legal name of this tax instrument. In the metropole (France), the *Contribution foncière* was created shortly after the French Revolution, later renamed *taxe foncière* (Cogneau *et al.* , 2020b; Gouvernement Général de l'Afrique Occidentale Française, 1946; France. Ministère de la marine et des colonies., 1887, 1897, 1923; Piketty, 2001).

Taxation modalities. The rules regarding this tax instrument have been the same ever since its introduction: the tax base is the rental value of the property (estimated or as per rental contracts) and the tax liability is a percentage of the tax base. There is an abatement for home-ownership. The tax is due by the owner. We can trace the existence of the tax in the Dakar region as early as 1897 (according to the sources at hand, note that Dakar was not the earliest urban center in colonial Senegal). The rate has varied over time – 4 percent in 1897, 6 percent in 1929, 15 percent in 1999, 5 percent at least since 2013 (France. Ministère de la marine et des colonies., 1897; Colonie du Sénégal., 1897, 1904, 1931; Gouvernement Général de l'Afrique Occidentale Française, 1946).

Administrative level of responsibility. At first, the tax is administered at the level of the colony (*Budget local*). After revenue collection, the different territorial units receive from the colony's budget property tax revenues accruing from their area. In 1925, the Circonscription of Dakar is created and establishes its own budget (starting 1942, the Circonscription collects property tax revenues itself). However, the Circonscription of Dakar is dissolved in 1946. Since independence, the property tax is managed nationally – it is therefore only partially a 'local' tax: revenues are indeed remitted by the Treasury to local governments, however, local governments cannot modify the fiscal rules. There is no specificity for the region of Dakar, apart from the fact that there is a lack of

[88] The survey variables do not allow to construct an equivalent of the $TaxNet$ dummy for income tax, and neither do we have a variable indicating whether the respondent paid the income tax in 2018

clarity in the legal framework regarding how the revenues of each city (Dakar, Pikine, Rufisque and Guediawaye) are supposed to be shared between the city's budget and the Communes' (districts within each city) budgets (Gouvernement Général de l'Afrique Occidentale Française, 1925, 1937, 1942; Ndiaye, 2020).

***Contribution mobilière* (Contribution on movable property)**

History. First mentioned in the colony of Senegal in 1831 (from the sources at hand), also referred to as *côte mobilière*. This tax together with the *côte personnelle* constitute the *contribution personnelle-mobilière*. In the metropole, the *contribution personnelle-mobilière* was created in the aftermath of the French Revolution, and later became the *taxe d'habitation* (occupancy tax). It seems that in Senegal, this transformation did not occur : there is no occupancy tax nowadays (Cogneau *et al.* , 2020b; Gouvernement Général de l'Afrique Occidentale Française, 1931; Piketty, 2001).

Taxation modalities. The tax is due by the occupant, whether it be the owner or a tenant. The tax base is the same as for the property tax. The rates varied across the territory of the colony. First, the tax existed only in proper Communes (*communes de plein exercice*, urban areas with a certain degree of autonomy), and was only extended to the whole colony after the 1930s (Gouvernement Général de l'Afrique Occidentale Française, 1931, 1946).

Administrative level of responsibility. The tax was managed at the national level, and revenues were remitted to the corresponding territorial units – except between 1942 and 1946 when the Circonscription of Dakar was also responsible for the collection of the tax.

***Taxe d'enlèvement des ordures ménagères* (TEOM or Garbage tax)**

History. My understanding is that under colonial rule, this tax was implemented according to each municipality's decision. For example, Gouvernement Général de l'Afrique Occidentale Française (1946) mentions that the revenues of some cities such as Dakar, Rufisque and Saint Louis include collections from municipal taxes such as the garbage tax. The TEOM was introduced nationally in 1958. The current legal framework stems from a piece of legislation from 1972. Today, it is the only (public) source of funding for waste management services (Gouvernement Général de l'Afrique Occidentale Française, 1946; Colonie du Sénégal. Commune de Dakar., 1947; Ndiaye, 2020).

Taxation modalities. The tax is due by the owner, and the base is the rental value of the property. The rate is higher in the region of Dakar (3.6 percent against 3 percent) and there is no abatement for home-ownership.

Administrative level of responsibility. Nowadays, the tax is managed nationally just like the property tax. There is here again some unclarity on how revenues are supposed to be divided between Cities and Communes (Ndiaye, 2020).

***Taxe d'habitation* (Occupancy tax)**

The only mention we found was in Colonie du Sénégal. Commune de Dakar. (1947). This suggests that this must have been a municipal tax in the colonial period. It no longer exists.

A.8.2 Computation of variables for the historical analysis

Figure A.19 displays the ratio of property tax revenue, and assimilated taxes, per capita for the region of Dakar. This section describes how the variables used in the ratio are computed.

Property tax revenue:

Referred to in the sources as *Impôt locatif* in years 1904 and 1895, and *Impôt foncier* in other years.

In terms of the geographical scope, we have to make some adjustments to ensure consistency between the numerator and denominator. In year 1897, we use the value corresponding to the *2e arrondissement* of the colony, which is the region of Dakar and Rufisque (and surroundings). For years 1925, 1936, and 1937, we were not yet able to recover revenue data for Rufisque. We therefore use total revenue for Dakar, and also restrict the population to Dakar instead of Dakar and Rufisque. In years 1931 and 1942, we combine Dakar, its suburbs (*Banlieue de Dakar*), and Rufisque. The population variable is the sum of the population of both cities. Revenues of each city are found in the national budget in 1931. By 1942, both cities are integrated in the Circonscription of Dakar - revenues are found in the budget of the Circonscription. In terms of currency, the original data is in current nominal francs. We convert to 2019 West African Francs (XOF) using an exchange rate and deflator based on AOF price indexes, and Senegalese price index after independence (source: Afristory dataset from Cogneau *et al.* (2020b)).

Assimilated tax revenue:
This revenue category corresponds to *Contribution mobilière* during the colonial period, and *Taxe d'enlèvement des ordures ménagères* (garbage tax) in the post-colonial period. We use revenues corresponding to *2e arrondissement* in year 1897. As for the property tax revenue, in years 1925 we restrict to Dakar and its suburb both for the revenue and for the population variables, and in years 1931 and 1942, we use total revenues and total population for Dakar, its suburb and Rufisque.

Local tax revenue:
In Table 4 we display the ratio of property tax revenues to total local tax revenues. In 1942 we use the figure of total tax revenue in the Circonscription of Dakar. This is a fraction of total revenue of the Circonscription, since a substantial share of revenues are transfers. We do not have the equivalent figure for 1931, so we extrapolate by applying to total revenues of Dakar and its suburbs the ratio of tax revenue over total revenue observed in 1942. In 2018, total local tax revenues are from the Ministry of Finance.

Total tax revenue:
For the colonial period, this refers to total tax revenues collected in the colony of Senegal. The source is *Total recettes ordinaires* in Colonie du Sénégal. (1897) for 1897, total tax in local budget from Cogneau *et al.* (2020b) for 1931 and 1942 (we use the figure for 1943), IMF (2019) for 2018.

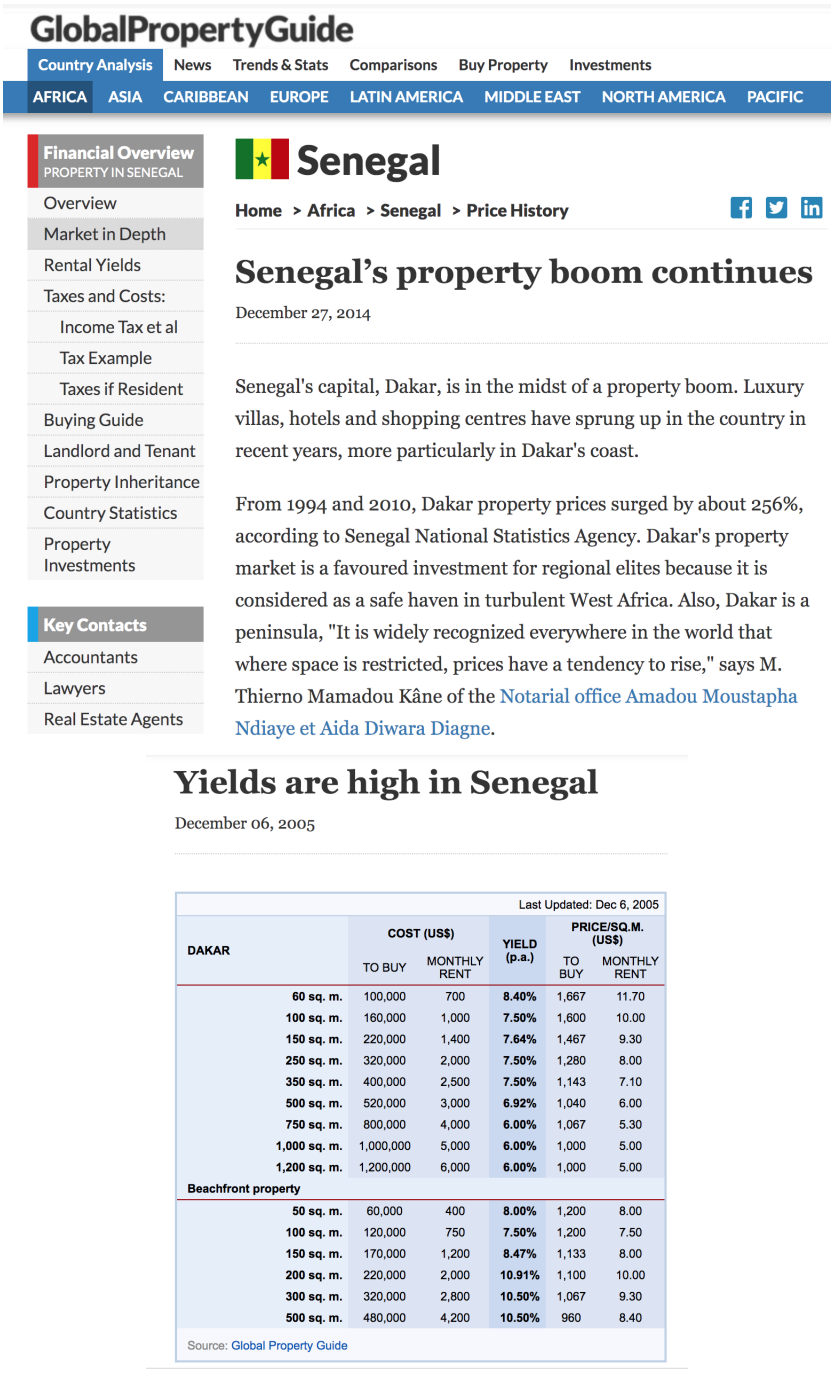
Population:
Colonial Dakar is characterized by a massive population increase throughout the first half of the XXth century, see Figure A.20 (Becker & Martin, 1981; Bouche, 1978; Seck, 1961). For Dakar and its suburb, and Rufisque, the data comes from the Africapolis dataset (Sahel and West Africa Club, 2018), which provides population by decade, starting 1920. Since the revenue data points are not on round number decade years, we interpolate assuming a constant growth rate between year T and yeat $T+10$. This means, for example, $pop_{1925} = pop_{1920} + 5 * \frac{pop_{1930} - pop_{1920}}{10}$. For year 1897, we interpolate between 1882 and 1904, assuming a constant growth rate, and the figures for these years are from Becker & Martin (1981). For recent years (2015 to 2019) the data comes from the national statistical agency (ANSD online resources).

Appendix Figures

FIGURE A.1
DAKAR REAL ESTATE BOOM IN THE PRESS

Notes: Articles from the press. Left: DakarActu (senegalese media outlet), the author is Pape Makhtar Diop, the CEO of Jiwall, a real estate investment platform, March 2020. Right: RFI (international radio and online media), February 2020.

FIGURE A.2
DAKAR REAL ESTATE BOOM IN A GLOBAL PROPERTY INVESTMENT GUIDE



GlobalPropertyGuide

Country Analysis | News | Trends & Stats | Comparisons | Buy Property | Investments

AFRICA | ASIA | CARIBBEAN | EUROPE | LATIN AMERICA | MIDDLE EAST | NORTH AMERICA | PACIFIC

Financial Overview
PROPERTY IN SENEGAL

- Overview
- Market in Depth
- Rental Yields
- Taxes and Costs:
  - Income Tax et al
  - Tax Example
  - Taxes if Resident
- Buying Guide
- Landlord and Tenant
- Property Inheritance
- Country Statistics
- Property Investments

Key Contacts

- Accountants
- Lawyers
- Real Estate Agents

Senegal

Home > Africa > Senegal > Price History

Senegal's property boom continues

December 27, 2014

Senegal's capital, Dakar, is in the midst of a property boom. Luxury villas, hotels and shopping centres have sprung up in the country in recent years, more particularly in Dakar's coast.

From 1994 and 2010, Dakar property prices surged by about 256%, according to Senegal National Statistics Agency. Dakar's property market is a favoured investment for regional elites because it is considered as a safe haven in turbulent West Africa. Also, Dakar is a peninsula, "It is widely recognized everywhere in the world that where space is restricted, prices have a tendency to rise," says M. Thierno Mamadou Kâne of the Notarial office Amadou Moustapha Ndiaye et Aida Diwara Diagne.

Yields are high in Senegal

December 06, 2005

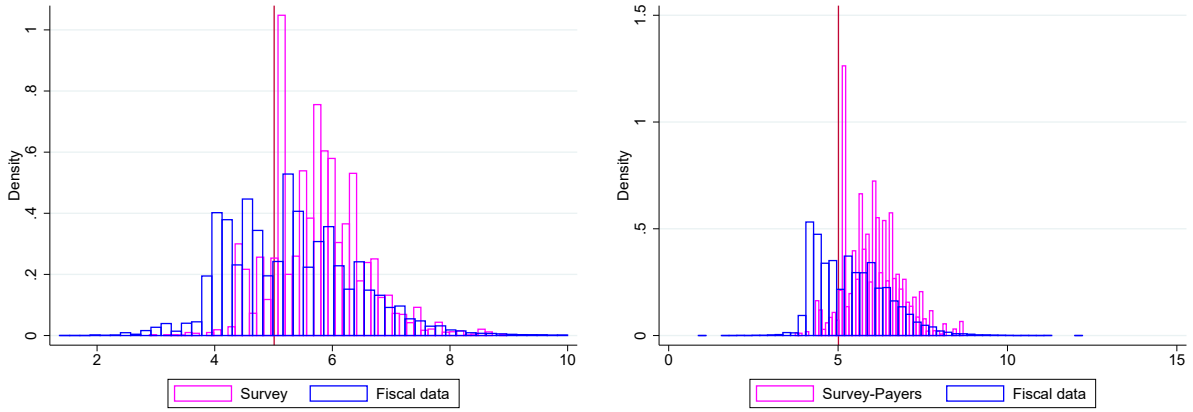
Last Updated: Dec 6, 2005

DAKAR	COST (US$) TO BUY	COST (US$) MONTHLY RENT	YIELD (p.a.)	PRICE/SQ.M. (US$) TO BUY	PRICE/SQ.M. (US$) MONTHLY RENT
60 sq. m.	100,000	700	**8.40%**	1,667	11.70
100 sq. m.	160,000	1,000	**7.50%**	1,600	10.00
150 sq. m.	220,000	1,400	**7.64%**	1,467	9.30
250 sq. m.	320,000	2,000	**7.50%**	1,280	8.00
350 sq. m.	400,000	2,500	**7.50%**	1,143	7.10
500 sq. m.	520,000	3,000	**6.92%**	1,040	6.00
750 sq. m.	800,000	4,000	**6.00%**	1,067	5.30
1,000 sq. m.	1,000,000	5,000	**6.00%**	1,000	5.00
1,200 sq. m.	1,200,000	6,000	**6.00%**	1,000	5.00
Beachfront property					
50 sq. m.	60,000	400	**8.00%**	1,200	8.00
100 sq. m.	120,000	750	**7.50%**	1,200	7.50
150 sq. m.	170,000	1,200	**8.47%**	1,133	8.00
200 sq. m.	220,000	2,000	**10.91%**	1,100	10.00
300 sq. m.	320,000	2,800	**10.50%**	1,067	9.30
500 sq. m.	480,000	4,200	**10.50%**	960	8.40

Source: Global Property Guide

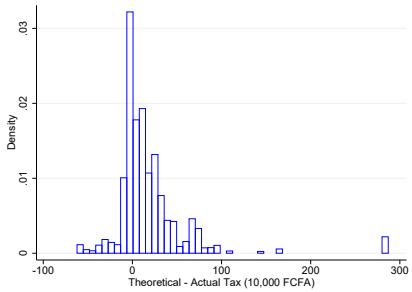
Notes: Articles from the property investment advisory platform, Global Property Guide.

FIGURE A.3
DISTRIBUTION OF PROPERTY VALUES IN ADMINISTRATIVE VS SURVEY DATA



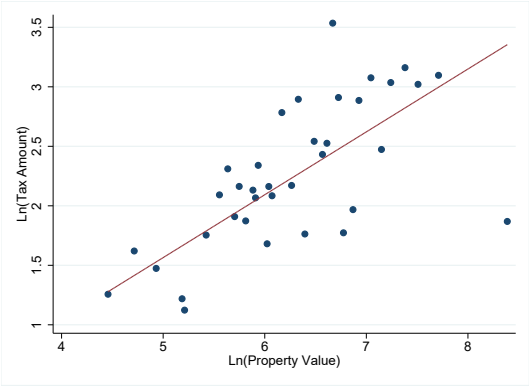
Notes: These histograms show the distribution of property values across two sources: administrative data from the property tax valuation roll (blue), and property owner survey responses (pink). The x-axis corresponds to $Ln(annual\ rental\ value)$, and the vertical red line indicates $annual\ rental\ value = XOF\ 1,500,000$, the abatement for owner-occupied properties. The left panel includes the universe of owners in the fiscal data ($N = 53,878$), and all survey respondents ($N = 2,443$, observations are weighted). The share below the abatement threshold is 15.5 percent in the survey, 41 percent in the fiscal data. The right histogram isrestricted to owners registered for the property tax in the fiscal data ($N = 39,934$), and to survey respondents in the tax net in 2018 ($N = 441$). The share below the abatement threshold is 7.8 percent in the survey, 42 percent in the fiscal data. In the survey data, property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, DGID property tax assessment data, 2018.

FIGURE A.4
DIFFERENCE BETWEEN THEORETICAL AND ACTUAL TAX PAYMENT



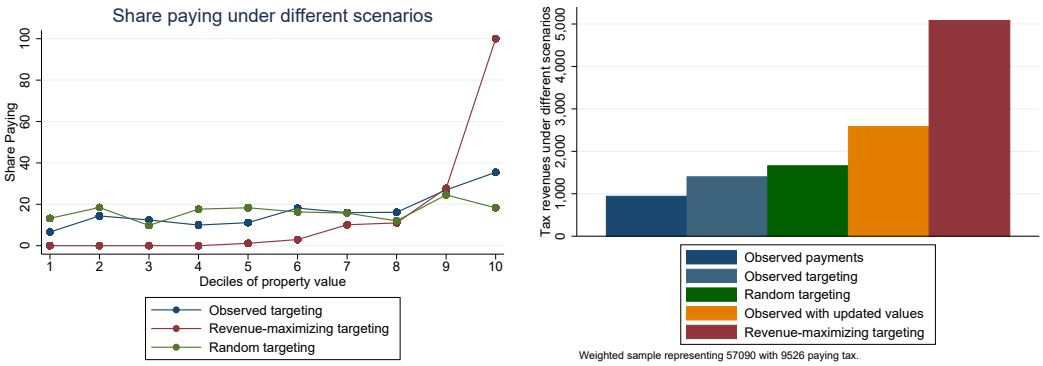
Notes: This figure shows the distribution of the difference between theoretical tax liability and tax amount paid, conditional on paying. The sample is restricted to respondents who declare paying and who provide a tax amount ($N = 193$). Tax amount paid is self-declared. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. Amounts are in XOF 10,000. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

FIGURE A.5
CORRELATION BETWEEN PROPERTY VALUE AND TAX AMOUNT IN SURVEY



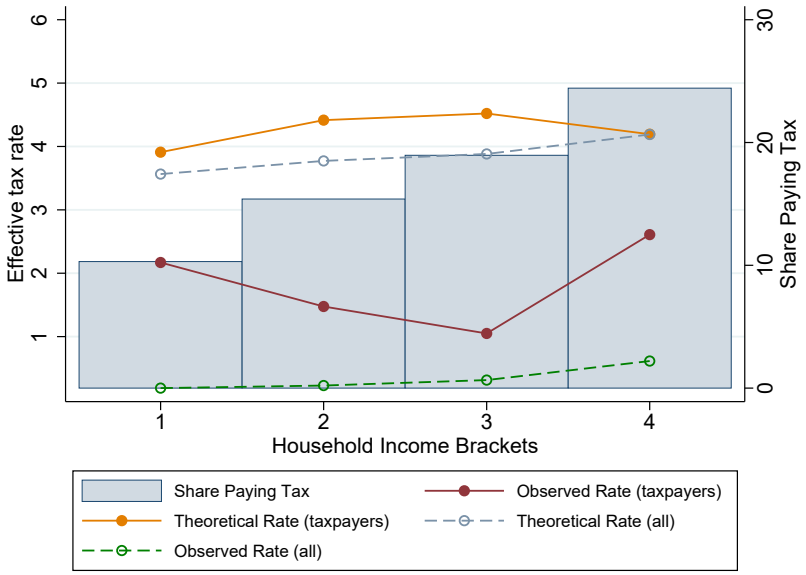
Notes: This graph plots amount paid over property value, both expressed in logarithm. Observations are averaged by bin of property value. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. Tax amounts are declared. For details on the data and computations, see Appendix A.4. When running the regression of $Ln(TaxAmount)$ on $Ln(PropertyValue)$, I find a positive and significant coefficient of 0.53, and a R2 of 19.6. Source: property owner survey 2018, observations are weighted.

FIGURE A.6
SIMULATING A REVENUE-MAXIMIZING TARGETING ACCOUNTING FOR COST OF DISTRIBUTION



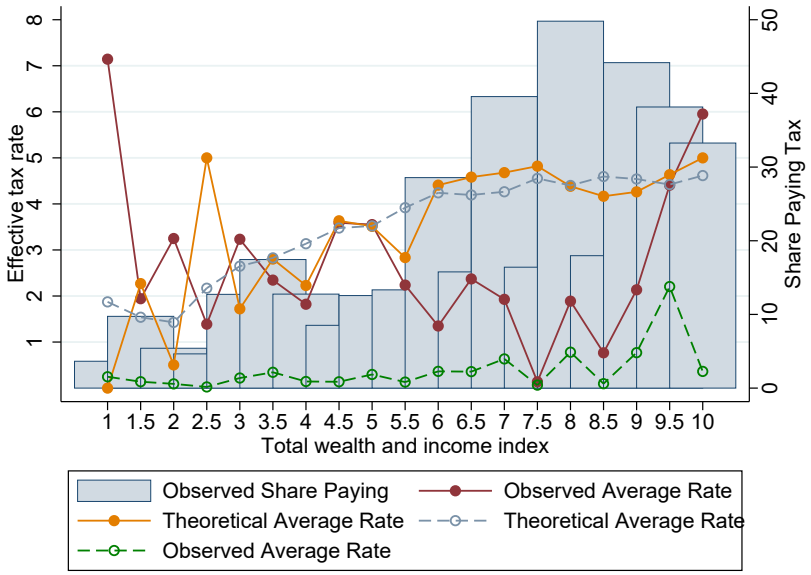
Notes: These figures compare different scenarios of targeting by the tax administration on the property owner survey sample, keeping constant the total administrative cost of distributing tax notifications. The cost of distributing a given tax notification is approximated by $C_i = 2 \cdot dist_i \cdot uc$. $dist_i$ is the number of kilometers between the property and the closest Treasury office. uc is the unit cost of driving one kilometer that we estimate at XOF 100 (we consider a vehicle which consumes 0.13 L per kilometer, and the price of fuel in Dakar is approx. 775 per Liter). 2 is a multiplication factor mimicking the journey back and forth between the property and Treasury. The left panel plots the share of payers by decile of property value under each scenario. The right panel shows total revenues under the different scenarios in XOF millions. *Observed payments* shows total tax revenues as declared in the survey. *Observed targeting* is the targeting scenario measured in the survey, including all respondents who declare paying the tax in 2018. For observations where tax amount paid is missing, it is imputed using theoretical tax. In the *Random targeting* scenario, the payer status is randomly assigned. In the scenario labeled *Observed with updated values* the payers are those counted in the survey, but tax amounts are switched to theoretical tax liabilities. Under the *Revenue-maximizing targeting* scenario the payer status is assigned such as to maximize the overall difference between tax revenues and administrative costs. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. Tax revenues under the observed scenario are computed using survey responses. Theoretical tax revenues under the other scenarios are computed using property value and fiscal modalities of the property tax. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

FIGURE A.7
FISCAL PRESSURE BY QUARTILE OF HOUSEHOLD INCOME



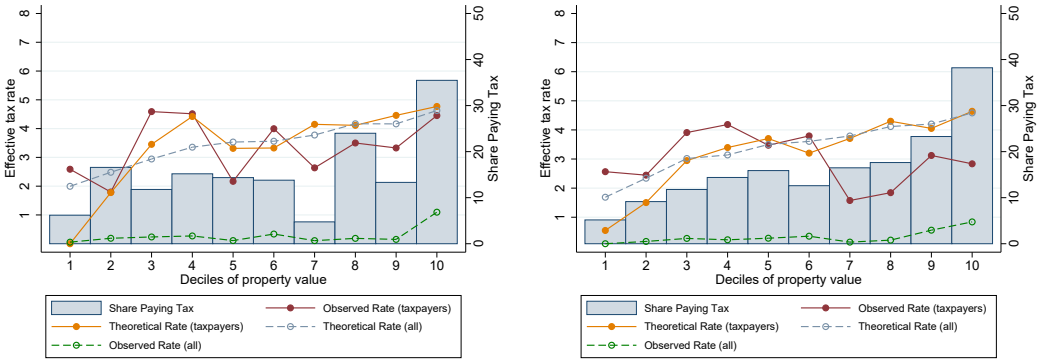
Notes: This graph shows different measures of fiscal pressure for each group of household income. The grey bars indicate the share of owners in each decile that paid the tax in 2018 (right-hand y-axis). The green dashed line represents the average effective tax rate observed in each quartile - computed as the ratio of total tax amounts over total property value. The grey dashed line shows the theoretical tax rate, if all owners paid their liability in full. The orange and red lines show respectively the observed effective tax rate, and the theoretical rate, restricting the sample to owners who do pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. Monthly household income is reported in the survey using brackets. I combine some of the categories to generate four groups of similar size. Income group 1: less than XOF 100,000; Income group 2: XOF 100 to 200,000; Income group 3: XOF 200 to 400,000; Income group 4: XOF 400,000 and above. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

FIGURE A.8
FISCAL PRESSURE BY TOTAL WEALTH AND INCOME SCORE



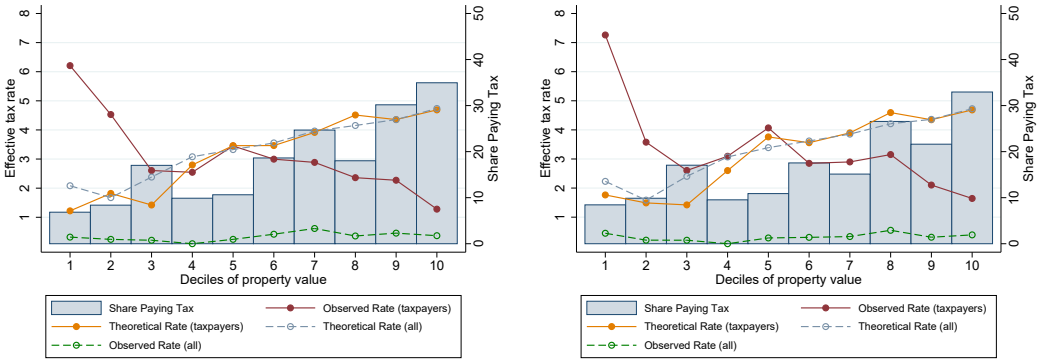
Notes: This graph shows different measures of fiscal pressure for each level of total wealth and income score. The grey bars indicate the share of owners for each score that paid the tax in 2018 (right-hand y-axis). The green dashed line represents the average effective tax rate observed in each quartile - computed as the ratio of total tax amounts over total property value. The grey dashed line shows the theoretical tax rate, if all owners paid their liability in full. The orange and red lines show respectively the observed effective tax rate, and the theoretical rate, restricting the sample to owners who do pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. The wealth and income score is computed as the average of the raw income score and total wealth decile. The latter is computed based on property value and other properties owned. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, weighted observations.

FIGURE A.9
FISCAL PRESSURE BY DECILE OF PROPERTY VALUE - ROBUSTNESS 1/3



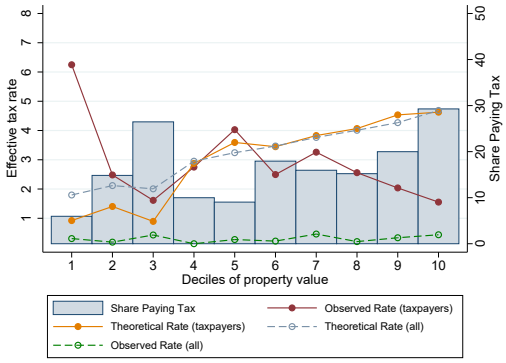
Notes: These graphs show different measures of fiscal pressure for each decile of property value. They are robustness check complements to Figure 6. Here, I use alternative measures of property value. In the left panel, property value is predicted using the valuation formula, the sample is restricted to owners for which declared property value is missing. In the right panel, property value is predicted using the valuation formula, for the full sample. The grey bars indicate the share of owners in each decile that paid the tax in 2018 (right-hand y-axis). The green dashed line represents the average effective tax rate observed in each quartile - computed as the ratio of total tax amounts over total property value. The grey dashed line shows the theoretical tax rate, if all owners paid their liability in full. The orange and red lines show respectively the observed effective tax rate, and the theoretical rate, restricting the sample to owners who do pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, weighted observations.

FIGURE A.10
FISCAL PRESSURE BY DECILE OF PROPERTY VALUE - ROBUSTNESS 2/3



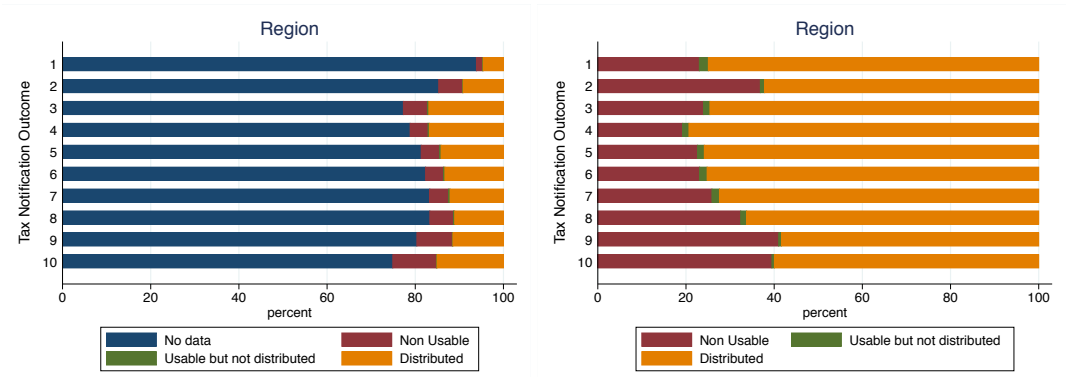
Notes: These graphs show different measures of fiscal pressure for each decile of property value. They are robustness check complements to Figure 6. Here, I use alternative measures of property value. In the left panel, property value is declared. I exclude observations for which declared value is missing. In the right panel, property value is declared, and if missing, it is replaced by the value predicted with the valuation formula augmented by a residual that is randomly drawn from the distribution of the observed residuals. The grey bars indicate the share of owners in each decile that paid the tax in 2018 (right-hand y-axis). The green dashed line represents the average effective tax rate observed in each quartile - computed as the ratio of total tax amounts over total property value. The grey dashed line shows the theoretical tax rate, if all owners paid their liability in full. The orange and red lines show respectively the observed effective tax rate, and the theoretical rate, restricting the sample to owners who do pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, weighted observations.

FIGURE A.11
FISCAL PRESSURE BY DECILE OF PROPERTY VALUE - ROBUSTNESS 3/3



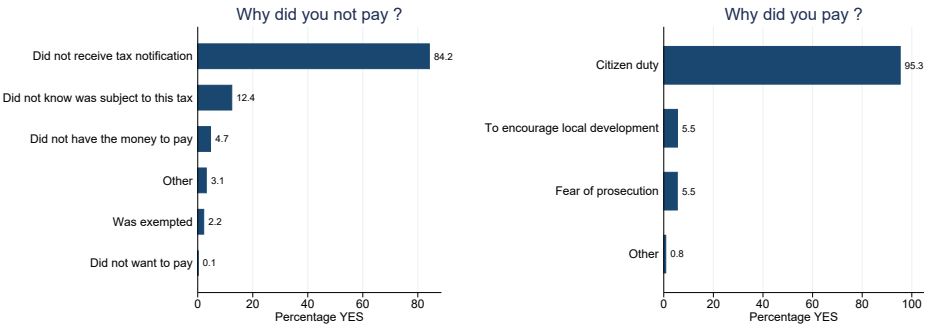
Notes: This graph show different measures of fiscal pressure for each decile of property value. It is a robustness check complement to Figure 6. Here, I restrict the sample to respondents who are not multiple owners. The grey bars indicate the share of owners in each decile that paid the tax in 2018 (right-hand y-axis). The green dashed line represents the average effective tax rate observed in each quartile - computed as the ratio of total tax amounts over total property value. The grey dashed line shows the theoretical tax rate, if all owners paid their liability in full. The orange and red lines show respectively the observed effective tax rate, and the theoretical rate, restricting the sample to owners who do pay the tax in 2018. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Tax amounts are declared in the survey. Theoretical tax is calculated using property value and fiscal modalities of the property tax. Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations, and imputed using the property valuation formula if missing. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

FIGURE A.12
TAX NOTIFICATION OUTCOME BY DECILE OF TAX LIABILITY



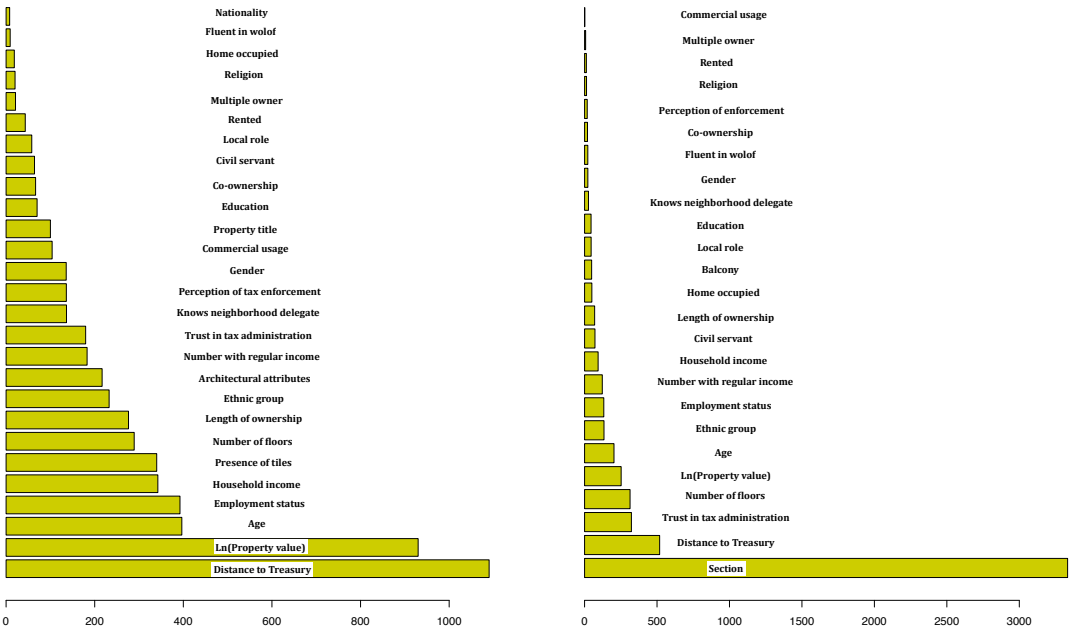
Notes: These graphs show tax notification outcomes by decile of tax liability for the region of Dakar in 2019. The left panel includes the whole universe of tax assessments ($N = 56,147$). The right panel is restricted to tax notifications for which an outcome was obtained through the survey administered to Treasury agents ($N = 10,208$). Outcomes are defined as: *No data*, no information on this tax notification in the distribution survey, *Non-Usable*, tax notification classified as non usable by Treasury agent, *Usable but not distributed*, tax notification not distributed for other reasons, *Distributed* Treasury agent declares having distributed the tax notification. Source: Tax notification outcome survey 2019-2020, DGID property tax assessment data, 2019.

FIGURE A.13
SURVEY RESPONSES FOR REASONS OF (NOT) PAYING PROPERTY TAX



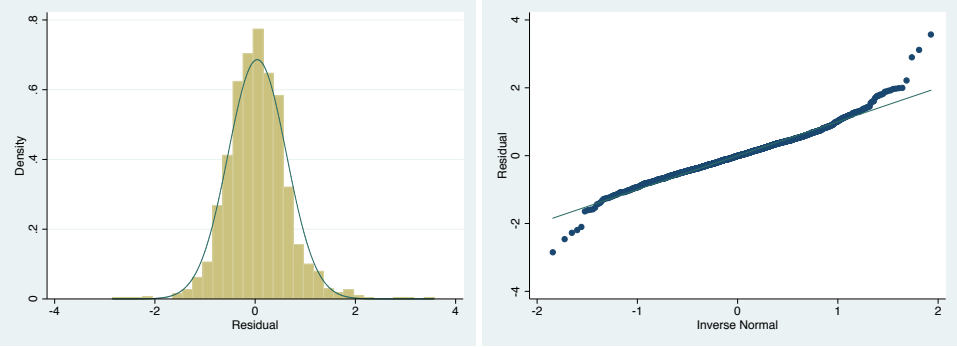
Notes: These graphs shows the percentage of respondents selecting a specific reason for not paying (left panel, $N = 2,106$) or paying (right panel, $N = 368$) the property tax in 2018. Respondents could select multiple reasons. Source: property owner survey 2018, observations are weighted.

FIGURE A.14
DETERMINANTS OF BEING IN THE TAX NET - VARIABLE IMPORTANCE RANKING FROM CLASSIFICATION TREE



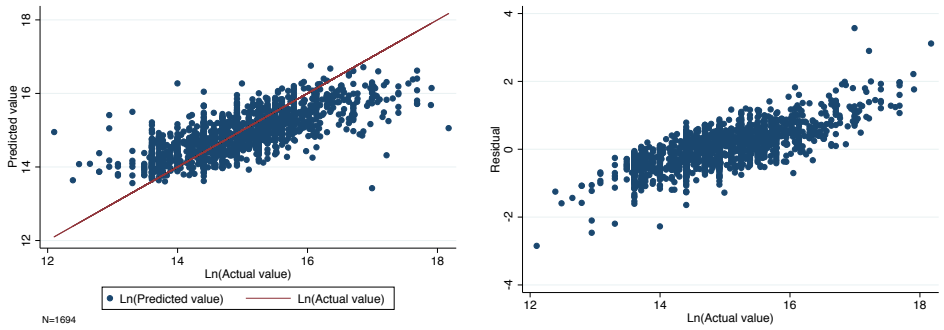
Notes: These figures show the variable importance ranking obtained from a classification tree model, where the dependent variable is a dummy indicating whether a property owner is in the tax net in 2018. The threshold complexity parameter of the tree is 0.001. In the left panel, the CART model excludes section fixed effects. In the right panel, the CART model includes section fixed effect. Variable importance of variable var is the mean decrease in node impurity from splits based on var. Node impurity is the extent to which a subset is far from including only one type of outcome value ($TaxNet = 1$ or $TaxNet = 0$). For details on the data and computations, see Appendix A.4, and see Table A.4 for full descriptive statistics. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

FIGURE A.15
RESIDUALS FROM PROPERTY VALUATION FORMULA



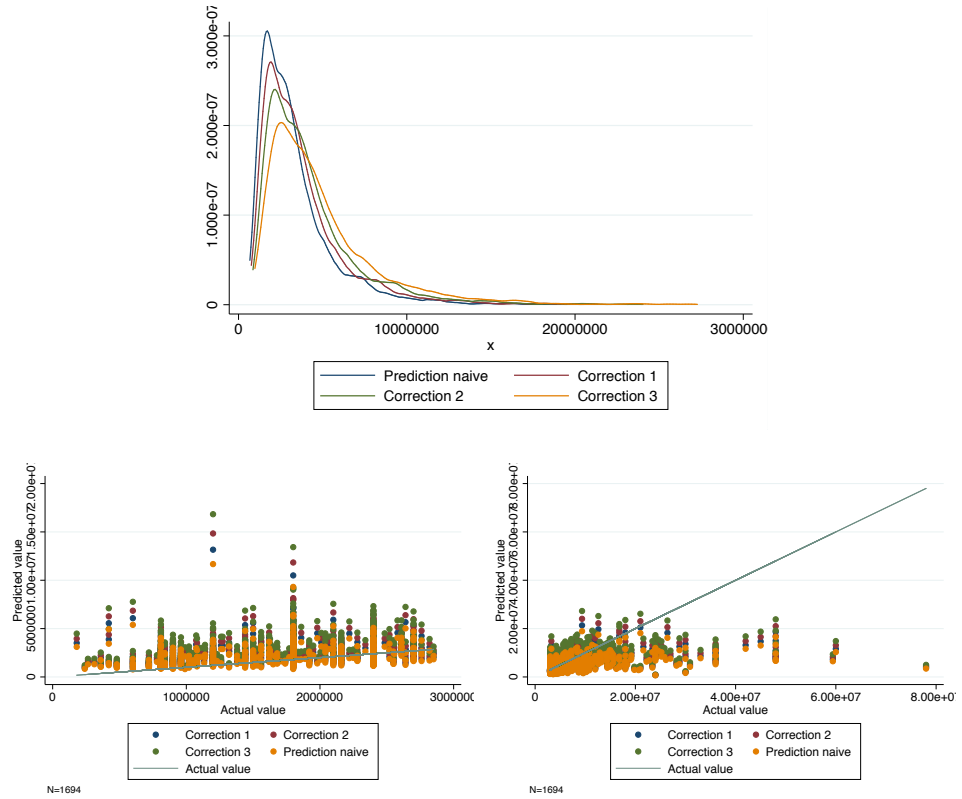
Notes: These figures compare the distribution of the residuals from the property valuation formula to a normal distribution. Residuals are computed as $r_i = Ln(Value_i) - \widehat{Ln(Value_i)}$ with $Value_i$ the observed values and $\widehat{Ln(Value_i)}$ the fitted values The left panel plots the histogram of the residuals, and the blue line is the normal distribution. The right panel plots the quantiles of the distribution of the residuals against those of a normal distribution.

FIGURE A.16
FITTED VALUES VERSUS ACTUAL VALUES IN PROPERTY VALUATION FORMULA



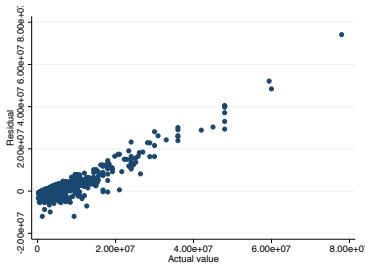
Notes: These figures display the fitted values and residuals of the property valuation formula. In the left panel, fitted values $\widehat{Ln(Value_i)}$ are plotted over actual values $Ln(Value_i)$. In the right panel, residuals are plotted over actual values. Residuals are computed as $r_i = Ln(Value_i) - \widehat{Ln(Value_i)}$ with $Value_i$ the observed values and $\widehat{Ln(Value_i)}$ the fitted values.

FIGURE A.17
CORRECTED FITTED VALUES



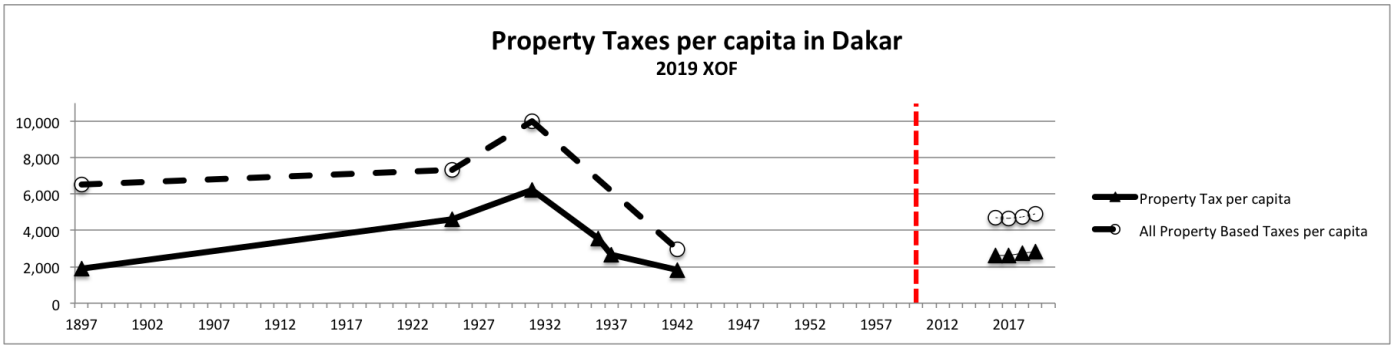
Notes: These Figures compare the distribution of predicted values from the property valuation formula, when using different correction procedures to exponentiate the fitted value. All values are in XOF. The upper panel plots the kernel density estimate of the distributions. The bottom panels plot predicted values over actual values, for values below median (left) and values above median (right). $\widehat{Value} = \alpha_c \cdot exp(\widehat{Ln(Value_i)})$ Prediction naive: $\alpha_c = 1$. Correction 1: $\alpha_1 = exp(\frac{\hat{\sigma^2}}{2})$. Correction 2: $\alpha_2 = n^{-1} \sum_{i=1}^{n} exp(\hat{u_i})$. Correction 3: $\alpha_3 = (\sum_{i=1}^{n} \hat{m_i^2})^{-1} (\sum_{i=1}^{n} \hat{m}_i y_i)$.

FIGURE A.18
RESIDUALS OVER ACTUAL VALUES WHEN APPLYING CORRECTION TO FITTED VALUE



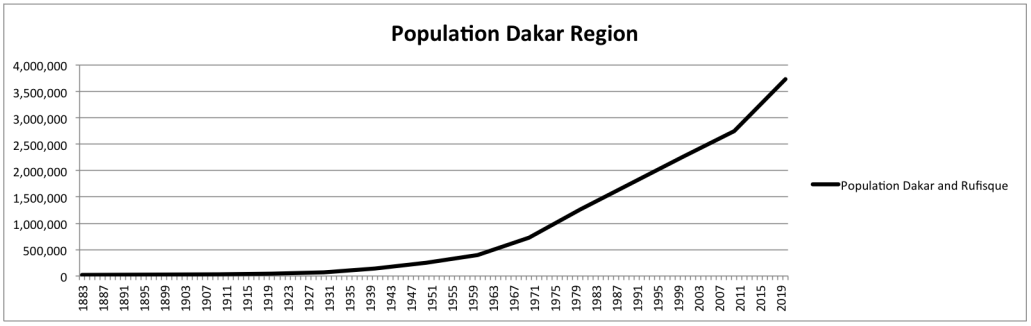
Notes: This figure plots residuals from the property valuation formula over actual values, for the subset of observations for which both are available. Fitted values are exponentiated using the following correction term: $\widehat{Value} = exp(\frac{\hat{\sigma^2}}{2})exp(\widehat{Ln(Value_i)})$. Residuals are computed as: $r_i = Value_i - \widehat{Value}$

FIGURE A.19
PROPERTY TAX REVENUE PER CAPITA IN DAKAR THROUGHOUT COLONIAL AND POST-COLONIAL TIMES



Notes: This figure shows the ratio of property and assimilated tax revenues, converted to 2019 XOF francs, over total population in the region of Dakar, between 1897 and 2019. The red line marks the year of independence. Assimilated taxes correspond to *Contribution mobilière* in the colonial period, and *Taxe d'Enlèvement des Ordures Ménagères* in the post-colonial period. The precise geographical delimitations of the agglomeration vary across time. Priority was given to making the geographical area for the numerator and the denominator of the ratio coherent. For details on the data and computations, see Appendix A.8. Source: Colonial budgets for Senegal and Dakar; Afristory data series; Africapolis data series.

FIGURE A.20
POPULATION IN DAKAR THROUGHOUT COLONIAL AND POST-COLONIAL TIMES



Notes: This figure shows population growth in the region of Dakar between 1883 and 2020. The region of Dakar includes the cities of Dakar, Rufisque, Pikine and Guédiawaye. Source: Afristory data series; Africapolis data series.

Appendix Tables

TABLE A.1
BALANCE CHECK FOR MISSING RESPONSE TO PROPERTY VALUE QUESTION

Declared rental value	Missing		Non-Missing		P-stat
	Mean	N	Mean	N	
Expert valuation	603,372	140	494,337	301	0.22
Formula prediction	340,889	828	318,833	1,556	0.06

Notes: This table aims to verify that there is no systematic difference in property values between cases where the respondent declared a value in the survey, and cases where the response to this question is missing. Property value (more precisely, annual rental value) is computed based on a series of survey questions. We have an external expert valuation for 441 properties, and we are able to compute property value using the property valuation formula for 2,384 properties. In the first row of the Table, we compare expert values across properties with missing vs non-missing survey responses. In the second row, we compare formula predictions across properties with missing vs non-missing survey responses. Mean values are monthly rental values in XOF. The third column shows the P-value for a significance test of the coefficient on a dummy for missing survey response, in a regression where the dependent variable is expert property value (respectively, formula predicted value). For details on the data and computations see Appendix A.2 and A.4. Source: property owner survey 2018, observations are weighted.

TABLE A.2
TAX ENFORCEMENT

	Percent
Ever subject to enforcement action	2.5
- conditional on ever receiving tax notification	5.0
In your city ...	
...None or few people pay the property tax	13.3
...Most or all people pay the property tax	9.8
...Does not know	76.9
If an owner doesn't pay the property tax, enforcement actions ...	
...Are very likely	25.9
...Are likely	23.3
...Are unlikely	21.2
...Does not know	29.6

Notes: This table displays descriptive statistics from the property owner survey, on questions relating to property tax enforcement. For details on the survey sampling and protocol, see Appendix A.2. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

TABLE A.3
CHARACTERISTICS OF THE REAL ESTATE MARKET

	Percent
Rented	30.3
Among which: fully rented	10.1
Share rented (value)	15.3
Previous connection Tenant-Owner	6.6
Time since rent last adjusted	
- past 3 years	49.8
- 3 to 10 years ago	37.4
- more than 10 years ago	12.9
Involvement of real estate agency	8.3
Commercial	10.9
Co-ownership	19.4
Owns multiple properties	10.5
Length of ownership: below 5 years	8.8
Length of ownership: 5 to 10 years	10.1
Length of ownership: over 10 years	78.9
Property title type: None	3.3
Property title type: Weak	52.8
Property title type: Strong	35.4
Acquisition: Donation	2.2
Acquisition: Inherited	19.9
Acquisition: Purchase	65.9

Notes: This table displays descriptive statistics from the property owner survey, on questions relating to the rental market, real estate ownership and transactions. Share rented is the share of total property value that is rented. Previous connection Tenant-Owner is the percent, among properties with a tenant, for which the owner declares that the tenant is either a family member or previous friend. For details on the survey sampling and protocol, see Appendix A.2. For details on the data and computations, see Appendix A.4. Source: property owner survey 2018, observations are weighted.

TABLE A.4
DESCRIPTIVE STATISTICS: PROPERTY, INDIVIDUAL AND TAX CHARACTERISTICS

N	677
N weighted	14,962
Variable	Mean
In tax net	0.22
Visible Property Characteristics	
Ln(Property value)	5.73
Rented	0.33
Co-ownership	0.20
Floors	0.79
Commercial	0.16
Local Public Service Index = 1	0.08
Local Public Service Index = 2	0.34
Local Public Service Index = 3	0.58
Balcony	0.62
Tiles	0.30
Architectural attributes	0.27
Personal Property Characteristics	
Main residence	0.68
Owns multiple properties	0.12
Length of ownership: below 5 years	0.08
Length of ownership: 5 to 10 years	0.13
Length of ownership: over 10 years	0.80
Property title type: None	0.03
Property title type: Weak	0.55
Property title type: Strong	0.41
Acquisition: Donation	0.03
Acquisition: Inherited	0.24
Acquisition: Purchase	0.73
Individual Characteristics	
Male	0.60
Age	56.87
Employment status: Formal	0.28
Employment status: Informal	0.05
Employment status: Non Employed	0.48
Employment status: Retired	0.18
Household revenue: $<$ 100,000	0.22
Household revenue: $<$ 300,000	0.49
Household revenue: $<$ 400,000	0.10
Household revenue: $>$ 400,000	0.20
Nationality: Senegalese	0.99
Ethnic group: Wolof	0.52
Ethnic group: Poular	0.20
Ethnic group: Other	0.27
Local role: Association	0.12
Local role: None	0.72
Local role: Political	0.03
Local role: Religious	0.12
Knows neighborhood delegate	0.67
Fluent in Wolof	0.98
Religion: Islam	0.92
Education: no primary completion	0.22
Education: no lower secondary completion	0.19
Education: no upper secondary completion	0.18
Education: upper secondary completion	0.12
Education: higher education	0.28
Individual Tax Characteristics	
Ever received any tax notification	0.49
Ever received property tax notification	0.41
Ever visited by tax agent	0.18
Trust in tax administration: No opinion	0.25
Trust in tax administration: Yes	0.38
Trust in tax administration: No	0.37
Enforcement: Very likely	0.24
Enforcement: Rather likely	0.55
Enforcement: Rather unlikely	0.21
Knows property tax	0.60

See notes on the next page.

Notes This table displays descriptive statistics from the property owner survey. For details on the survey sampling and protocole, see Appendix A.2. This table is restricted to the subset of observations included in the classification tree analysis of the determinants of being in the tax net, for which all covariates were non-missing. *In Tax Net* is defined as having paid the tax and/or having received a tax notification, for the property tax in 2018. *Property value* is the value at which the property could be rented at market prices and is defined whether or not the property is actually rented. It is computed based on a series of survey questions, with some corrections applied using external valuations. Apart from *property value*, *floors*, and *age*, all variables are categorical. *Local public service index* is a score between 1 and 3 with an additional point for each of the following features of the property: connected to electricity, connected to running water, connected to a sewage system. *Property title* is classified as strong if it is a formal full ownership title (*Titre foncier*), weak if it corresponds to long-term but not permanent tenancy titles (*Bail, permis d'occuper*). *Enforcement* refers to whether the respondent thinks the tax administration would take measures against a non-compliant property owner. For details on the data and computations see Appendix A.4. Source: property owner survey 2018, observations are weighted.

TABLE A.5
SURVEY SAMPLE COMPARED TO CENSUS

Variable	Survey	Census
Owner-occupied	0.64	0.41
Including: shared ownership	0.19	0.04
Age	55.60	54.39
	(15.49)	(13.70)
Restricting census to owner occupied homes and household head and survey to respondents who are the owner		
Age	61.26	54.39
	(12.87)	(13.70)
Age Oldest		58.66
		(13.85)
Employment: Employed	0.32	0.51
Employment: Non Employed	0.41	0.27
Employment: Retired	0.27	0.22
Education: less than primary	0.27	0.48
Education: primary	0.34	0.33
Education: secondary	0.11	0.13
Education: higher education	0.27	0.07
Water: Piped in dwelling	0.76	0.70
Water: Piped outside dwelling	0.22	0.19
Water: No	0.02	0.11
Trash collection	0.96	0.87
Sewage:	0.78	0.87
Electricity:	0.98	0.95
N. Rooms:	7.61	4.76
	(4.12)	(2.51)

Notes This Table compares descriptive statistics from our survey with statistics from the 2013 national census. I restrict the census data to the region of Dakar, and in rows three and following, to respondents who are property owners (own the property they live in). I restrict the survey data to owners who live in the surveyed property. For all variables, I display the mean value, and the standard deviation in brackets. *Age* is the age of the respondent, and *age oldest* is the age of the oldest household member. I perform a Chi-Square test of the difference in means across both samples, but results are not displayed here: the difference across the two samples is always significant at the 5 percent level. Source: property owner survey 2018, observations are weighted. 2013 Census, Senegal National Statistical Agency (accessed via IPUMS).

TABLE A.6
GEOGRAPHIC DETERMINANTS OF BEING IN THE TAX NET

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable (0,1)	TaxNet	TaxNet	TaxNet	TaxNet	TaxNet	TaxNet
Street name		-0.038	-0.020		-0.048	-0.027
		(0.041)	(0.038)		(0.040)	(0.036)
Distance to closest Treasury		0.005	-0.158		-0.008	-0.153
		(0.016)	(0.154)		(0.014)	(0.154)
Distance to closest Tax center		0.015	0.105		0.023	0.109
		(0.021)	(0.159)		(0.021)	(0.156)
Ln(PropertyValue)				0.089***	0.112***	0.089***
				(0.022)	(0.022)	(0.022)
Section FE	Yes	No	Yes	Yes	No	Yes
N	1108	1108	1108	1108	1108	1108
Adj R2	0.26	0.00	0.26	0.28	0.05	0.28
Mean of dep.	0.20	0.20	0.20	0.20	0.20	0.20
0.5 Cutoff						
% Correct	0.86	0.83	0.86	0.86	0.83	0.86
Sensitivity	0.25	0.00	0.29	0.31	0.01	0.32
Specificity	0.98	1.00	0.98	0.97	1.00	0.97
Index	0.23	0.00	0.27	0.28	0.01	0.29
0.3 Cutoff						
% Correct	0.84	0.83	0.84	0.85	0.79	0.85
Sensitivity	0.52	0.00	0.52	0.57	0.18	0.57
Specificity	0.91	1.00	0.91	0.91	0.92	0.91
Index	0.43	0.00	0.43	0.48	0.10	0.48

Notes: This table shows results from OLS regressions analyzing geographic determinants of fiscal pressure at the property owner level. The dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). I control for section fixed effects in column (1), (3), (4), and (6). *Property value* (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. *Street name* is a proxy for cadastral quality and is a dummy equal to one if a street name was entered. *Distance to closest Treasury* is the number of kilometers between the property and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the number of kilometers between the property and the closest Tax office, where agents who conduct registration and assessment activities work from. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, observations are weighted.

TABLE A.7
ECONOMIC AND GEOGRAPHIC DETERMINANTS OF BEING IN THE TAX NET (LOGIT)

	(1)	(2)	(3)
Dependent Variable (0,1)	TaxNet	TaxNet	TaxNet
Paid or received notif. in 2018			
Ln(PropertyValue)	0.467***	0.623***	
	(0.181)	(0.159)	
Multiple owner	0.675**	0.434	
	(0.265)	(0.353)	
Income Group 2	0.159	0.395	
	(0.309)	(0.334)	
Income Group 3	0.388*	0.739**	
	(0.232)	(0.362)	
Income Group 4	0.414	1.341***	
	(0.332)	(0.400)	
Rented	0.235	0.428	
	(0.278)	(0.317)	
N with regular income	-0.174*	-0.187	
	(0.097)	(0.122)	
Informal	-0.410	-0.629	
	(0.574)	(0.570)	
Non Employed	-0.780**	-0.932***	
	(0.338)	(0.357)	
Retired	-0.303	-0.393	
	(0.318)	(0.355)	
Street name			-0.203
			(0.311)
Distance to closest Treasury			-1.507
			(1.297)
Distance to closest Tax center			0.877
			(1.350)
Section FE	No	Yes	Yes
N	1108	831	831
Adj R2			
Mean of dep.	0.20	0.24	0.24
0.5 Cutoff			
% Correct	0.83	0.82	0.81
Sensitivity	0.07	0.34	0.24
Specificity	0.99	0.95	0.96
Index	0.06	0.29	0.20
0.3 Cutoff			
% Correct	0.79	0.81	0.79
Sensitivity	0.32	0.60	0.48
Specificity	0.89	0.86	0.87
Index	0.21	0.46	0.35

See notes on the next page.

Notes: This table shows results from Logit regressions analyzing economic (columns 1 and 2) and geographic (column 3) determinants of fiscal pressure at the property owner level. The dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). I control for section fixed effects in columns (2) and (3). *Property value* (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. *Street name* is a proxy for cadastral quality and is a dummy equal to one if a street name was entered. *Distance to closest Treasury* is the number of kilometers between the property and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the number of kilometers between the property and the closest Tax office, where agents who conduct registration and assessment activities work from. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.8
ECONOMIC DETERMINANTS OF FISCAL PRESSURE AT THE INDIVIDUAL LEVEL - TAX AMOUNT

	(1)	(2)	(3)
		OLS	
Dependent Variable	TaxAmnt	TaxAmnt	TaxAmnt
Ln(PropertyValue)	0.455***	0.167	-0.043
	(0.156)	(0.165)	(0.495)
Multiple owner		0.821**	0.360
		(0.318)	(0.812)
Income Group 2		0.369	0.453
		(0.355)	(0.895)
Income Group 3		0.287	0.501
		(0.321)	(0.644)
Income Group 4		0.796**	1.103
		(0.361)	(0.684)
Rented		0.392*	0.423
		(0.204)	(0.595)
N with regular income		-0.027	0.021
		(0.071)	(0.164)
Informal		0.181	-0.693
		(0.340)	(0.603)
Non Employed		0.430	-0.372
		(0.265)	(0.613)
Retired		0.026	-0.789
		(0.244)	(0.859)
Section FE	No	No	Yes
N	118	118	118
Adj R2	0.14	0.29	0.30
Mean of dep.	1.96	1.96	1.96

Notes: This table shows results from OLS regressions analyzing economic correlates of fiscal pressure at the property owner level. The dependent variable is $ln(TaxAmnt)$, the inverse hyperbolic sine transformation of tax amount paid. I control for section fixed effects in column (3). The sample is restricted to owners who pay the tax. *Property value* (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. Standard errors clustered at the section level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.9
ECONOMIC DETERMINANTS OF EVER RECEIVING AN INCOME VERSUS A PROPERTY TAX NOTIFICATION (OLS)

Dependent Variable (0,1)	(1) EverProperty	(2) EverProperty	(3) EverIncome	(4) EverIncome
Ln(PropertyValue)	0.028	0.031	0.024	0.009
	(0.039)	(0.034)	(0.018)	(0.020)
Multiple owner	0.092	0.031	0.066	0.046
	(0.068)	(0.055)	(0.083)	(0.074)
Income Group 2	-0.031	0.085	-0.012	-0.051
	(0.056)	(0.054)	(0.022)	(0.034)
Income Group 3	-0.104**	0.022	0.028	0.023
	(0.051)	(0.049)	(0.029)	(0.031)
Income Group 4	-0.062	0.154*	0.039	0.039
	(0.080)	(0.078)	(0.041)	(0.034)
Rented	0.096*	0.101**	-0.026	0.009
	(0.054)	(0.049)	(0.031)	(0.022)
N with regular income	0.008	0.002	-0.008	-0.001
	(0.014)	(0.012)	(0.005)	(0.005)
Informal	-0.072	-0.129*	-0.090*	-0.041
	(0.109)	(0.076)	(0.048)	(0.041)
Non Employed	-0.032	-0.024	-0.072**	-0.043*
	(0.064)	(0.049)	(0.029)	(0.025)
Retired	0.020	0.038	-0.023	-0.013
	(0.069)	(0.063)	(0.036)	(0.032)
Section FE	No	Yes	No	Yes
N	1085	1085	1085	1085
Adj R2	0.01	0.27	0.05	0.19
Mean of dep.	0.41	0.41	0.06	0.06
0.5 Cutoff				
% Correct	0.78	0.69	0.82	0.82
Sensitivity	0.21	0.58	0.00	0.02
Specificity	0.90	0.71	1.00	0.99
Index	0.11	0.29	0.00	0.01
0.3 Cutoff				
% Correct	0.22	0.50	0.82	0.81
Sensitivity	0.96	0.84	0.00	0.05
Specificity	0.06	0.43	1.00	0.97
Index	0.02	0.27	0.00	0.02

Notes: This table shows results from OLS regressions analyzing economic determinants of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner ever received an income tax notification. In columns (3) and (4), the dependent variable is a dummy equal to one if the property owner ever received a property tax notification. I control for section fixed effects in columns (2) and (4). *Property value* (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. For details on the data and computations, see Appendix A.4. *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.10
GEOGRAPHIC DETERMINANTS OF EVER RECEIVING AN INCOME VERSUS PROPERTY TAX NOTIFICATION (OLS)

Dependent Variable (0,1)	(1) EverProperty	(2) EverProperty	(3) EverIncome	(4) EverIncome
Street name		0.075*		-0.011
		(0.041)		(0.019)
Distance to closest Treasury		0.082		-0.114
		(0.213)		(0.123)
Distance to closest Tax center		0.186		-0.072
		(0.246)		(0.128)
Section FE	Yes	Yes	Yes	Yes
N	1085	1085	1085	1085
Adj R2	0.25	0.25	0.16	0.17
Mean of dep.	0.41	0.41	0.06	0.06
0.3 Cutoff				
% Correct	0.69	0.69	0.93	0.93
Sensitivity	0.90	0.91	0.26	0.23
Specificity	0.55	0.55	0.98	0.98
Index	0.45	0.46	0.24	0.21

Notes: This table shows results from OLS regressions analyzing geographic determinants of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner ever received an income tax notification. In columns (3) and (4), the dependent variable is a dummy equal to one if the property owner ever received a property tax notification. I control for section fixed effects in all columns. *Street name* is a proxy for cadastral quality and is a dummy equal to one if a street name was entered. *Distance to closest Treasury* is the number of kilometers between the property and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the number of kilometers between the property and the closest Tax office, where agents who conduct registration and assessment activities work from. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.11
DETERMINANTS OF TAX ENFORCEMENT AT THE CADASTER SECTION LEVEL - ROBUSTNESS 1

Dependent Variable	At least one in Tax Net	Share in Tax Net	Share in Tax Net
Ln(Average Property Value)	0.001	0.127***	0.146***
	(0.072)	(0.038)	(0.039)
Share with street name	0.098	0.052	0.025
	(0.105)	(0.056)	(0.058)
Distance to closest Treasury	-0.032**	-0.006	0.002
	(0.016)	(0.009)	(0.009)
Distance to closest Tax office	0.047	0.038**	0.030*
	(0.030)	(0.016)	(0.016)
N	155	155	129
R2	0.03	0.10	0.17
Mean of dep.	0.83	0.21	0.25

Notes: This table shows results from an OLS regression at the cadastral section level. In column (1) the dependent variable is a dummy equal to one if at least one respondent in the section is in the tax net in 2018. Being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared. In columns (2) and (3) the dependent variable is the share of respondents in the section which are in the tax net. In column (3) the sample is restricted to sections where at least one respondent is in the tax net. **Property value is predicted using the valuation formula described in Appendix A.2, and the sample is restricted to observations where the declared value is missing.** *Share with street name* is a proxy for cadastral quality and is computed as the share of surveyed properties for which a street name was entered. *Distance to closest Treasury* is the average number of kilometers between properties in a given section and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the average number of kilometers between properties in a given section and the closest Tax office, where agents who conduct registration and assessment activities work from. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018.

TABLE A.12
ECONOMIC DETERMINANTS OF FISCAL PRESSURE AT THE INDIVIDUAL LEVEL - ROBUSTNESS 1

	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
Dependent Variable (0,1)	TaxNet	TaxNet	TaxAmnt	TaxAmnt	TaxAmnt	TaxAmnt
lnval_r1	-0.170**	-0.045	-0.107	0.014	0.250	-0.000
	(0.071)	(0.146)	(0.100)	(0.254)	(0.412)	(.)
Multiple owner	-0.111	0.044	-0.080	0.221	0.171	-0.179
	(0.133)	(0.148)	(0.248)	(0.346)	(0.155)	(.)
Income Group 2	0.088	0.059	-0.122	-0.070	-0.536**	-0.947
	(0.079)	(0.097)	(0.108)	(0.251)	(0.225)	(.)
Income Group 3	0.237**	-0.067	0.084	-0.280	-0.017	-1.072
	(0.113)	(0.123)	(0.221)	(0.385)	(0.245)	(.)
Income Group 4	0.414***	0.246	0.090	-0.156	-0.110	-1.189
	(0.140)	(0.189)	(0.188)	(0.482)	(0.323)	(.)
Rented	0.150*	0.142	0.206	0.287	0.564*	-0.421
	(0.081)	(0.098)	(0.155)	(0.249)	(0.255)	(.)
N with regular income	-0.033	-0.013	-0.045	-0.056	-0.397*	-0.000
	(0.020)	(0.028)	(0.028)	(0.051)	(0.191)	(.)
Informal	-0.100	0.057	-0.425*	-0.402		
	(0.077)	(0.107)	(0.216)	(0.302)		
Non Employed	0.027	0.006	-0.401*	-0.380	-0.574**	-1.715
	(0.087)	(0.125)	(0.210)	(0.245)	(0.228)	(.)
Retired	-0.020	0.120	-0.456**	-0.354	-1.186***	-2.379
	(0.090)	(0.188)	(0.199)	(0.252)	(0.136)	(.)
Section FE	No	Yes	No	Yes	No	Yes
N	182	182	182	182	12	12
Adj R2	0.08	0.54	0.05	0.44	0.70	.
Mean of dep.	0.20	0.20	0.18	0.18	2.28	2.28
0.5 Cutoff						
% Correct	0.84	0.93				
Sensitivity	0.10	0.71				
Specificity	0.99	0.98				
Index	0.09	0.69				
0.3 Cutoff						
% Correct	0.79	0.94				
Sensitivity	0.45	0.84				
Specificity	0.85	0.96				
Index	0.30	0.80				

Notes: This table shows results from OLS regressions analyzing economic determinants of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). In columns (3) to (6), the dependent variable is $asinh(TaxAmnt)$, the inverse hyperbolic sine transformation of tax amount paid. I control for section fixed effects in columns (2), (4), and (6). In columns (5) and (6), the sample is restricted to owners who pay the tax. **Property value is predicted using the valuation formula described in Appendix A.2, restricted to cases for which declared value is missing.** *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.13
DETERMINANTS OF TAX ENFORCEMENT AT THE CADASTER SECTION LEVEL - ROBUSTNESS 2

Dependent Variable	At least one in Tax Net	Share in Tax Net	Share in Tax Net
Ln(Average Property Value)	0.077	0.186***	0.206***
	(0.083)	(0.047)	(0.050)
Share with street name	0.154	0.045	-0.025
	(0.102)	(0.058)	(0.061)
Distance to closest Treasury	-0.047***	-0.011	0.002
	(0.017)	(0.010)	(0.010)
Distance to closest Tax office	0.041	0.017	0.012
	(0.030)	(0.017)	(0.018)
N	181	181	143
R2	0.07	0.08	0.18
Mean of dep.	0.79	0.22	0.28

Notes: This table shows results from an OLS regression at the cadastral section level. In column (1) the dependent variable is a dummy equal to one if at least one respondent in the section is in the tax net in 2018. Being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared. In columns (2) and (3) the dependent variable is the share of respondents in the section which are in the tax net. In column (3) the sample is restricted to sections where at least one respondent is in the tax net. **Property value is predicted using the valuation formula described in Appendix A.2.** *Share with street name* is a proxy for cadastral quality and is computed as the share of surveyed properties for which a street name was entered. *Distance to closest Treasury* is the average number of kilometers between properties in a given section and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the average number of kilometers between properties in a given section and the closest Tax office, where agents who conduct registration and assessment activities work from. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018.

TABLE A.14
ECONOMIC DETERMINANTS OF FISCAL PRESSURE AT THE INDIVIDUAL LEVEL - ROBUSTNESS 2

	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
Dependent Variable (0,1)	TaxNet	TaxNet	TaxAmnt	TaxAmnt	TaxAmnt	TaxAmnt
lnval_r2	0.095**	0.117***	0.372***	0.384***	0.243	0.941*
	(0.038)	(0.039)	(0.093)	(0.101)	(0.183)	(0.525)
Multiple owner	0.081*	0.055	0.222	0.163	0.813**	-0.017
	(0.047)	(0.046)	(0.150)	(0.145)	(0.313)	(0.706)
Income Group 2	0.024	0.043	0.037	0.053	0.343	0.293
	(0.040)	(0.035)	(0.080)	(0.082)	(0.350)	(0.627)
Income Group 3	0.040	0.076*	0.067	0.152	0.297	0.082
	(0.036)	(0.044)	(0.093)	(0.114)	(0.289)	(0.757)
Income Group 4	0.069	0.127**	0.232*	0.309**	0.753**	0.768
	(0.048)	(0.049)	(0.120)	(0.129)	(0.315)	(0.733)
Rented	0.058	0.037	0.094	0.062	0.413**	0.107
	(0.036)	(0.034)	(0.079)	(0.068)	(0.189)	(0.645)
N with regular income	-0.019*	-0.017	-0.035	-0.035	-0.037	0.001
	(0.010)	(0.011)	(0.025)	(0.025)	(0.071)	(0.143)
Informal	-0.063	-0.067	0.073	0.092	0.145	-0.694
	(0.085)	(0.088)	(0.230)	(0.209)	(0.388)	(0.655)
Non Employed	-0.093*	-0.082**	-0.218**	-0.207**	0.430	-0.611
	(0.048)	(0.040)	(0.110)	(0.100)	(0.265)	(0.578)
Retired	-0.034	-0.041	-0.141	-0.233*	0.063	-1.048
	(0.048)	(0.042)	(0.110)	(0.127)	(0.243)	(0.834)
Section FE	No	Yes	No	Yes	No	Yes
N	1076	1076	1074	1074	117	117
Adj R2	0.06	0.27	0.10	0.31	0.29	0.37
Mean of dep.	0.19	0.19	0.33	0.33	2.68	2.68
0.5 Cutoff						
% Correct	0.83	0.86				
Sensitivity	0.01	0.34				
Specificity	1.00	0.97				
Index	0.01	0.31				
0.3 Cutoff						
% Correct	0.80	0.85				
Sensitivity	0.31	0.63				
Specificity	0.90	0.90				
Index	0.21	0.53				

Notes: This table shows results from OLS regressions analyzing economic determinants of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). In columns (3) to (6), the dependent variable is $asinh(TaxAmnt)$, the inverse hyperbolic sine transformation of tax amount paid. I control for section fixed effects in columns (2), (4), and (6). In columns (5) and (6), the sample is restricted to owners who pay the tax. **Property value is predicted using the valuation formula described in Appendix A.2.** *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.15
DETERMINANTS OF TAX ENFORCEMENT AT THE CADASTER SECTION LEVEL - ROBUSTNESS 3

Dependent Variable	At least one in Tax Net	Share in Tax Net	Share in Tax Net
Ln(Average Property Value)	0.062	0.083***	0.104***
	(0.048)	(0.027)	(0.031)
Share with street name	0.196*	0.017	-0.093
	(0.104)	(0.059)	(0.062)
Distance to closest Treasury	-0.043***	-0.006	0.003
	(0.015)	(0.009)	(0.009)
Distance to closest Tax office	0.038	0.015	0.013
	(0.029)	(0.017)	(0.017)
N	174	174	140
R2	0.08	0.04	0.14
Mean of dep.	0.80	0.22	0.27

Notes: This table shows results from an OLS regression at the cadastral section level. In column (1) the dependent variable is a dummy equal to one if at least one respondent in the section is in the tax net in 2018. Being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared. In columns (2) and (3) the dependent variable is the share of respondents in the section which are in the tax net. In column (3) the sample is restricted to sections where at least one respondent is in the tax net. **Property value (more precisely, annual rental value) is computed based on a series of survey questions.** *Share with street name* is a proxy for cadastral quality and is computed as the share of surveyed properties for which a street name was entered. *Distance to closest Treasury* is the average number of kilometers between properties in a given section and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the average number of kilometers between properties in a given section and the closest Tax office, where agents who conduct registration and assessment activities work from. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018.

TABLE A.16
ECONOMIC DETERMINANTS OF FISCAL PRESSURE AT THE INDIVIDUAL LEVEL - ROBUSTNESS 3

	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
Dependent Variable (0,1)	TaxNet	TaxNet	TaxAmnt	TaxAmnt	TaxAmnt	TaxAmnt
lnval_r3	0.090***	0.069***	0.256***	0.196***	0.194	-0.008
	(0.032)	(0.021)	(0.063)	(0.063)	(0.171)	(0.492)
Multiple owner	0.145**	0.051	0.107	-0.022	0.833**	0.452
	(0.059)	(0.045)	(0.161)	(0.223)	(0.334)	(0.858)
Income Group 2	-0.001	0.025	0.082	0.170	0.416	0.568
	(0.041)	(0.039)	(0.096)	(0.120)	(0.382)	(0.929)
Income Group 3	0.014	0.060	0.033	0.146	0.340	0.554
	(0.039)	(0.046)	(0.092)	(0.130)	(0.368)	(0.746)
Income Group 4	-0.007	0.099*	0.286**	0.394***	0.884**	1.084
	(0.064)	(0.059)	(0.121)	(0.145)	(0.383)	(0.770)
Rented	0.021	0.043	0.118	0.119	0.268	0.144
	(0.048)	(0.039)	(0.077)	(0.089)	(0.244)	(0.763)
N with regular income	-0.017	-0.014	-0.030	-0.034	-0.025	0.042
	(0.012)	(0.012)	(0.028)	(0.031)	(0.073)	(0.173)
Informal	-0.067	-0.028	0.198	0.153	0.260	-0.503
	(0.101)	(0.083)	(0.232)	(0.231)	(0.362)	(0.700)
Non Employed	-0.150**	-0.121***	-0.174*	-0.223*	0.563**	-0.232
	(0.063)	(0.044)	(0.103)	(0.116)	(0.277)	(0.763)
Retired	-0.063	-0.063	-0.068	-0.245	0.091	-0.799
	(0.062)	(0.046)	(0.106)	(0.154)	(0.262)	(1.037)
Section FE	No	Yes	No	Yes	No	Yes
N	926	926	924	924	106	106
Adj R2	0.09	0.29	0.10	0.26	0.30	0.27
Mean of dep.	0.20	0.20	0.35	0.35	2.72	2.72
0.5 Cutoff						
% Correct	0.83	0.86				
Sensitivity	0.03	0.32				
Specificity	1.00	0.97				
Index	0.03	0.29				
0.3 Cutoff						
% Correct	0.78	0.85				
Sensitivity	0.38	0.66				
Specificity	0.86	0.89				
Index	0.24	0.55				

Notes: This table shows results from OLS regressions analyzing economic determinants of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). In columns (3) to (6), the dependent variable is $asinh(TaxAmnt)$, the inverse hyperbolic sine transformation of tax amount paid. I control for section fixed effects in columns (2), (4), and (6). In columns (5) and (6), the sample is restricted to owners who pay the tax.**Property value (more precisely, annual rental value) is computed based on a series of survey questions.** *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.17
DETERMINANTS OF TAX ENFORCEMENT AT THE CADASTER SECTION LEVEL - ROBUSTNESS 4

Dependent Variable	At least one in Tax Net	Share in Tax Net	Share in Tax Net
Ln(Average Property Value)	0.047	0.119***	0.145***
	(0.057)	(0.033)	(0.037)
Share with street name	0.156	0.050	-0.037
	(0.102)	(0.058)	(0.062)
Distance to closest Treasury	-0.045***	-0.007	0.004
	(0.016)	(0.009)	(0.010)
Distance to closest Tax office	0.041	0.017	0.013
	(0.030)	(0.017)	(0.018)
N	181	181	143
R2	0.07	0.07	0.17
Mean of dep.	0.79	0.22	0.28

Notes: This table shows results from an OLS regression at the cadastral section level. In column (1) the dependent variable is a dummy equal to one if at least one respondent in the section is in the tax net in 2018. Being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared. In columns (2) and (3) the dependent variable is the share of respondents in the section which are in the tax net. In column (3) the sample is restricted to sections where at least one respondent is in the tax net. **Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. When missing, the value is replaced by the predicted property value using the valuation formula described in Appendix A.2, to which I add a residual drawn from a normal distribution with a mean of zero and a standard deviation equal to the standard deviation on observed residuals.** *Share with street name* is a proxy for cadastral quality and is computed as the share of surveyed properties for which a street name was entered. *Distance to closest Treasury* is the average number of kilometers between properties in a given section and the closest Treasury office, where which agents who distribute tax notifications work from. *Distance to closest Tax office* is the average number of kilometers between properties in a given section and the closest Tax office, where agents who conduct registration and assessment activities work from. Standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Source: property owner survey 2018.

TABLE A.18
ECONOMIC DETERMINANTS OF FISCAL PRESSURE AT THE INDIVIDUAL LEVEL - ROBUSTNESS 4

	(1)	(2)	(3)	(4)	(5)	(6)
			OLS			
Dependent Variable (0,1)	TaxNet	TaxNet	TaxAmnt	TaxAmnt	TaxAmnt	TaxAmnt
lnval_r4	0.008*	0.008*	0.004	0.009	-0.009	-0.004
	(0.004)	(0.004)	(0.007)	(0.008)	(0.027)	(0.085)
Multiple owner	0.150**	0.066	0.164	0.038	0.856***	0.334
	(0.065)	(0.043)	(0.154)	(0.214)	(0.302)	(0.639)
Income Group 2	0.028	0.048	0.073	0.130	0.427	0.456
	(0.042)	(0.033)	(0.093)	(0.100)	(0.334)	(0.836)
Income Group 3	0.083**	0.084**	0.132	0.139	0.355	0.473
	(0.035)	(0.041)	(0.091)	(0.125)	(0.281)	(0.645)
Income Group 4	0.115**	0.163***	0.463***	0.446***	0.877***	1.086
	(0.045)	(0.048)	(0.112)	(0.125)	(0.300)	(0.678)
Rented	0.067*	0.064*	0.236***	0.186**	0.533***	0.384
	(0.038)	(0.037)	(0.081)	(0.086)	(0.176)	(0.688)
N with regular income	-0.025**	-0.016	-0.040	-0.034	-0.033	0.020
	(0.012)	(0.011)	(0.025)	(0.025)	(0.068)	(0.163)
Informal	-0.077	-0.029	0.106	0.167	0.083	-0.686
	(0.103)	(0.087)	(0.231)	(0.218)	(0.326)	(0.643)
Non Employed	-0.125*	-0.083**	-0.223**	-0.214**	0.401	-0.380
	(0.064)	(0.038)	(0.103)	(0.099)	(0.259)	(0.583)
Retired	-0.065	-0.047	-0.151	-0.257**	0.056	-0.813
	(0.064)	(0.041)	(0.106)	(0.130)	(0.254)	(0.967)
Section FE	No	Yes	No	Yes	No	Yes
N	1108	1108	1106	1106	118	118
Adj R2	0.06	0.29	0.06	0.26	0.28	0.29
Mean of dep.	0.20	0.20	0.32	0.32	2.68	2.68
0.5 Cutoff						
% Correct	0.83	0.87				
Sensitivity	0.01	0.34				
Specificity	1.00	0.97				
Index	0.01	0.31				
0.3 Cutoff						
% Correct	0.79	0.84				
Sensitivity	0.28	0.58				
Specificity	0.89	0.90				
Index	0.17	0.48				

Notes: This table shows results from OLS regressions analyzing economic determinants of fiscal pressure at the property owner level. In columns (1) and (2), the dependent variable is a dummy equal to one if the property owner is in the tax net in 2018 (being in the tax net is defined as having paid the property tax and/or having received a tax notification for the property tax and is self-declared). In columns (3) to (6), the dependent variable is $asinh(TaxAmnt)$, the inverse hyperbolic sine transformation of tax amount paid. I control for section fixed effects in columns (2), (4), and (6). In columns (5) and (6), the sample is restricted to owners who pay the tax. **Property value (more precisely, annual rental value) is computed based on a series of survey questions, with some corrections applied using external valuations. When missing, the value is replaced by the predicted property value using the valuation formula described in Appendix A.2, to which I add a residual drawn from a normal distribution with a mean of zero and a standard deviation equal to the standard deviation on observed residuals.** *Multiple owner* is a dummy equal to one if the owner possesses another property in the region. *Income group* is a categorical variable splitting the sample into groups of similar size, using responses to a bracket based survey question on total monthly household income. *Rented* is a dummy equal to one if at least part of the property is rented. *N with regular income* indicates the number of people within the household earning a regular income. *Informal*, *Non-employed* and *Retired* are modalities of a categorical variable where the reference category is *Formal employment*. The bottom panel of the Table displays predictive performance measures of the different models. Under the 0.5 (respectively, 0.3) cut-off, the predicted outcome is $TaxNet = 1$ if the combination of covariates and coefficients is larger than or equal to 0.5 (resp., 0.3). *% Correct* is the percentage of correct predictions, equal to the number of correct predictions over total number of observations. *Sensitivity* is equal to correct predicted positives over total true positives. *Specificity* is equal to correct predicted negatives over total true negatives. *Index* corresponds to the Youden index and is computed as $Sensitivity + Specificity - 1$. Standard errors clustered at the section level are in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.10. Source: property owner survey 2018, restricted to respondents who are either the owner or a close family member of the owner, observations are weighted.

TABLE A.19
CODEBOOK PROPERTY CHARACTERISTICS

		N	Mean	Min	Max	SD
Annual Rental Value (CFA)		1458	3,868,226	240,000	48,000,000	4,119,224
Built Area		1458	317	24	2585	250
		N	**%**			
Sector						
	10000	205	14.1			
	40000	82	5.6			
	50000	64	4.4			
	53000	254	17.4			
	65000	3	0.2			
	70000	13	0.9			
	80000	15	1			
	90000	1	0.1			
	100000	287	19.7			
	110000	357	24.5			
	150000	49	3.4			
	200000	109	7.5			
	220000	4	0.3			
	300000	15	1			
Usage: commercial or mixed						
Residential		1120	76.8			
Commercial		8	0.5			
Mixed		330	22.6			
Fence: Type						
None		861	59.1			
Metal		9	0.6			
Wall		517	35.5			
Wall w. wrought iron		71	4.9			
Fence: state						
Very Good		191	13.1			
Average		1137	78			
Bad		130	8.9			
Wall: Cement						
No		6	0.4			
Yes		1452	99.6			
Cladding: type						
Coating: Wis		171	11.7			
Coating: Plain		63	4.3			
Paint		802	55			
Tiles		291	20			
Stone		18	1.2			
None		113	7.8			
Cladding: state						
Very Good		293	20.1			
Average		833	57.1			
Bad		332	22.8			
Tiles						
No		1071	73.5			
Yes		387	26.5			
Quality Doors and Windows						
Very Good		290	19.9			
Average		835	57.3			
Bad		333	22.8			
Landscape Improvement						
No		1179	80.9			
Yes		279	19.1			
Architectural Improvement						
No		1136	77.9			
Yes		322	22.1			
Garage						
No		965	66.2			
Yes		493	33.8			
Garage						
Simple		410	28.1			
Double		83	5.7			
None		965	66.2			
Shop						
No		1191	81.7			
Yes		267	18.3			
Balcony						
No		544	37.3			
Yes		914	62.7			
Floors						
	0	640	43.9			
	0.5	4	0.3			
	1	561	38.5			
	1.5	12	0.8			
	2	185	12.7			
	3	42	2.9			
	4	14	1			
Main Road						
On Main Road		374	25.7			
Near Main Road		388	26.6			
Off Main Road		696	47.7			
Pavement						
Tarmac		468	32.1			
Pavements		134	9.2			
Gravel		46	3.2			
Sand		806	55.3			
None (narrow)		4	0.3			
Sidewalk						
No		548	37.6			
Yes		910	62.4			
Angle		0	0			
No		1058	72.6			
Yes		400	27.4			
Street Lights						
No		370	25.4			
Yes		1088	74.6			

Notes: This table displays descriptive statistics on property characterisctics, which were collected during the property owner survey. The characteristics were determined through collaborative work with the Cadaster and Valuation department of the DGID, and the African Property Tax Initiative. This table is restricted to the sample used for the calibration of the property valuation formula.

TABLE A.20
COMPARING PERFORMANCE STATISTICS

	RMSE	R2
OLS	0.51	0.56
OLS with interactions	0.55	0.54
LASSO	0.51	0.56
LASSO with interactions	0.51	0.56
Elastic net	0.51	0.55
Elastic net with interactions	0.51	0.56

Notes: This table displays performance statistics for different specifications of the property valuation models. In all specifications, the logarithm of the property value is regressed on all retained characteristics. Results are obtained through 100-fold cross validation, and each estimation is run two times, one where the RMSE is minimized, and one where the R2 is maximized. The first (resp. second) column displays the average RMSE (resp. R2) over the 100 iterations. In the models with interactions, each observable property characteritic is interacted with a categorical variable indicating whether the Sector is of high, standard or low value, based on average price per square meter. See Appendix A.2 for details on the valuation model.

TABLE A.21
ELASTIC NET REGRESSION
PERFORMANCE STATISTICS

Mean R2	56%
Mean RMSE	0.49
MAPE	41
Freddie Mac 10%	16
Freddie Mac 20%	32
Freddie Mac 40%	62

Notes: This table displays performance statistics for the retained property valuation model. MAPE is the mean absolute percentage error and is computed as the mean of the percentage difference between predicted and actual values. Freddie Mac 10% refers to the Freddie Mac Criterion, the share of predicted values which fall within 10% of the actual value. I also display the equivalent for 20 and 40% thresholds. See Appendix A.2 for details on the valuation model.

TABLE A.22
PREDICTING PROPERTY VALUE: ELASTIC NET REGRESSION RESULTS

	Log(Value)
Log(Area)	0.31
Sector	
10000	−0.39
40000	−0.19
50000	−0.04
53000	−0.34
65000	0.00
70000	0.00
80000	−0.24
90000	−0.18
100000	−0.08
110000	0.00
150000	0.06
200000	0.20
220000	0.00
300000	0.31
Floors	0.21
Usage	
Residential	0.00
Commercial	−0.19
Mixed	0.15
Road: Type	
Tarmac	0.05
Pavements	0.08
Gravel	0.08
Sand	0.00
None	−0.04
Sidewalk	0.00
Angle	0.06
Street Lights	0.04
Constant	12.66
N	1458
Elastic Net Regression	
Cross-validation Iterations	100
Mean RMSE	0.49
Mean R2	0.56
Fence: Type	
None	0.00
Metal	0.00
Wall	0.06
Wall w. wrought iron	0.00
Fence: State	
Very Good	0.00
Average	0.00
Bad	−0.06
Wall: Cement	0.13
Cladding: Type	
Wis	−0.01
Plain	−0.19
Paint	0.00
Tiles	0.00
Stone	0.14
None	−0.06
Cladding: State	
Very Good	0.02
Average	0.00
Bad	−0.04
Any Tiles	0.05
Quality Doors and Windows	
Very Good	0.14
Average	0.00
Bad	0.00
Landscape	0.13
Architecture	0.00
Garage	
Simple	0.10
Double	0.17
None	0.00
Balcony	0.05
Road: Location	
On Main Road	0.10
Near Main Road	0.07
Off Main Road	0.00

Notes: This table displays the resulting coefficients for the property valuation formula. The coefficients are estimated through a cross-validation methodology: an elastic net regression is run 100 times, using different calibration and test subsamples in each iteration. The retained coefficients are the median of all iterations. The coefficients which are equal to zero were found not to have a significant impact in the elastic net regression. See Appendix A.2 for details on the valuation model.

TABLE A.23
PROPERTY VALUATION FORMULA ILLUSTRATION

Built Area	116.88
Floors	2
Total Area	350.62
Sector	100,000
Usage	Mixed
Fence: Type	Absence
Fence State	Standard
Wall: Ciment	Yes
Cladding: Type	Paint
Cladding: state	Bad
Tiles	No
Quality Doors and Windows	Bad
Landscape Improvement	No
Architectural Improvement	No
Garage	Absence
Shop	No
Balcony	Yes
Main Road	Off Main Road
Pavement	Tarmac
Sidewalk	Yes
Angle	No
Street Lights	No

Notes: This is an illustrative case for the property valuation formula. The property under study is displayed on the left, and the box on the right lists the observable characteristics that are entered during the agent's field work. This example comes from the baseline survey, more precisely from the municipality of Medina in Dakar. See Appendix A.2 for details on the valuation model.

TABLE A.24
TAX AMOUNT AND TAX NOTIFICATION OUTCOME

Dependent Variable (0,1)	(1) In Data	(2) In Data	(3) Distributed	(4) Distributed
Ln(Tax liability)	0.028***	0.019***	-0.041***	-0.002
	(0.001)	(0.001)	(0.004)	(0.004)
Office FE	No	Yes	No	Yes
N	55953	55953	10017	10017
Adj R2	0.01	0.13	0.01	0.08
Mean of dep.	0.18	0.18	0.70	0.70

Notes: This table shows results from OLS regressions where the dependent variable in columns (1) and (2) is a dummy defined for each tax notification of the property tax assessment valuation roll, equal to one when information was obtained about the tax notification's outome during the distribution phase. In columns (3) and (4), the sample is restricted to tax notifications for which the outcome is known ($In\ Data = 1$), and the dependent variable is a dummy equal to one if the tax notification was distributed. The distribution status is reported by Treasury agents. The only covariate included is the logarithm of tax liability amount. Treasury office fixed effects are included in columns (2) and (4). Source: Tax notification outcome survey 2019-2020, DGID property tax assessment data, 2019.