Discretion versus Algorithms: Bureaucrats and Tax Equity in Senegal*

Justine Knebelmann
(Sciences Po Paris)

Victor Pouliquen
(University of Essex)

Bassirou Sarr
(Ministry of Finance, Senegal)

October 16, 2024

Abstract

The implementation of government programs requires a list or register of individuals who are eligible for the program. Building these registers accurately is a challenge in contexts with low administrative capacity, and is often the responsibility of bureaucrats. We study the impacts of giving such bureaucrats more (or less) discretion in building these registers in the context of the creation of the first digitized property tax register in Senegal. We randomly assign neighborhoods to valuation methods with different degrees of bureaucrat discretion and compare the registered property values against a benchmark of market values provided by licensed real estate assessors. Bureaucrats in full discretion areas undervalue properties, and more so for higher-value properties, resulting in a regressive tax profile. The median tax rate is 3.8% in the lowest quintile and 1.7% in the top quintile, instead of the expected 4.4% and 8.6% rates based on the tax code. In contrast, a rule-based system where bureaucrats record property characteristics (not values) that an algorithm then uses to compute values, significantly reduces this tax gap. A pure rule with no bureaucrat inputs yields the highest accuracy and equity. We show this is due to bureaucrats' lack of knowledge about high-end properties and their fairness concerns, and not due to collusion between bureaucrats and property owners.

*We thank the Senegalese Tax Administration (*Direction Générale des Impôts et Domaines* or DGID) for excellent collaboration. We thank Denis Cogneau, Marc Gurgand, Tavneet Suri for invaluable advice and constructive suggestions. This project was funded by: Economic Development and Institutions, J-PAL DigiFI and the Fund for Innovation in Development (ICTD, J-PAL GI, EUR-PGSE, IGC provided complementary funding). We thank the *Centre National d'Etudes Spatiales* for making high resolution satellite images accessible through their Geosud program. This RCT was registered as AEARCTR-0002277. We thank Samuel Allain, Aramata Badji, Adrien Ciret, Bilal Choho, Ndeye Rokhaya Diao, Papa Daouda Diene, Oumar Fara Diop, Mor Tacko Hane, Ousseynou Niang, Nicolas Orgeira for their contributions to project operations and research assistance. The views expressed in this paper are our own and do not reflect the views of the DGID nor of our employers. All errors are our own.

1 Introduction

The ability of a government to gather accurate individual-level information is a crucial component of state capacity (Lee & Zhang, 2017; Scott, 1999). Indeed, both the implementation and the effectiveness of many public policies rely on the availability of comprehensive administrative data. This is equally important for targeted anti-poverty programs (Hanna & Olken, 2018), and for taxation (Pomeranz, 2015; Bowles, 2023), as both require the identification and registration of eligible individuals, as well as an evaluation of the amount they should receive or contribute. In low-income countries, weak administrative capacity can lead to such registers built by the government being incomplete and outdated if at all they exist (Banerjee *et al.*, 2019). This can have far-reaching implications, for instance if many of the poorest households are not registered for an anti-poverty program, or if the wealth of the richest is strongly underestimated by the tax administration.

To build or update registers, governments in low-income countries often lack harmonized protocols, and instead rely on the discretion of bureaucrats. Discretion has advantages if bureaucrats have access to granular information that might not be integrated in a systematized process (Duflo *et al.*, 2018). However, discretion can be costly if bureaucrats' preferences are misaligned with the government's objective (Bandiera *et al.*, 2023) or if they engage in rent-seeking (Niehaus *et al.*, 2013). Discretion might also pose challenges if bureaucrats lack the required knowledge or skills (Rogger & Somani, 2023), or if they are heterogeneous and difficult to screen. Rule-based processes that incorporate algorithms may be implemented to reduce bureaucrats' discretion (Aghion & Tirole, 1997). Data-driven rule-based processes are increasingly available thanks to digitization, the effects of which can be transformative for governance outcomes (Muralidharan *et al.*, 2016; Banerjee *et al.*, 2020, 2023; Dzansi *et al.*, 2022).

In this paper, we investigate how the discretion of bureaucrats involved in the expansion of a tax register affects its quality and equity, and we assess whether a rule-based process leads to more accurate and fairer outcomes. We do so by leveraging the roll-out of the first digitized property tax census carried out in Dakar, the capital city of Senegal. We experimentally vary the degree of discretion given to bureaucrats in determining the rental value of properties, which serves as the base of the property tax.[1] The objective of the administration is to tax to its full potential by registering all properties, assigning values as close as possible to market values, while preserving horizontal equity – two properties with the same underlying value should face a similar tax bill – and vertical equity – more expensive properties should face higher tax bills.

We randomly assign 94 cadastral sections (small neighborhoods) covering approximately 40,000 properties into two treatment arms. In the first arm, which we call the *discretionary* arm, the property's value is determined by bureaucrats' judgment of this value. Discretion is the status quo in this context. It is also a reasonable approach since there is no pre-existing information on real estate values, and bureaucrats can see properties up close, get a sense of neighborhood amenities, and potentially ask occupants about rents.[2] In the second arm, the *rule-based* arm, an algorithm predicts property values by combining information from satellite images with observable characteristics entered by bureaucrats. The characteristics are easy to verify for the administration, which limits the bureaucrats' discretion.[3] We also compare these two arms to a *pure rule* which predicts property values using only variables that can be recovered remotely, without any bureaucrat input.[4] The pure rule has the advantages of totally removing the bureaucrats' role and drastically reducing costs. Finally, to compare these three approaches, we use a benchmark of market valuations by nationally certified real estate assessors that we obtained for a subset of 2,290 properties.[5]

[1]The tax base is the market rental value of the property, the value that is or could be obtained from the property rented at market prices.

[2]It is standard for property tax liabilities to be determined in a discretionary way in low- and middle-income countries, due to the absence of comprehensive real estate data, and a lack (or lack of affordability for governments) of certified assessors – see Chapter 2 in Franzsen & McCluskey (2017) for details.

[3]We developed this innovative rule-based system by tailoring Computer Assisted Mass Appraisal methods to the context, collaborating with the administration and international practitioners. The algorithm includes eighteen directly observable characteristics, a location fixed-effect, and built area. Other cities in low- and middle-income countries are considering adopting similar approaches – see Knebelmann (2022) for a review of property tax digitalization initiatives.

[4]The R^2 of the rule with all covariates is 0.90, and the R^2 of the pure rule, using only section fixed-effect and built area, is 0.87. We calibrate the rules on a sample of 4,921 valuations by real estate assessors, using an elastic-net regression and five-fold cross-validation.

[5]There is no pre-existing data with wide coverage on real estate prices in Dakar which is a common feature in low-income countries (Behr *et al.*, 2023). Reassuringly, we find a 62 percent correlation between these assessor values and self-reported values from our owner baseline survey. We also scrape values from online rental listings, with the caveat that we can only situate these properties within a broad set of neighborhoods

The deployment of the 268 bureaucrats involved in the census is orthogonal to treatment and they are assigned to properties in a quasi-random way. These operations being a first-of-a-kind in Dakar, tax office managers do not have any prior information on bureaucrats' performance they could use for assignments, nor are they aware of the valuation method assignment *ex ante*.[6] Therefore, our setting offers a unique opportunity to measure the influence of bureaucrats and their discretion on the shaping of the tax roll.

Our first main result is that under discretion, the bureaucrats substantially undervalue properties, and more so for expensive ones, leading to a regressive tax profile. The median assessment ratio, defined as the bureaucrats' value over market value, is 0.50. It varies substantially which harms horizontal equity. The median effective tax rate is 3 percent (instead of the expected 8.2 percent based on the tax code), it decreases from 3.8 percent in the lowest quintile of property values to 1.7 percent in the top quintile, against the expected 4.4 and 8.6 percent respectively.[7] We exploit the quasi-random assignment of the bureaucrats' to properties to estimate bureaucrat fixed-effects, and find that they explain 40 percent of the variation in the tax base gap – defined as the absolute difference between bureaucrat value and market value.

Second, the rule-based process yields substantial improvements in both accuracy and tax equity. The median assessment ratio is 0.95. Leveraging our randomization, we estimate that in comparison, discretion increases the tax base gap by 83 percent (significant at the 1 percent level).[8] The median effective tax rate is 6.8 percent. It is 7.4 percent in the lowest quintile of properties and 5.2 percent in the top quintile. Rule-based valuations are

due to a lack of addressing information. As expected, on average, these values are closest to the top quintile of assessors' values, since online postings tend to cover a selected high-end segment of the real estate market.

[6]Once the census starts in a section, it needs to be completed before bureaucrats move to the following area. This prevents any sorting of bureaucrats once the valuation method is known. In Section 4, we show that bureaucrat characteristics are balanced across treatment arms and do not correlate with market values.

[7]The statutory tax rate is 8.6 percent, but there is a reduction in the tax base by a fixed amount if the property is the owner's main residence. The expected effective tax rate is thus increasing in property values, since for cheaper properties, the reduction represents a larger share of total property value.

[8]Two factors allow us to be confident that these results are not mechanically driven by the fact that market values that are used as the benchmark and market values that are used to calibrate the rule are from the same source (assessor valuations). First, the sample used for the calibration of the rule includes 4,921 properties, only half of which are in our analysis sample, and the calibration is done through cross-validation. Second, we perform a robustness check with an alternative rule that we calibrate on self-reported values from our owner baseline survey. Although this algorithm is of lower quality ($R^2 = 0.33$), we still find that discretion significantly widens the tax base gap compared to this rule.

less dispersed and less subject to bureaucrat heterogeneity, improving horizontal equity.

Third, the pure rule strictly dominates the rule-based process. The median assessment ratio is in the vicinity of 1 throughout the distribution, except for the lowest quintile of properties (1.26). In comparison, discretion increases the tax base gap by 166 percent, and the rule-based process by 65 percent (significant at the 1 percent level in both cases). The median effective tax rate is 7.3 percent. It is 6.9 percent in the lowest quintile of properties and 7 percent in the top quintile. The pure rule performs better than the rule-based process in spite of the fact that its model-fit statistics on the calibration sample are lower. This occurs because of the distortions created by the bureaucrats: in some instances, the bureaucrats do enter incorrect property characteristics. Hence, even limited discretion has large effects on the tax profile.

Finally, we investigate the mechanisms underlying why the bureaucrats may make inaccurate valuations. In a lab-in-the-field experiment, bureaucrats are shown pictures of properties and asked to report their best estimates of the value. On average, a high-end property is undervalued by so much that it is 'shifted down' from the fifth to the third quintile of market values.[9] This highlights the role of the knowledge channel. Second, we rule out rich owners bribing bureaucrats in exchange for lower tax liabilities. The above finding when there are no possible (corruption) gains from undervaluation is a first piece of evidence. A second is that we find the same undervaluation gradient irrespective of whether or not the bureaucrat met the owner during the field visit. Lastly, we find that the bureaucrats are biased by their perceptions of fairness. In the lab-in-the-field valuations, the bureaucrats who are told that the owner is retired (versus employed – retired being associated with vulnerability in this setting) provide a 38 percent lower value.

These results have significant implications for local public finance outcomes. Total tax liabilities amount to 8 billion FCFA in the discretionary arm instead of 19.6 billion FCFA if we extrapolate the benchmark market values. In the rule-based arm, liabilities amount to 11 billion FCFA with the rule-based process and 19.7 billion FCFA in the case of the pure rule. In a country with 3.4 billion FCFA of total property tax collections, the

[9] We do not find evidence of learning. A randomized information treatment aiming to update bureaucrats' beliefs on the distribution of market values is not enough to correct their misconceptions. We also find that there is no learning in the field over time, nor from exposure to the rule.

implications from undervaluation are immense.[10] Importantly, under discretion, 49.6 percent of the tax burden is owed by the 10 percent most expensive properties, while this figure is 63 percent under the rule, and 70 percent under the pure rule. Removing discretion shifts the tax burden from low-value to high-value property owners. Does this mean that the optimal strategy for the administration should be to rely solely on a pure rule? Overall, the pure rule yields the most reliable values.[11] But for the cheapest properties a trade-off exists: even if it is still more accurate than discretion strictly speaking, the pure rule is likely to overvalue in this segment of the market.[12] At the same time, it is precisely for these properties that bureaucrats are performing relatively better.[13] If the government wants to minimize the risk of overvaluing the cheapest properties, while maximizing accuracy, an optimal policy is to use discretion for the lowest quintile of the market, and the pure rule elsewhere.[14] The government may aim to avoid overtaxing the owners of the cheapest properties, likely the poorest owners, due to welfare considerations.

1.1 Contribution to the Literature

To the best of our knowledge, our study is the first experimental comparison of an algorithm with discretion in the creation of a tax roll at scale.[15] In doing so, we contribute to the literature questioning the respective advantages of strict rules versus discretionary

[10]These amounts correspond to: 13 million USD instead of 32 million USD; 18 million USD and 32 million USD; 5.5 million USD. Baseline municipal revenues in Dakar from all local taxes are 28 billion FCFA or 45 million USD (Delbridge *et al.*, 2022).

[11]Under discretion, some gains could be made by selecting more highly educated bureaucrats, since this characteristic correlates with bureaucrats' fixed-effects. However, at best, top bureaucrats perform as well as the rule.

[12]This is an inherent feature of property valuation models due to unobserved variable bias, see McMillen & Singh (2020); Berry (2021); Amornsiripanitch (2023) for demonstrations in the United States context where rule-based methods are widespread.

[13]We show that private information obtained by bureaucrats under discretion (when the property is rented, or when they meet the owner) helps improve valuations only for properties with lower values.

[14]One could be worried about dynamic adjustments by property owners if they learn which characteristics used in the algorithm trigger higher tax liabilities. Using a pure rule which relies only on location and built area mitigates this risk.

[15]Casaburi & Troiano (2015) show that property tax revenues increase after an Italian government program successfully detected unregistered buildings with aerial imagery, but the paper focuses on municipality-level financial and political outcomes, and not on bureaucrats' behavior when generating the tax roll. Battaglini *et al.* (2022) in Italy and Black *et al.* (2022) in the United States provide simulations from algorithm-based audit selection. Their studies rely on data from former audits rather than an experiment, and again, do not allow to study tax officials' behavior.

judgment in the work of organizations (Aghion & Tirole, 1997; Dessein, 2002). Our findings contrast with empirical evidence from other contexts where discretion was shown to yield better outcomes (Duflo *et al.*, 2013, 2018; Bandiera *et al.*, 2021; Kala, 2023; Decarolis *et al.*, 2021). It is worth noting that the agency problem in taxation differs in nature from other settings (Gordon, 2017): while tax codes invariably contain rules on paper, discretion always comes into play at one point or another of the taxation process. This accentuates the importance of exploring tax officials' discretionary decision-making.[16] A distinctive advantage of our setting lies in our ability to directly quantify the effects of different degrees of discretion on outcomes since we can independently observe the tax base. Our study also provides novel insights into the potential of algorithmic systems adopted by governments (Kleinberg *et al.*, 2018; Haseeb & Vyborny, 2022; Greenstone *et al.*, 2022). While discretion fares significantly worse in terms of tax fairness, it's important to acknowledge that the rules are more likely to overvalue properties owned by poorer individuals (Avenancio-León & Howard, 2022; McMillen & Singh, 2020; Elzayn *et al.*, 2023). Our findings on the rule-based process suggest that as soon as the implementation of a rule is delegated, social planners must expect possible deviations (Niehaus *et al.*, 2013; Björkegren *et al.*, 2022).[17]

Furthermore, our study sheds new light on how bureaucrats shape policy outcomes. While former studies have measured the influence of bureaucrat heterogeneity (Best *et al.*, 2023; Fenizia, 2022; Limodio, 2021) including on tax enforcement (Bergeron *et al.*, 2022; Khan *et al.*, 2016, 2019), our study innovates by allowing to measure bureaucrats' individual performance with respect to a benchmark. This perspective allows us to delve deeper into the "black-box" of bureaucrat discretion.[18] Our finding that the loss in tax

[16] Bachas *et al.* (2021) provide evidence on the effect of reduced bureaucrat discretion for the targeting of tax audits in Senegal, and Okunogbe & Pouliquen (2022) for tax filing in Tadjikistan. While we study discretion in the creation of the tax register, in high-income countries, where there is more information on the tax base, this trade-off is perhaps more relevant in other stages of the taxation process such as management of delinquent taxpayers and audits.

[17] Important dimensions that are beyond the scope of this paper but that we hope to address in follow-up work are the potential of interactions between algorithmic predictions and human decisions (Sadka *et al.*, 2018; Agarwal *et al.*, 2023) as well as political resistance to the adoption of of algorithms (Browne *et al.*, 2023).

[18] Former studies often measure outcomes that are more aggregate (at the office level, or the bureaucrat-team level, for instance), and that depend on a combination of decisions made by the bureaucrat and another actor (for example, realized procurement prices).

liabilities under discretion results from bureaucrats' insufficient knowledge rather than collusion aligns with the results in Rogger & Somani (2023) and resonates with the concept of passive waste developed in Bandiera *et al.* (2009).[19] Finally, we break new ground by showing how misconceptions about the distribution of wealth on the part of those implementing tax policy can decrease tax fairness building on Hvidberg *et al.* (2023); Hoy (2022); Stantcheva (2021) who measure these misconceptions among the general population.

Lastly, we contribute to the literature on state capacity through the lense of property taxation (Besley & Persson, 2009; Weigel, 2020; Balan *et al.*, 2022; Okunogbe, 2021; Dray *et al.*, 2023). While many studies have demonstrated the potential of informational investments, such as cadastral updates, to boost tax revenue (Martínez, 2023; D'Arcy & Nistotskaya, 2018; Christensen & Garfias, 2021; vom Hau *et al.*, 2023), we address the key question of *how* governments can build this capacity. By providing evidence on the different modalities through which technology can be leveraged for a large and equitable expansion of the tax net, our study paves the way for other promising applications in taxation (Okunogbe & Santoro, 2023).[20] We focus on "upstream" investments in state capacity, i.e. building a reliable information set, while Dzansi *et al.* (2022) explore the role of technology in distributing property tax bills in Ghana, addressing a subsequent stage in the taxation process.

[19] Bandiera *et al.* (2023) also find that delegation to bureaucrats makes a policy less progressive, but for different reasons (social ties).

[20] Levitt & Syverson (2008) quantify the effects of asymmetric information in the reporting of property prices in the United States. Studies with our level of granularity of real estate data are very rare in low- and middle-income countries (an exception being Anagol *et al.* (2022) in India). Brockmeyer *et al.* (2021) in Mexico and Bergeron *et al.* (2023) in the Democratic Republic of the Congo provide insights on how the government should set optimal property tax liabilities, but do not pose the question of how the administration can generate information on the tax base at scale to start with.

2 Context and Experimental Design

2.1 Background

In Senegal, all property owners are subject to a property tax. The national tax administration is responsible for generating the tax bills.[21] The tax base is the annual market rental value, the value that is or could be obtained from the property if rented at market prices. The rate is 8.6 percent, with an abatement for owner occupied properties, for which the share of the tax base below 1.5 million FCFA is only taxed at 3.6 percent. In theory, owners are supposed to come to the tax administration office once a year to declare their property's value, but only a very small minority respect this obligation. The administration is legally enabled to conduct field work to register new properties. In doing so, before the digital property tax census, bureaucrats used their own judgment to determine property values: the status quo is what we call discretionary valuation. In practice, these field operations were very rare before the program. As such the values on the pre-program tax roll originate either from self-declarations by owners, or from discretionary valuations by bureaucrats at one point in time, with no efficient strategy for expansions or updates.

Indeed, there is no systematic information on real estate prices. This is typical in cities of low-income countries (Behr *et al.*, 2023). Although 41 percent of properties are rented, only 2.6 percent have some form of contract, and these are rarely reported to the administration.[22] Rental agreements are typically channeled through informal brokers and real estate agencies mostly cover high-end segments of the market. When there are property sales outside of informal inheritance, they are rarely reported to the administration.[23] More generally, a significant challenge to use any type of data on property prices is the lack of a systematic addressing system, making it difficult to match entries across sources. Before the program, the administration did not make use of any data-driven strategy to improve the property tax roll, whether it be by computing neighborhood-level

[21] In Appendix section A.1, we provide additional details on the institutional context.

[22] Source: data from the property tax census operations described in Section 2.2. In our baseline property owner survey the share of properties that are rented is 30 percent.

[23] The record of property transfers the administration has only includes around 2,000 transactions reported by a dozen of notary offices over the past ten years.

rental values, by using built area measurements, or even by assigning precise location details to already registered properties.

The tax strongly underperforms, a first order problem being the low registration ratio (around 16 percent): in the region of Dakar, approximately 61,000 properties were on the pre-program tax roll against an estimated total of 370,000. The payment ratio is also low: in 2022, a payment was recorded for 10 percent of the tax bills, and collections amounted to 16 percent of liabilities. Total revenues for the region were 3.4 billion FCFA or 5.8 million USD in 2022.[24] This underperformance of the property tax is a common feature in African countries (Franzsen & McCluskey, 2017).[25] However, in a context of recent decentralization policies and of willingness to expand the tax net while leveraging digitization, there is now in Senegal strong political will to strengthen the property tax.

2.2 Experimental Design

Our experiment is embedded in the roll-out of the first digital property tax census in the region of Dakar. We developed a new property tax management software in collaboration with the administration. It is innovative along two dimensions. First, it allows bureaucrats to conduct the census on tablets, and incorporates pre-loaded plot identifiers[26] and GIS coordinates. Second, it enables a semi-automatized valuation method based on an algorithm. Figure 2 provides illustrations.[27]

We introduce experimental variation in the degree of discretion bureaucrats have in the determination of property values. Our unit of randomization is the cadastral section. Sections are small neighborhoods with 417 plots on average, and are the unit the adminis-

[24]These numbers are computed using administrative data from the tax administration and the national treasury. See Knebelmann (2021) for a more in-depth exploration of the challenges in the functioning of the property tax before the program.

[25]Recent micro-estimates of property tax compliance rates highlight a similar problem in other countries: 10 percent of properties pay their bill in Kampala, Uganda (Manwaring & Regan, 2023), as few as 4.4 percent of properties are registered and 2.2 percent pay in some areas of Monrovia, Liberia (Okunogbe, 2021).

[26]A plot is the piece of land on which a property is built. Each plot has a precise delimitation and a unique identifying number. These were established by the administration long before the census started. We use the words "plot" and "property" interchangeably.

[27]In Appendix Section A.2, we provide additional details on the new property tax management system. Similar digital applications for property taxation with GIS functionalities are still very new on the African continent but are being increasingly adopted (Dzansi *et al.*, 2022; Knebelmann, 2022).

tration uses to organize field work. In 48 sections, valuation follows a fully discretionary method: bureaucrats use their knowledge and judgment to assign values. In 48 other sections, properties are valued using the new rule-based method combining characteristics collected by bureaucrats and data from satellite images. Third, we apply to this second arm an alternative algorithm, a pure rule which predicts property values using only variables that can be computed remotely, without any bureaucrat input. Figure 3 shows the geographical scope of the study and the experimental design. Taken together the two treatment arms include 42,423 plots. There are an additional 97 pure control sections, in which the property tax census did not take place.[28] The following variables were used for stratification: tax office, section size, and share of plots eligible for taxation according to the baseline survey.[29]

The bureaucrats are 268 temporary employees selected, trained, managed and paid by the administration. Bureaucrat assignment to plots is quasi-random. Assignment to sections is planned by the tax office and is orthogonal to the randomization (72.5 percent worked in both arms). The valuation method is programmed directly in the application. Neither tax office managers nor the field bureaucrats are aware of it before starting a given section, and once a section is started, it must be completed before the bureaucrats are deployed to the next. The program being rolled out for the first time, there is no prior information on census outcomes nor on bureaucrat performance that the tax office could be using to assign bureaucrats. There are on average 14 bureaucrats working in a given section with one supervisor who geographically assigns each bureaucrat to a set of plots on the map. A given property is visited by only one bureaucrat whose identification number is recorded in the application. In Section 4, we verify that bureaucrat characteristics are balanced across arms, and do not correlate with market values.

A plot visit lasts around ten to fifteen minutes. The bureaucrat first takes a picture and indicates whether the property is eligible for taxation.[30] Next, the detailed steps vary

[28] As a whole the broader experiment spans 83,360 plots. The 193 sections were selected by the administration based on tax potential, excluding informal settlements and industrial areas. The comparison between the census sections and the pure control sections will be addressed in a follow-up paper.

[29] There are eight tax offices covering the whole region. A plot is *not* eligible for property taxation if it is vacant, if it is a building totally under construction, if it belongs to the state, if it is a public school or religious institution.

[30] In Section 4 we verify that the probability of being covered by the census and of being classified eligible

by arm. In the 48 discretionary sections, the bureaucrat tries to speak to the owner and/or the tenants and asks about their identification details and about monthly rental values. If the bureaucrat is not able to ask occupants, or if the values provided seem unrealistic, she uses her own judgment to provide the best estimate of the monthly rental value. Rents paid by tenants and the value of owner-occupied parts are entered separately. In the 48 rule-based sections, the bureaucrat starts by entering eighteen observable characteristics of the property, visible from the outside. This takes around three minutes. These inputs are used to automatically generate a predicted property value for the tax roll, which is not visible on the tablet.[31] In these areas too, the bureaucrat tries to speak to occupants to recover their identification details, and writes down rental values if she recovers any, although these are not used to compute tax liabilities.

The algorithm used in the rule-based arm is inspired by Computer Assisted Mass Appraisal (CAMA) methods for property valuation.[32] The characteristics were selected with the administration; they are all visible from the outside, for the valuation to be possible even if occupants don't let bureaucrats in. They are reported on tablets either by ticking 'Yes' or 'No' boxes, or by selecting a modality in drop-down menus. For some of the characteristics, the answer requires making a judgment on quality, for example indicating whether the fence is in a 'Good', 'Medium' or 'Bad' state.

The property tax census first started in 2019 with seven sections covered between November 2019 and February 2020 (less than three percent of the plots in our sample). There was an interruption due to the Covid-19 pandemic. The census resumed in June 2021 and ended in April 2023.[33]

do not differ across arms.

[31] We provide details on the calibration of the algorithm in Section 4.2. The eighteen characteristics are listed in Table A3.

[32] CAMA methods are widespread in some developed countries such as the United States and Canada for instance, but only strongly established in Sub-Saharan Africa in the case of some South African cities. We worked with practitioners having made suggestions on how to adapt these methods to the African context Franzsen & McCluskey (2017); Davis *et al.* (2012); McCluskey *et al.* (2013); Guan *et al.* (2011); Moore (2005); Ali *et al.* (2018); Fish (2018); International Association of Assessing Officers (2022).

[33] The tax bills generated from the new census operations will be distributed all at once at the end of 2023.

3 Conceptual Framework

The objective of the administration is to register the maximum number of properties with values reflecting market rents ("accurate") in order to tax to the full potential, while ensuring tax fairness, more precisely vertical equity – which implies that more expensive properties should always face a higher tax bill than cheaper properties – and horizontal equity – which implies that two properties of similar market value should face similar tax bills. Overvaluation error is worse for low value properties, while undervaluation is worse for high value properties; how strongly this is true depends on the government's preference for vertical equity. Dispersion – large differences in valuations for a given market value – is problematic anywhere in the distribution; how strongly this is true depends on the government's preference for horizontal equity.

Bureaucrats' discretion in the field can have advantages: they may have local knowledge and they may be able to recover private information from occupants such as actual rents being paid. They may generate more equity if they adjust values to idiosyncratic information on the property or its owner that could not be captured in a systematized process. They may be better equipped to value some types of properties, for instance, within price ranges with which they are more familiar, or where occupants are more cooperative with the government.

On the other hand, discretionary bureaucrats could lack the required expertise to recover reliable values. They may be inconsistent in their valuations, they may be strongly heterogeneous, all of which would generate dispersion. They may also be biased towards or against certain types of occupants, and their values could intentionally or not be affected by their perceptions of what is fair. Finally, bureaucrats could have objectives that differ from those of the administration. There could be disutility associated with getting accurate values, if it implies additional effort, conflicts with occupants, or reduces the number of plots covered per day. Finally, they could engage in rent-seeking and try to collude with owners, asking for bribes in exchange for lower future tax liabilities.

A rule-based system is more systematic and predictable. Even with a rule implemented by bureaucrats, core components of the predictions (built area and location) do not depend on their input. This might allow the mitigation of the effects of heterogeneity and inconsistency. Collusion under the rule implemented by bureaucrats is less straightforward since it requires sophisticated manipulation by bureaucrats who are not aware of the weight of each characteristic.

Finally, a pure-rule process implemented using only data "from the office" removes all drawbacks due to delegation, but may lead to losses in accuracy if the characteristics collected in the field increase precision. There may also be a cost for the government of not collecting local information about occupants.

4 Data

4.1 Data Sources

Baseline property owner survey. Our baseline owner survey covered 2,474 plots, half being in the sections covered by the digital property tax census (spanning both arms) and half in the pure control sections. We collected socio-demographic information on owners, rental status and declared property values. The survey was conducted in 2018, before the beginning of the property tax census.

GIS mapping of plots. We compiled a geocoded dataset with plot identification number, plot area and ground built area measurement, comprising all 83,360 plots of the census and control sections.[34] Plot identifiers and delimitations were provided by the tax administration. Built area was recovered by GIS experts hired by the research team relying on high resolution images from the Pléiades satellites (50cm resolution) made available by the French Space Agency.[35]

Market rental values. We hired licensed assessors from the real estate section of the Senegalese National Order of Experts[36] to build a dataset of market values for a representative sample of properties. We prepared a random selection of 5,806 plots, half in the census sections (spanning both the discretionary and rule-based arms), and half in the pure

[34] This is the largest geocoded dataset with built area measurements and administrative plot identification made available to the Senegalese administration to date.

[35] *Centre National d'Etudes Spatiales* or CNES

[36] See: https://www.experts-ones.com/

control sections.[37] Figure A1 shows how our different samples overlap, and Appendix Section B.1 provides additional details on the data collection. Assessors went to see each property in the field and provided a market rental value, as well as an upper and lower bound. Properties were valued from the outside to avoid any biases due to non-response. Assessors also collected the observable property characteristics we used in the calibration of the rule. This work was carried out between June and August 2022, and resulted in a valuation for 4,921 properties.[38] Figure 1 shows the range of properties and their market values.

Property tax census data from the digital tool. Out of the 96 targeted sections, 94 (comprising 41,609 plots) were covered by the census.[39] 38,227 (91.9 percent) of these plots were registered by a bureaucrat. In Table A9, we show that implementation did not differ across arms.[40] For each plot covered by the census, we have the bureaucrat's identification number and all variables collected by the bureaucrat: identification details of owners and tenants, usage of the property, number of floors, numbe of rooms, rental status, main residence status, rent value, value of owner-occupied parts, whether the bureaucrat met the owner. The observable property characteristics used for the rule are only collected in rule-based sections. In Appendix B.3 we provide details on the data cleaning we apply to software extractions.[41] Henceforth, we refer to the property values from the census as census values or tax roll values.

[37] The daily fee of licensed assessors from the private sector is ten times the daily fee paid to field bureaucrats hired for the program. For this reason, it is not feasible to directly hire assessors to value all properties.

[38] In Appendix Table A1 we show correlations with values from other sources. Although the samples are small, we are reassured by the 0.62 correlation between assessor valuations and self-reported values from the owner baseline survey for properties that were at least partially rented out, and the 0.72 correlation with rents of properties with a rental contract recovered in the census. In Figure A3, we show comparisons with values from online rental listings. We can only situate these properties within a broad set of sections due to imprecise address information. The average values are typically similar to or higher than the averages for the top quintile of assessor values, which is expected since online postings tend to cover high-end properties.

[39] For two sections (one in each arm), the census was interrupted in the first days because of pre-existing tensions between the local population and the tax administration regarding property titles (11 out of 89 and 16 out of 404 plots were covered respectively). The associated tax bills will not be generated. We drop these two sections from our analysis sample.

[40] We regress the number of plots covered per day in a given section and find no significant difference across arms. In Panel (A) results are for the full sample, and in Panel (B) we restrict the sample to properties for which we have market values.

[41] This includes replacements we conduct for a subset of properties for which some observable characteristics were not collected following a technical problem. In Appendix C, we present two robustness checks to ensure that the replacement strategy is not affecting our results.

Analysis sample. Our main analysis sample is restricted to plots from the discretionary and rule-based arms that were not classified as non-eligible for taxation by bureaucrats[42] and for which we have both a market value from the assessors' dataset and built area measurement.[43] The analysis sample includes 2,290 plots – 1,166 in the rule-based arm and 1,124 in the discretionary arm – and 93 percent were covered by the census.[44]

Bureaucrat surveys. Bureaucrats filled in a short baseline survey during their training, with information on their background, their social preferences and views about taxation. The also completed an endline survey with several modules: satisfaction with the job, personality traits, socio-emotional skills, cognitive skills, persuasion exercise, tax knowledge. Finally, the endline survey included experimental lab-in-the-field property valuations we use to test mechanisms. We also surveyed supervisors and asked them to evaluate the bureaucrats on different dimensions. In Appendix Section B.5 we define the variables of the bureaucrat surveys. Out of a total of 268 bureaucrats, 247 completed the baseline survey, and 180 completed the endline.[45]

4.2 The Rule: Property Valuation Algorithm

We calibrate a property valuation model on the sample of all assessor valuations ($N = 4,921$, spanning both census arms as well as the pure control sections) using an elastic-net regression and five-fold cross-validation. The functional form is:

$$Ln(Value_{ij}) = \alpha + \beta Ln(BuiltArea_{ij}) + \gamma floors_{ij} + \sum_k \theta_k X_{k,ij} + Sec_j + \epsilon_{ij} \tag{1}$$

[42]We show that this is balanced across arms in Table A9 column (3).

[43]In Panel (A) of Table 1 we show that this is balanced across arms. Built area is missing for 7 percent of plots, this can be due to the plot being vacant land, or a mismatch between satellite images and plot borders for cases where there was a recent change.

[44]We show that this is balanced across arms in Table A9 column (2).

[45]Among the 38,227 plots covered by the census, 98.8 percent were covered by a bureaucrat for which we have a baseline survey, and 81.4 percent were covered by a bureaucrat for which we have an endline survey. The lower number of bureaucrats completing the endline survey is due to the fact that bureaucrats who stopped over the pandemic period were more difficult to track down.

where $Value_{ij}$ is the annual rental value of property i in section j, $BuiltArea$ is total built area (ground area multiplied by the number of floors), $floors$ is the number of floors, and the X_k variables are the property characteristics visible from the outside as reported by assessors. Sec_j is a section fixed effect.[46] For the main rule implemented in the application in the rule-based arm, we include the X_k covariates. Table A3 shows the resulting coefficients and Table A2 shows the performance statistics. See Appendix B.2 for additional details and references. We use the standard calibration indicators from the property valuation literature: the R^2 is 90 percent, the out-of-sample mean absolute percentage error (MAPE) is 33.8, and 59.6 percent of out-of-sample predicted values fall within 30 percent of the market value.[47][48] The application automatically computes predicted values by applying coefficients from the model to bureaucrats' inputs, and to pre-loaded section indicators and built area measurements (see Appendix B.2 for details).

The second rule, the pure rule, uses only covariates that could be recovered from the office: built area, number of floors, and section fixed-effect. Table A4 displays the performance statistics and summarizes the coefficients. The R^2 decreases to 0.87 and the MAPE increases to 41.4 percent. This suggests that section fixed-effects and built area explain a large share of the variation in property values, although the additional characteristics used in the main rule still add some valuable information.[49]

[46]The X_k covariates are: usage, type of fence, state of the fence, type of cladding, state of the cladding, cement wall, presence of decorative tiles, quality of doors and windows, landscape improvement, architectural improvement, presence and type of garage, balcony, location with respect to main road, type of road, presence of sidewalk, whether the property is at an angle, presence of street lights. All variables and response categories are listed in Table A3.

[47]We also test a random forest prediction model, the out-of-sample R^2 is 0.83 and the MAPE is 43. The best performing model we test is an elastic-net-regression estimated through cross-validation using all covariates and adding quadratic and cubic terms for built area and number of floors, as well as interactions between all section fixed-effects and built area. 65.6 percent of predictions fall within 30 percent of the market value. We do not implement this rule in the application because the gains in precision seem limited when compared to the increased complexity.

[48]Although our performance indicators are slightly lower than those observed in high-income contexts with superior data quality, they are high when compared to results from similar settings. Ali *et al.* (2018) find and R^2 around 56 percent in Rwanda, Franklin (2019) finds an R^2 of 85 percent in Addis Ababa, Ethiopia. Behr *et al.* (2023) find an MAPE ranging between 0.30 and 0.64 in South Africa.

[49]In the current version of the pure rule we still rely on the number of floors reported by assessors to compute total built area. Indeed, a formula calibrated using only the ground floor built area is much less precise (R^2 of 0.62 and MAPE of 99). However, it is technically feasible to recover the number of floors directly from more sophisticated satellite images. The Senegalese administration has started the acquisition of such images which should become available in 2024.

Additionally, we calibrate three rules used for robustness checks. The first is calibrated using values from our property owner baseline survey as the outcome variable ($N = 1,293$). The R^2 is 0.33 and the MAPE is 75.4. The second is a pure rule also calibrated on baseline values (only section fixed-effects and built area as regressors). The R^2 is 0.29 and the MAPE is 72. The third is calibrated using assessor values but restricting the sample to pure control sections ($N = 2,458$).[50] The R^2 is 0.88 and the MAPE is 51.6.

4.3 Balance Checks

In Table 1 we verify that section and bureaucrat characteristics are balanced across the rule-based and discretionary arms. Section characteristics are balanced thanks to our randomization, while the balance of bureaucrat characteristics results from their quasi-random assignment to areas. In Panel (A), we regress plot characteristics (from the cadastral data, the assessors' dataset, and the owner baseline survey) on a dummy taking value one for discretionary sections. In Panel (B), we regress baseline bureaucrat characteristics on a dummy taking value one for discretionary sections, observations are at the bureaucrat X section level. In Panel (C), we regress bureaucrat characteristics on market values, observations are at the plot level. We can rule out joint significance of all characteristics in the three Panels. None of the p-values for the coefficient of interest are below 0.05, except for one out thirteen of characteristics in Panel (C) but with a very small magnitude.[51] These verifications confirm that we can draw causal interpretations of the effects on the valuation roll of: (i) giving bureaucrats different degrees of discretion; (ii) bureaucrat characteristics.

5 Results

[50] We use strata fixed-effects instead of section fixed-effects in order to be able to apply this rule to our census sections. Additionally, we control for the value of a standardized 200 m^2 house, which was a section-level variable integrated in the assessor dataset.

[51] A property with a one percent higher market value is 0.04 percentage points (0.09 percent) less likely to be visited by a bureaucrat with three years or more of higher education.

5.1 Removing Discretion Increases Accuracy and Tax Equity

Valuation profile by arm. In Figure 4, we show the scatter plot of property values on the tax roll over market values. The first striking observation is the strong dispersion of valuations under discretion (Panel (A)). We add a 6th-degree polynomial fit curve with its 95 percent confidence interval. It reveals a significant undervaluation gap under discretion that increases with market value. These two problems seem to be mitigated by the rule-based process (Panel (B)), even more so by the pure rule (Panel (C)). In the remaining of this section we confirm and quantify these visual results.

First, we plot the median assessment ratio – computed as tax roll value over market value[52] – by quintiles of market value, separately for each valuation method. Results are shown in Figure 5, and in Table 2, we test whether the differences across quintiles are significant. There is a strong undervaluation gradient under discretion: the median assessment ratio is 0.83, 0.73, 0.50, 0.44 and 0.23 in quintiles one to five respectively, and these differences are significant at the 1 percent level (Panel (A) of Table 2).[53] Under the rule-based system, the ratio is in the vicinity of one for quintiles two to four, but significanly higher (1.25) in the bottom quintile and significantly lower (0.64) in the top quintile (Panel (B) of Table 2). The pure rule (Panel (C) of Figure 5) displays a median assessment ratio in the vicinity of one for all quintiles except the bottom one (1.26).[54] We provide additional results on the assessment ratios in Appendix Table A8.[55]

[52]The assessment ratio is a widely used indicator in the property tax literature (Avenancio-León & Howard, 2022; McMillen & Singh, 2020; Dray *et al.*, 2023).

[53]More precisely, the significant differences are the ones between quintiles one and two versus three and four, and between quintile four versus quintile five. In Table A5 we conduct the following robustness checks: considering mean assessment ratios instead of the median (Panel (A)), grouping properties by quintile of market value per square meter (Panel (B)), grouping properties by quintile of predicted market value using a prediction on pure control areas where there is no census.

[54]In Appendix Figure A5, we show the assessment ratio by quintile for our calibration sample with assessor inputs. It still displays a slightly higher (resp. lower) assessment ratio in the bottom (resp. top) quintile, but to a much lower extent than when the rule is implemented by bureaucrats. Some degree of regressivity is inherent to the rules because of unobserved variables (such as property features the assessors saw in the field but not captured in the characteristics, or area characteristics at a lower level than the section). This has been shown extensively for property tax valuation models applied in the United States (McMillen & Singh, 2020; Berry, 2021; Amornsiripanitch, 2023).

[55]Appendix Table A8 reports horizontal and vertical equity statistics that are commonly used in the property valuation literature (McMillen & Singh, 2020; International Association of Assessing Officers, 2013). The price-related differential (PRD) is computed as the ratio of average assessment ratio over the market value-weighted average assessment ratio. The larger it is, the weaker vertical equity is. As a benchmark, the International Association of Assessing Officers suggests that the PRD should lie between 0.98 and 1.03

Second, we measure the correlation between a property's rank in market values and its rank on the tax roll – a slope closer to one suggests stronger vertical equity.[56] In Figure 6, properties are sorted into 20 bins of market values on the x-axis, and the y-axis shows the mean rank for properties of a given bin. The rank-preservation slope is 0.28 under discretion, 0.69 under the rule, and 0.94 with the pure rule.

Finally, we plot the median effective tax rates by quintile in Figure 7, and add the benchmark tax profile obtained by applying the tax code to market values. Panels (B), (C) and (D) show that the standard deviation of the tax rate in each quintile increases with each degree of discretion, which harms horizontal equity. The expected tax rate in the bottom quintile is 4.4 percent. The observed tax rates are 3.8 percent in the discretionary arm, 7.4 percent in the rule-based arm, and 6.9 percent if applying the pure rule to the rule-based arm. The expected tax rate in the top quintile is 8.6 percent. The observed tax rates are 1.7 percent under full discretion, 5.2 percent under the rule and 7 percent under the pure rule. We provide additional results and robustness checks in Appendix Figure A7.[57]

Regression results for tax base gap outcomes. We define the tax base gap as the tax roll value minus market value. In the discretionary arm, the tax roll value is bureaucrat's discretionary value. In the rule-based arm, the tax roll value is subsequently the rule-based value and the pure rule value. The sign and value of the tax base gap indicate whether there is over- or undervaluation and by how much, in monetary amounts. We estimate:

$$Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk} \tag{2}$$

(International Association of Assessing Officers, 2013). We find that it is 1.60 under full discretion, 1.28 under the rule, and 1.02 using the pure rule. Second, we compute the coefficient of dispersion (COD), a measure of horizontal equity calculated as the average percentage deviation of the assessment ratio from its median. We display this statistic both overall and by quintile: dispersion in valuations is always the strongest under full discretion, and always the lowest under the pure rule.

[56] Methodologically, we follow the social mobility literature (Chetty *et al.*, 2019, 2020).

[57] Panel (A) of Figure A7 shows the *mean* tax rates by quintile, in Panel (B) we calculate the share of property value subject to the owner-occupied abatement using the number of rooms in both arms (instead of using the owner-occupied value directly provided by bureaucrats in the discretionary arm), in Panel (C) we plot the median tax rate by $deciles$, and in Panel (D) we restrict the sample to properties for which a positive valuation was made.

where Y_{ijk} is the outcome for plot i of section j and strata k, D_{jk} is a dummy taking value one for discretionary sections, and S_k is a strata fixed effect. Standard errors are clustered at the section level. Results are shown in Table 3, where Y_{ijk} is in turn the tax base gap (column (1)), the median tax base gap (column (2)), the absolute tax base gap (column (3)), and the assessment ratio (column (4)). Under the rule-based system, the mean (respectively, median) gap is -2.33 (resp. -0.16) millions FCFA, and the mean absolute tax base gap is 4.67 millions FCFA (Panel (B)).[58] The estimated $\hat{\beta}_{Discretion}$ shows that the gap is widened by -4.61, the median gap by -1.87, and the absolute tax base gap by 3.88 (or 83 percent) when switching to discretion, all coefficients being significant at the 1 percent level. In Panel (C), we run the same regression but applying the pure rule, and find larger $\hat{\beta}_{Discretion}$ coefficients: compared to the pure rule, discretion increases the absolute tax base gap by 4.71 million FCFA or 166 percent. In Appendix Table A6, we provide additional results and robustness checks.[59] In Tables A9 and A7, we decompose the effect of discretion into its extensive and intensive margins and carry out a Lee bounds exercise.[60]

Heterogeneity by market value. Next, we split the sample into low market value (quintiles one and two) and high market value (quintiles three to five) properties and re-estimate regression 2. Results are in Table 3. For low value properties, the rule-based process slightly overvalues while discretion slightly undervalues. The coefficient for the absolute tax base gap is still significant, but much smaller: discretion increases the tax

[58] All amounts are winsorized at the one percent level. The median annual rental market value is 15.80 millions FCFA.

[59] First, we add bureaucrat fixed-effects and find that the difference in tax base gap across arms is of the same order of magnitude within bureaucrat (Panel (A) of Table A6). Second, one might wonder whether part of the result is mechanical since the algorithm is calibrated on assessor values, and assessor values are also used as the benchmark. This problem is mitigated by the fact that the algorithm calibration relies on cross-validation, and also has half of its observations in pure control sections which do not overlap with our analysis sample. In addition, in Panels (B) and (C) of Table A6, we re-estimate regression 2 replacing values in the rule-based arm by predictions using two rules calibrated on self-reported owner baseline values. The firs rule uses all covariates and the second one is a pure rule (see Section 4.2). Discretion still significantly widens the median tax base gap even compared to these low quality rules (, results are noisier when the outcome is the absolute tax base gap, the coefficient is still positive but not significant).

[60] In Table A9, we find that discretion leads to a 17.8 percentage point lower probability of being assigned a positive value, and this is driven entirely by values of owner-occupied parts. In Table A7 we verify that the rule outperforms discretion in terms of accuracy on the intensive margin as well. In Panel (A), we estimate equation 2 on the sample of properties with a positive value, the results are indicative and not causal. In Panel (B), we carry out a Lee bounds exercise and find that discretion significantly widens the tax base gap even with extreme assumptions on the nature of selection into being assigned a positive value.

base gap by 25 percent, against 72 percent for high-value properties. The pure rule slightly outperforms discretion even for low value properties, but again, its strongest advantage compared to discretion is for the upper part of the distribution.

Tax liabilities. Translating these results into tax liabilities illustrates how massive the effects of different degrees of discretion are for the tax burden and its distribution. Total liabilities amount to 8 billion FCFA in the discretionary arm (16,651 eligible plots), against 19.6 billion if extrapolating the benchmark market values to all eligible plots; 11 billion FCFA in the rule-based arm (16,026 eligible plots), and 19.7 billion FCFA in this second arm if using the pure rule. Based on the subsample with market values, the share of liabilities due by the bottom 10 percent of properties is: 1.1, 0.95 and 0.64 percent respectively. The share due by the top 10 percent is: 49.6, 63 and 70 percent respectively.

5.2 Removing Discretion Reduces Bureaucrat-Induced Variability

Estimating bureaucrat fixed-effects. To measure bureaucrat-induced variability in the quality of valuations, we estimate bureaucrat fixed-effects in the following specification, separately on the discretionary and rule-based arms:

$$|Gap|_{ijb} = \alpha_b + Val_j + \epsilon_{ijb} \tag{3}$$

where $|Gap|_{ijb}$ is the absolute value of the tax base gap (in millions of FCFA and winsorized at the one percent level) for property i of section j covered by bureaucrat b, α_b is the bureaucrat fixed-effect, and Val_j is a section-level control for market values.[61] Errors are clustered at the section level. In the discretionary arm, there are 1,055 properties and 198 bureaucrats; in the rule-based arm, there are 1,063 properties and 190 bureaucrats.[62] Thanks to the quasi-random assignment of bureaucrats to plots, the estimated α_b are an unbiased measure of the quality of a bureaucrats' valuations: a larger α_b means bureaucrat b drives values further away from market values. The best performing bureaucrats are the ones with the smallest α_b. Next, we shrink each estimated fixed-effect proportionally to

[61] Val_j is a categorical variable indicating in which decile of market value per square meter section j is.

[62] Plots that are not covered by the census are dropped since no bureaucrat is associated to them. As shown in Table A9 column (2), the probability of being covered does not differ across arms.

the noise with which it is estimated using an empirical Bayes procedure, this yields the vector of adjusted fixed-effects $\alpha_{b,EB}$.[63] Finally, we measure the share of variance in the tax base gap explained by bureaucrats in each arm.[64] The kernel density estimates of the distribution of the fixed-effects are pictured in Panel (A) of Figure 8.[65] Summary statistics are reported in Table 4. The variance of the absolute tax base gap is much larger under discretion than under the rule (217.6, column (1), against 103.65, column (2)), but so is the variance of the estimated $\alpha_{b,EB}$, and to a larger extent. As a result, the share of variance in the tax base gap explained by bureaucrat fixed-effects is 40 percent under discretion, against 13 percent under the rule.[66]

Screening bureaucrats. Could discretion outperform the rule if the administration were able to successfully screen bureaucrats? First, we quantify how strongly the administration would need to screen for the difference in the tax base gap between rule-based and discretionary valuations to fade. In Panel (B) of Figure 8, we sort bureaucrats by their $\alpha_{b,EB}$ estimated in the discretionary arm. We run specification 2 starting by the sample with all bureaucrats, and iterating by removing the worst bureaucrats one by one. The coefficient on discretionary sections is no longer significant after removing the 81 (41 percent) worst bureaucrats, suggesting that restricting to top bureaucrats performs at best as well as the rule-based process.[67] Second, we investigate whether there are observable characteristics that predict being a top bureaucrat, defined as having a negative fixed-effect in the discretionary arm $\alpha_{b,EB} < 0$ (Figure A9). Having three years or more of

[63]Our estimates may suffer from noise due to the limited number of observations for each bureaucrat. For the shrinkage procedure, we follow the methodology developed in Chandra *et al.* (2016), in line with Kane & Staiger (2008); Morris (1983).

[64]We can compute the share of variance as $Var(\alpha_{b,EB})/Var(|Gap|)$ since quasi-random assignment leads to $Cov(\alpha_{b,EB}, Val_j) = 0$.

[65]Figure A8 plots the correlation between each bureaucrats' fixed-effect under the rule and under discretion.

[66]As a comparison, Bergeron *et al.* (2022) find that bureaucrats explain 21 percent of variance in tax compliance across neighborhoods in the Democratic Republic of Congo; Fenizia (2022) finds that managers of social security offices in Italy explain 9 percent of variation in productivity, measured by the efficiency in processing insurance claims; Best *et al.* (2023) find that bureaucrats explain 39 percent of variation in quality-adjusted prices in the Russian public procurement system. One reason why our results align with the highest among these findings is that in our setting, the outcome depends fully on what the bureaucrat is doing in the field. It does not depend for instance on the subsequent reactions of taxpayers, of other co-workers, etc.

[67]The estimation becomes noisier as we remove bureaucrats since the number of observations decreases and we may be underpowered to conclude with precision once we have removed a large share of bureaucrats.

higher education is significantly associated with a 20 pp (35 percent) higher probability of being a top bureaucrat, and is also a straightforward variable to use for screening.[68] Next, we use a k-means clustering procedure to divide bureaucrats into the two groups with the strongest possible difference in all characteristics, and find that one group has a 25 pp higher probability (significant at the 10 percent level) of being a top bureaucrat.[69] Finally, in Appendix D, we show how the tax profile differs for top versus bottom bureaucrats. First, we directly split the sample based on the bureaucrat fixed-effects. Next, we split the sample based on observables, first by long higher education (Figure A21), and subsequently based on the k-means clustering exercise (Figure A22), to see whether these screening procedures replicate the same difference in tax profiles obtained when directly using bureaucrat fixed-effects. In both cases, the predicted 'low performance' group generates lower tax rates, a more regressive profile, and much stronger dispersion.

6 Mechanisms

The role of the knowledge channel. A main driver of the strong inaccuracies and the undervaluation gradient found under discretion is bureaucrats' lack of expertise with respect to property values. We first shed light on this by identifying *bureaucrats' implicit algorithm*: we show that their values are poorly explained by objective property characteristics. We regress discretionary valuations on property characteristics, following the exact same methodology as the one used for our main algorithm. Results are shown in Appendix Table A10. The R^2 is 0.25, the elastic-net procedure assigns a value of zero to 15 out of 34 coefficients and 18 out of 48 section fixed-effects (against 3 out of 34 and 22 out of 193 in the main calibration), and the coefficient on built area 0.43 is (against 0.57).

Second, to isolate the knowledge channel, we implement a lab-in-the-field valuation

[68] 98 percent of bureaucrats have some higher education, 40 percent have three years or more. Other significant correlations show that a one standard deviation stronger baseline preference for widespread taxation is associated with a 11.4 percent higher probability of being a top bureaucrat (significant at the 10 percent level). Conversely, a one standard deviation higher emotions reading score (resp., agreeableness score) is associated with a 19 percent (resp., 15.7 percent) lower probability of being a top bureaucrat. See Appendix Section B.5 for a detailed description of the variables from the bureaucrat surveys, and see Appendix D for additional results on bureaucrats' skills and different measures of their performance.

[69] See Appendix Section D for details.

exercise. Bureaucrats are shown the picture of a property (with an indication of its neighborhood) and are asked to provide their best estimate of the monthly rental value. The exercise is done twice by each bureaucrat, for a low value property and a high value property.[70] The distribution of answers is plotted in Figure A15. The bottom property is given a value that is on average 73 percent higher than the true value, but still within the lowest quintile. It is accurately valued in 11 percent of cases, overvalued (respectively, undervalued) in 77 (resp., 15) percent of cases. On the other hand, the high value property is undervalued in 98 percent of cases, it is given an average value that is 70 percent lower than the true value, bringing it down from the highest to the third quintile of market values. We compare relative accuracy for the low versus high value property for a given bureaucrat by estimating

$$AR_{ib} = \alpha_b + \beta High_{ib} + \epsilon_{ib} \tag{4}$$

where AR_{ib} is the assessment ratio for property i and bureaucrat b, α_b is a bureaucrat fixed-effect, and $High_{ij}$ is a dummy taking value one for the high value property. Results are shown in Table 6, column (1), and show that that a given bureaucrat is 68 percent less accurate for the high value property. While very few bureaucrats are property owners (4.8 percent), those who are tenants (49 percent) report paying an average rent of 73,067 FCFA, which is in the bottom 5 percent of market values. This lack of exposure to expensive properties could explain the results. The dispersion of the answers is also telling: considering the high-end property, the hypothetical tax bill would vary between 83 and 4,158 USD depending on the bureaucrat.

Third, we use an information treatment to test whether there are avenues for learning. Half of the bureaucrats are shown a fact sheet providing key numbers from the distribution of market values (see Figure A14).[71] In column (2) of Table 6, we estimate regression

[70]The properties used in the experimental question are shown in Figure 1, Panels (A) and (C) respectively. These survey questions were not incentivized, they were presented as a means for the research team to better understand how the program was being implemented. We have no reason to believe that bureaucrats would strategically provide wrong answers, they were not aware that we had collected benchmark market values.

[71]The display of the fact sheet resembles the one used in Hoy (2022) to inform respondents about income distributions. Treatment assignment was randomized with a stratification on observed accuracy in the census data, gender, and education level. We verified bureaucrats' comprehension of the fact sheet using two

4 with a dummy for receiving the information treatment. We find that the information treatment does not improve valuations, neither for low nor high value property.[72] Our findings echo those in Hvidberg *et al.* (2023) showing that people tend to underestimate inequality, and Stantcheva (2021); Hoy (2022) showing individuals' misconceptions of income and wealth distributions.

Ruling out the collusion channel. An alternative mechanism for the undervaluation gradient could be collusion, if owners of expensive properties offer bribes in exchange of lower tax liabilities.[73] We rule this out using suggestive evidence. First, the lab-in-the-field finding proves that undervaluation is strong even when there are no possible gains from corruption. The hypothetical median effective tax rate based on bureaucrats' answers is 1.8 percent – strikingly close to the 1.7 percent found in the census for properties of the top quintile. Second, we test whether we find any difference in the undervaluation gradient in cases where the owner was met during the field visit, with the assumption that collusion would occur when the owner is met.[74] To do so we estimate:

$$AR_i = \alpha + \sum_{n=1}^{5} \beta_n Q(n)_i + \sum_{n=1}^{5} \gamma_n Q(n)_i \cdot M_i + \epsilon_i \tag{5}$$

where AR_i is the assessment ratio for property i, the $Q(n)$ dummies are indicators for

simple interpretation questions that needed to be answered correctly before moving to the valuation.

[72]We also test whether we observe learning over time in the census data. First, we verify how the absolute tax base gap and tax roll values evolve with the number of days and the number of properties covered by a bureaucrat (Table A12). We find that the tax base gap *increases* with the number of days under discretion, while it decreases under the rule. This suggests that improvements in collecting observable characteristics might be easier to achieve compared to improvements in bureaucrats' ability to recover market values in a fully discretionary way. Second, we check whether bureaucrats having already been exposed to the rule do better under discretion than non-exposed colleagues, which could occur if they learn which characteristics matter (Figure A17). We don't find any difference in the assessment ratio by quintile across the two groups.

[73]One feature of the setting which makes corruption less likely is that there is no direct exchange of money between the bureaucrats and the occupants. Furthermore, according to our owner baseline survey, the vast majority of owners (85 percent) never received a tax bill before. For them to pay bribes, they would need to trust that there will be enforcement later on.

[74]This variable being reported by bureaucrats, one could worry that it is manipulated. However, we expect this to be unlikely: (i) bureaucrats are not aware we are making comparisons with market values; (ii) bureaucrats are incentivized to meet owners and recover their identification details, their monthly bonus takes into account the share of owners for which they recover this information; (iii) bureaucrats' supervisor spends the day with them in the neighborhood, knows more or less where each team member is at any given point in time, and also reviews the forms in the evening before submitting them to the server – if a bureaucrat tends to spend time discussing with owners asking for bribes, and reports not meeting them, this would likely be detected by the supervisor.

each quintile of the distribution of market values, and M_i is a dummy taking value 1 if the bureaucrat met the owner. Errors are clustered at the section level. Results are plotted in Figure 9. We find that the interaction coefficients between each quintile and the dummy for meeting the owner are never significantly different from zero.[75]

Behavioral biases based on owner status and perceptions of fairness. We use our bureaucrat endline survey to show that bureaucrats are biased by what they consider fair. For each picture in the lab-in-the-field valuations, half of the bureaucrats are told that the owner is retired, while the other half are told that the owner is employed.[76] In the socioeconomic context of Senegal, retired people are considered as vulnerable and deserving support. We then estimate:

$$Ln(Value_{ibk}) = \alpha + \beta_1 Retired_{ibk} + \beta_2 High_{ibk} + \beta_3 Retired_{ibk} \cdot High_{ibk} + A_k + \epsilon_{ibk} \quad (6)$$

$Value_{ibk}$ is the value given by bureaucrat b for property i, $Retired_{ibk}$ is a dummy taking value 1 if the bureaucrat received the information that the owner of i is retired, $High_{ibk}$ is a dummy indicating the high value property, A_k is a fixed effect for the strata used for bureaucrat randomization. Results are shown in Table 6: bureaucrats provide a value that is 37.8 percent lower when the owner is retired (column (3)), and this is driven entirely by the low value property (column (4)). In Appendix Table A13, we test whether we find this correlation in the census data. Retired owner (compared to employed) correlates with a lower bureaucrat value, while there is no correlation with household income.[77]

[75] Additionally, although self-declarations are of course to be considered with caution, direct survey responses suggest that corruption is at most very rare. When asked whether they were "offered arrangements by owners" (without any reference to whether or not they accepted), 79 percent of bureaucrats replied *Never*, 20 percent *Once or Twice*. When asked whether this happened to their colleagues, 64 percent replied to *None*, 20 percent *Almost None* and 14 percent a *Minority*.

[76] This variation is randomized using the same stratification variables as for the information treatment, although both randomizations are independent.

[77] More precisely, we regress bureaucrat discretionary value on owner and owner X bureaucrat characteristics and control for property value, using the pure rule prediction. We use predicted value instead of market value because restricting the sample to observations for which we have market values would yield a too small number of observations for which we also have data from the baseline owner survey. In columns (1) and (2), we use the full sample of properties in the discretionary arm, and owner characteristics are reported by the bureaucrat. Meeting the owner, a deceased owner, a retired owner, a female owner (all relative to male owner) are correlated with a lower value, while multiple ownership and the area being the bureaucrat's home commune are correlated with a higher value. In columns (3) and (4), we rely instead on owner characteristics from our baseline survey.

Additionally, we rely on direct survey responses to illustrate how perceptions of fairness might affect valuations. When asked which direction of error is worse, 23 percent of bureaucrats consider that it is worse for bureaucrats to overvalue; while only 8 percent consider it is worse to undervalue.[78] Bureaucrats are almost symmetrically divided about whether it is fair for a retired person to pay a tax if she owns a property: 44.8 percent agree while 42 percent disagree.[79] When asked which types of owners can benefit from tax rebates, only 36 percent select owner-occupied properties (which is the most common true rebate); 46 percent select retired owners; 11 percent select single mothers - although there is absolutely nothing in the tax code for this category of owners. Finally, bureaucrats differ in their 'naiveness' towards owner-declarations: 17 percent reply that owner values are somewhat overvalued, 53 percent that they are about correct, and 29 percent that they are somewhat undervalued.

Big effects of a small degree of discretion. The tax base gap we observe with limited discretion, under the rule-based system, has two components: the first originates from prediction errors which do not depend on bureaucrats;[80] the second is caused by bureaucrats, if they enter erroneous property characteristics. This second component is informative on the effects of a limited degree of discretion, and we measure it by comparing predictions computed using assessor inputs with predictions relying on bureaucrats' inputs.[81] We estimate:

$$Y_{irjk} = \alpha + \beta RuleBur_{irjk} + S_k + \epsilon_{irjk} \tag{7}$$

where Y_{itjk} is the outcome for plot i of section j and strata k under rule r, $RuleBur_{irjk}$ is a dummy taking value one if r is the rule as implemented by bureaucrats, and zero if

[78]Importantly, this topic is never mentioned in the training nor by the supervisors, therefore the asymmetry really reflects bureaucrats' individual perceptions. There are 70 percent who consider that overvaluation and undervaluation are equally problematic.

[79]In the tax code, retired civil servants with a formal pension may benefit from a reduction in their tax bill. This applies only to a minority if owners in the region.

[80]As shown in the calibration statistics in Appendix Table A2, the share of predictions that fall within 30 percent of the market value in the test sample is 59.6 percent.

[81]To support our assumption that assessors' characteristics are the correct ones, we conduct verifications a random subset of 100 pictures and find that assessors' reported number of floors is 2.2 times more likely to be correct than bureaucrats.

r is the benchmark rule. In Panel (A), the benchmark rule is the pure rule, and in Panel (B) the benchmark rule is the rule with assessor inputs. Each property appears twice in the dataset, and the discretionary arm is excluded. S_k is a strata fixed effect and standard errors are clustered at the section level. Results are shown in Table 7. The rule with bureaucrat inputs increases the absolute tax base gap by 1.84 million FCFA or 65 percent compared to the pure rule, and by 1.89 million FCFA or 68 percent compared to the rule with assessor inputs (column (3)).[82][83]

What makes a successful field visit? The role of private information. In Table 5, we assess whether top bureaucrats behave differently during field work (these results are suggestive and not causal). We find that compared to other bureaucrats, they are: 12.9 percentage points (22 percent) more likely to report a positive value (column (1)), 4.5 pp (19.6 percent) more likely to report meeting the owner (column (2)), 5.6 pp (18 percent) more likely to recover owner identification details,[84] and 10.8 pp (22 percent) more likely to leave a comment. Conditional on leaving a comment, they are 5.8 pp (32 percent) more likely to indicate that they used their own estimation for valuation instead of relying solely on what occupants said.

Next, we look at whether the difference in tax base gap across rule-based and discretionary valuations closes when restricting the sample in turn to specific types of field visits. Results are shown in Table A14, both with and without bureaucrat fixed-effects. For cases where the owner is met (Panel (A)) and where the property is rented at least in

[82] Furthermore, we split the sample by market value. Looking at Panel (B), we see that the rule by bureaucrats slightly increases overvaluation at the bottom (although this is not significant for all outcomes), while significantly driving values down in the upper part of the distribution (the median gap decreases from -0.33 to -1.28 million FCFA). Thus even *partial* delegation generates more regressivity in the tax profile than what would occur with a correctly implemented rule.

[83] In Appendix Figure A12, we show the contribution of each characteristic X_k to the aggregate differences between the rule implemented by bureaucrats versus assessors measured in Panel (B) of Table 7 as $\sum_i |\gamma_k X_{k,RuleBur,i} - \gamma_k X_{k,RuleAss,i}|$, where subscript $RuleBur$ indicates the value taken by X_k when entered by bureaucrat and $RuleAss$ the value taken by X_k when entered by assessors for property i, γ_k is the coefficient for X_k in the formula. We find that the most contributing characteristic is cladding type. This is likely due to the fact that (i) it has six modalities, (ii) it takes some technical expertise to differentiate them. The second most contributing characteristic is area, which originates from a difference in the number of floors entered. When the number of floors differs, in 56 percent of cases this is because the bureaucrat counted the ground floor as one instead of zero. In Appendix Table A11, we provide descriptive statistics on the share of observations for which each characteristic matches across the two sources

[84] This dummy takes value one if the bureaucrat recovered the name and/or national ID number of the property owner.

part (Panel (B)) – the two most obvious instances where the bureaucrat accessed valuable private information – we further split the sample into low and high value properties.[85] We find that the difference in tax base gap across rule and discretion is no longer significant (and also displays very small coefficient sizes) in these two instances, for low-value properties. This suggests that bureaucrats are able to leverage valuable local information, but this only proves helpful in the lower part of the distribution.

7 Optimal Policy

Cost-benefit analysis. In Table A15, we lay out the costs and total tax liabilities, assuming each method in turn is applied to all eligible plots ($N = 32,677$). There are two types of costs to consider: field costs – 118.1 million FCFA[86] – and rule-specific costs – 16 million FCFA[87]. We abstract from fixed program costs that are common across all methods.[88] The the liabilities-to-costs ratios vary immensely: x133 under full discretion, x166 with the rule-based system, and x2416 with the pure rule.

Optimal policy. The pure rule is the optimal policy if the government wishes to maximize overall accuracy.[89] However, since the pure rule tends to overvalue the cheapest properties, and since this segment is also precisely the one where bureaucrats' valuations are relatively more accurate, there is a trade-off. If the government wants to maximize accuracy, while minimizing the risk of overvaluations at the bottom, the optimal policy is to predict which properties belong to the lowest quintile based on location and built area

[85]In Figure A13 we plot the probability of meeting the owner (Panel (A)) and of the property being rented (Panel (B)) by quintile of market value.

[86]Corresponding to 8,560 bureaucrats-days with a 5,000 FCFA daily fee, and 10,040 supervisors-days with a 7,500 FCFA daily fee.

[87]Corresponding to the hiring of licensed assessors for the calibration sample, and to what is paid to GIS experts recovering built area measurement from satellite images.

[88]These include: office managers, software development, data server, mobile data, training costs. This means the tax bill-to cost ratios presented in this section are only meaningful in relative terms.

[89]Sending bureaucrats in the field may still prove useful to get identification details on owners and tenants. One might be worried of costs in terms of job satisfaction and motivation of removing delegation for the determination of the tax base, but survey responses do not seem to suggest this. Among bureaucrats who did only the rule, 98 percent would participate in similar operations again, and 83 percent feel they have autonomy overall (these shares are similar when considering all bureaucrats). According to bureaucrats, the best strategy to recover accurate market values at scale is *Declared by tenant* (ranked first) and *With an automatic formula* (ranked second), *Bureaucrat's own knowledge* is ranked the lowest. See Figure A16.

(the assumption being that the administration has no prior knowledge on values), send bureaucrats for discretionary valuations of these cases, and apply the pure rule elsewhere. In Figure A18, we plot the resulting tax profile.[90]

Policy uptake. Following the results from the randomized experiment, the administration has decided to incorporate rule-based methods in the general functioning of the property tax, and has asked the research team for support.[91]

8 Conclusion

We introduce experimental variation in the degree of discretion bureaucrats have to value properties in a large-scale digitized property tax census conducted in Dakar by the Senegalese tax administration. Under full discretion, bureaucrats' property valuations are significantly below market values. Bureaucrats also generate a regressive tax profile by undervaluing expensive properties to a higher extent, which harms vertical equity. Finally, their valuations display strong dispersion harming horizontal equity. Even with partial delegation under a rule-based system, which relies on an algorithm incorporating bureaucrat inputs, the valuation profile is distorted and more regressive than with a pure rule.

We investigate the mechanisms of discretionary valuations, and show that bureaucrats' lack of knowledge plays a fundamental role. We also find that bureaucrats are biased by their perceptions of fairness. We use suggestive evidence to rule out collusion as an important driver of the results. Long higher education is the only characteristic that the government could easily use for selection and which correlates with bureaucrats' relative ability to approximate market values. However, at best, top bureaucrats perform as well as the rule. Overall, a pure rule is the most promising strategy for an equitable expansion of the tax net. The administration would only want to keep some discretion for low-value properties, if its preference for minimizing the risk of overvaluation is stronger than its

[90]There are two caveats. First, the prediction of being in the lowest quintile yields 18 percent of false negatives – these properties will be overtaxed. Second, using discretion will necessarily imply more dispersion and lower horizontal equity in that segment of the market.

[91]Investments that the administration should consider are: (i) working with licensed real estate assessors to enlarge the calibration sample to the whole region; (ii) high quality satellite images to recover built area measurement for all properties.

preference for horizontal equity.

Following the results from our randomized experiment, the administration has asked the research team for support to expand the use of rule-based methods to the whole region. Our findings also shed light on directions for future research. There are potentially many other promising applications of digitization for increased fiscal capacity in Africa (Okunogbe & Santoro, 2023). Second, whether the algorithm generates sufficient levels of tax compliance and political acceptability in the long run remains an open question, which we will address in a follow-up paper. Finally, it seems important to assess whether our finding that under a status quo discretionary process, bureaucrats shape the tax register in a regressive way extend to other examples of public policies, and how this can be mitigated beyond our setting.

References

AGARWAL, NIKHIL, MOEHRING, ALEX, RAJPURKAR, PRANAV, & SALZ, TOBIAS. 2023. Combining Human Expertise with Artificial Intelligence: Experimental Evidence from Radiology. *National Bureau of Economic Research Working Paper 31422*, July. 17

AGHION, PHILIPPE, & TIROLE, JEAN. 1997. Formal and Real Authority in Organizations. *Journal of Political Economy*, **105**(1), 1–29. 1, 1.1

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. 32, 48, B.2

AMORNSIRIPANITCH, NATEE. 2023. Why are Residential Property Tax Rates Regressive? *Working Paper*. 12, 54

ANAGOL, SANTOSH, BALASUBRAMANIAM, VIMAL, RAMADORAI, TARUN, & UETTWILLER, ANTOINE. 2022. A Bad Bunch: Asset Value Under-Reporting in the Mumbai Real Estate Market. *SSRN working paper*, December. 20

AVENANCIO-LEÓN, CARLOS F, & HOWARD, TROUP. 2022. The Assessment Gap: Racial Inequalities in Property Taxation. *The Quarterly Journal of Economics*, **137**(3), 1383–1434. 1.1, 52

BACHAS, PIERRE, BROCKMEYER, ANNE, FERREIRA, ALIPIO, & SARR, BASSIROU. 2021. How to Target Enforcement at Scale? Evidence from Tax Audits in Senegal. *EDI Working Paper*. 16

BALAN, PABLO, BERGERON, AUGUSTIN, TOUREK, GABRIEL, & WEIGEL, JONATHAN. 2022. Local Elites as State Capacity: How City Chiefs use Local Information to increase Tax Compliance in the Democratic Republic of the Congo. *American Economic Review*, **112**, 762–97. 1.1

BANDIERA, ORIANA, PRAT, ANDREA, & VALLETTI, TOMMASO. 2009. Active and Passive Waste in Government Spending: Evidence from a Policy Experiment. *American Economic Review*, **99**(4), 1278–1308. 1.1

BANDIERA, ORIANA, BEST, MICHAEL, KHAN, ADNAN, & PRAT, ANDREA. 2021. The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats. *The Quarterly Journal of Economics*, **136**(4), 2195–2242. 1.1

BANDIERA, ORIANA, BURGESS, ROBIN, DESERRANNO, ERIKA, MOREL, RICARDO, SULAIMAN, MUNSHI, & RASUL, IMRAN. 2023. Social Incentives, Delivery Agents, and the Effectiveness of Development Interventions. *Journal of Political Economy Microeconomics*, **1**(1), 162–224. 1, 19

BANERJEE, ABHIJIT, NIEHAUS, PAUL, & SURI, TAVNEET. 2019. Universal Basic Income in the Developing World. *Annual Review of Economics*, **11**(1), 959–983. 1

BANERJEE, ABHIJIT, DUFLO, ESTHER, IMBERT, CLÉMENT, MATHEW, SANTHOSH, & PANDE, ROHINI. 2020. E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India. *American Economic Journal: Applied Economics*, **12**(4), 39–72. 1

BANERJEE, ABHIJIT, HANNA, REMA, OLKEN, BENJAMIN A., SATRIAWAN, ELAN, & SUMARTO, SUDARNO. 2023. Electronic Food Vouchers: Evidence from an At-Scale Experiment in Indonesia. *American Economic Review*, **113**(2), 514–47. 1

BATTAGLINI, MARCO, GUISO, LUIGI, LACAVA, CHIARA, MILLER, DOUGLAS L, & PATACCHINI, ELEONORA. 2022. Refining Public Policies with Machine Learning: The Case of Tax Auditing. *National Bureau of Economic Research Working Paper 30777*, December. 15

BEHR, DANIELA, CHEN, LIXUE, GOEL, ANKITA, HAIDER, KHONDOKER, SINGH, SANDEEP, & ZAMAN, ASAD. 2023. Estimating House Prices in Emerging Markets and Developing Economies: A Big Data Approach. *World Bank Working Paper*. 5, 2.1, 48

BERGERON, AUGUSTIN, BESSONE, PEDRO, KABEYA, JOHN, TOUREK, GABRIEL, & WEIGEL, JONATHAN. 2022. Optimal Assignment of Bureaucrats: Evidence from Randomly Assigned Tax Collectors in the DRC. 1.1, 66

BERGERON, AUGUSTIN, TOUREK, GABRIEL, & WEIGEL, JONATHAN. 2023. The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC. *Working Paper*. 20

BERRY, CHRISTOPHER. 2021. Reassessing the Property Tax. *Working Paper*. 12, 54

BESLEY, TIMOTHY, & PERSSON, TORSTEN. 2009. The Origins of State Capacity: Property Rights, Taxation, and Politics. *American Economic Review*, **99**(4), 1218–44. 1.1

BEST, MICHAEL CARLOS, HJORT, JONAS, & SZAKONYI, DAVID. 2023. Individuals and Organizations as Sources of State Effectiveness. *American Economic Review*, **113**(8), 2121–67. 1.1, 66

BJÖRKEGREN, DANIEL, BLUMENSTOCK, JOSHUA, & KNIGHT, SAMSUN. 2022. Manipulation-Proof Machine Learning. *mimeo*. 1.1

BLACK, EMILY, ELZAYN, HADI, CHOULDECHOVA, ALEXANDRA, GOLDIN, JACOB, & HO, DANIEL. 2022. Algorithmic Fairness and Vertical Equity: Income Fairness with IRS Tax Audit Models. *Working Paper in ACM Conference on Fairness, Accountability and Transparency*, June. 15

BOWLES, JEREMY. 2023. Identifying the Rich: Registration, Taxation, and Access to the State in Tanzania. *American Political Science Review*, 1–17. 1

BROCKMEYER, ANNE, ESTEFAN, ALEJANDRO, SERRATO, JUAN CARLOS SUÁREZ, & ARRAS, KARINA RAMIREZ. 2021. Taxing Property in Developing Countries: Theory and Evidence from Mexico. *NBER Working Paper*. 20

BROWNE, OLIVER, GAZZE, LUDOVICA, GREENSTONE, MICHAEL, & ROSTAPSHOVA, OLGA. 2023. Man vs. Machine: Technological Promise and Political Limits of Automated Regulation Enforcement. *National Bureau of Economic Research Working Paper 30816*, January. 17

CASABURI, LORENZO, & TROIANO, UGO. 2015. Ghost-House Busters: The Electoral Response to a Large Anti–Tax Evasion Program. *The Quarterly Journal of Economics*, **131**(1), 273–314. 15

CHANDRA, AMITABH, FINKELSTEIN, AMY, SACARNY, ADAM, & SYVERSON, CHAD. 2016. Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector. *American Economic Review*, **106**(8), 2110–44. 63

CHETTY, RAJ, HENDREN, NATHANIEL, JONES, MAGGIE R, & PORTER, SONYA R. 2019. Race and Economic Opportunity in the United States: an Intergenerational Perspective*.

The Quarterly Journal of Economics, **135**(2), 711–783. 56

Chetty, Raj, Friedman, John N, Saez, Emmanuel, Turner, Nicholas, & Yagan, Danny. 2020. Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics*, **135**(3), 1567–1633. 56

Chioda, Laura, Contreras-Loya, David, Gertler, Paul, & Carney, Dana. 2021. Making Entrepreneurs: Returns to Training Youth in Hard versus Soft Business Skills. *NBER Working Paper 28845*, May. B.5

Christensen, Darin, & Garfias, Francisco. 2021. The Politics of Property Taxation: Fiscal Infrastructure and Electoral Incentives in Brazil. *The Journal of Politics*, **83**(4), 1399–1416. 1.1

D'Arcy, Michelle, & Nistotskaya, Marina. 2018. The early modern origins of contemporary European tax outcomes. *European Journal of Political Research*, **57**(1), 47–67. 1.1

Davis, Peadar, McCluskey, William, Grissom, Terry, & McCord, Michael. 2012. An empirical analysis of simplified valuation approaches for residential property tax purposes. *Property Management*, **30**(3), 232–254. 32, B.2

Decarolis, Francesco, Fisman, Raymond, Pinotti, Paolo, & Vannutelli, Silvia. 2021. Rules, Discretion, and Corruption in Procurement: Evidence from Italian Government Contracting. *Working Paper*. 1.1

Delbridge, Victoria, Sarr, Khady Dia, Harman, Oliver, Haas, Astrid, & Venables, Tony. 2022. Enhancing the financial position of cities: evidence from Dakar. *IGC Financing Sustainable Urban Development Case Study 3*. 10

Dessein, Wouter. 2002. Authority and Communication in Organizations. *Review of Economic Studies*, **69**, 811–838. 1.1

Dodell-Feder, David, Ressler, Kerry, & Germine, Laura. 2020. Social cognition or social class and culture? On the interpretation of differences in social cognitive performance. *Psychol Med*, **50**(1), 133–145. B.5

Dray, Sacha, Landais, Camille, & Stantcheva, Stefanie. 2023. Wealth and Property Taxation in the United States. *National Bureau of Economic Research Working Paper 31080*, March. 1.1, 52

Duflo, Esther, Greenstone, Michael, & Ryan, Nicholas. 2013. Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India. *The Quarterly Journal of Economics*, **128**(4), 1499–1545. 1.1

Duflo, Esther, Greenstone, Michael, Pande, Rohini, & Ryan, Nicholas. 2018. The Value of Discretion in the Enforcement of Regulation: Experimental Evidence and Structural Estimates from Environmental Inspections in India. *Econometrica*, **86**(6), 2123–2160. 1, 1.1

Dzansi, James, Jensen, Anders, Lagakos, David, & Telli, Henry. 2022. Technology and Local State Capacity: Evidence from Ghana. *NBER working paper29923*. 1, 1.1, 27

Elzayn, Hadi, Smith, Evelyn, Hertz, Thomas, Ramesh, Arun, Goldin, Jacob, Ho, Daniel, & Fisher, Robin. 2023. Measuring and Mitigating Racial Disparities in Tax Audits. *Stanford Institute for Economic Policy Research Working Paper*. 1.1

FENIZIA, ALESSANDRA. 2022. Managers and Productivity in the Public Sector. *Econometrica*, **90**(3), 1063–1084. 1.1, 66

FISH, PAUL. 2018. Practical Guidance Note: Training Manual for Implementing Property Tax Reform with a Points-Based Valuation. ICTD African Tax Administration Paper 2. 32, B.2

FRANKLIN, SIMON. 2019. The demand for government housing: evidence from lotteries for 200,000 homes in Ethiopia. *mimeo*. 48

FRANZSEN, RIEL, & MCCLUSKEY, WILLIAM. 2017. *Property Tax in Africa: Status, Challenges and Prospects*. Lincoln Institute of Land Policy, Cambridge, Massachusetts. 2, 2.1, 32, B.2

GORDON, ROGER. 2017. Rules versus discretion in tax policy. *Vox Dev*. 1.1

GREENSTONE, MICHAEL, HE, GUOJUN, JIA, RUIXUE, & LIU, TONG. 2022. Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution. *American Economic Review: Insights*, **4**(1), 54–70. 1.1

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. 32, B.2

HANNA, REMA, & OLKEN, BENJAMIN A. 2018. Universal Basic Incomes versus Targeted Transfers: Anti-Poverty Programs in Developing Countries. *Journal of Economic Perspectives*, **32**(4), 201–26. 1

HASEEB, MUHAMMAD, & VYBORNY, KATE. 2022. Data, discretion and institutional capacity: Evidence from cash transfers in Pakistan. *Journal of Public Economics*, **206**, 104535. 1.1

HOY, CHRISTOPHER. 2022. How Does the Progressivity of Taxes and Government Transfers Impact People's Willingness to Pay Tax? Experimental Evidence across Developing Countries. *World Bank*. 1.1, 6, 71

HVIDBERG, KRISTOFFER B, KREINER, CLAUS T, & STANTCHEVA, STEFANIE. 2023. Social Positions and Fairness Views on Inequality. *The Review of Economic Studies*, 03, rdad019. 1.1, 6

INTERNATIONAL ASSOCIATION OF ASSESSING OFFICERS. 2013. Standard on Ratio Studies. A criterion for measuring fariness, quality, equity and accuracy. 55

INTERNATIONAL ASSOCIATION OF ASSESSING OFFICERS. 2022. A Review of the Methods, Applications and Challenges of Adopting Artificial Intelligence in the Property Assessment Office. 32, B.2

KALA, NAMRATA. 2023. The Impacts of Managerial Autonomy on Firm Outcomes. *MIT Sloan Working Paper 6004-19*. 1.1

KANE, THOMAS J, & STAIGER, DOUGLAS O. 2008. Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation. *National Bureau of Economic Research Working Paper 14607*, December. 63

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. 1.1

KHAN, ADNAN Q., KHWAJA, ASIM I., & OLKEN, BENJAMIN A. 2019. Making Moves

Matter: Experimental Evidence on Incentivizing Bureaucrats through Performance-Based Postings. *American Economic Review*, **109**(1), 237–70. 1.1

KLEINBERG, JON, LAKKARAJU, HIMABINDU, LESKOVEC, JURE, LUDWIG, JENS, & MULLAINATHAN, SENDHIL. 2018. Human Decisions and Machine Predictions*. *The Quarterly Journal of Economics*, **133**(1), 237–293. 1.1

KNEBELMANN, JUSTINE. 2021. The (Un)Hidden Wealth of the City: Property Taxation under Weak Enforcement in Senegal. *mimeo*. 24

KNEBELMANN, JUSTINE. 2022. Digitalisation of property taxation in developing countries: Recent advances and remaining challenges. *ODI Report*. 3, 27

LEE, MELISSA M., & ZHANG, NAN. 2017. Legibility and the Informational Foundations of State Capacity. *The Journal of Politics*, **79**(1), 118–132. 1

LEVITT, STEVEN D., & SYVERSON, CHAD. 2008. Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions. *The Review of Economics and Statistics*, **90**(4), 599–611. 20

LIMODIO, NICOLA. 2021. Bureaucrat Allocation in the Public Sector: Evidence from the World Bank. *The Economic Journal*, **131**(639), 3012–3040. 1.1

MANWARING, PRIYA, & REGAN, TANNER. 2023. Public Disclosure and Tax Compliance: Evidence from Uganda. Working Paper. 25

MARTÍNEZ, LUIS R. 2023. Natural Resource Rents, Local Taxes, and Government Performance: Evidence from Colombia. *The Review of Economics and Statistics*, 04, 1–28. 1.1

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. 32, B.2

MCMILLEN, DANIEL, & SINGH, RUCHI. 2020. Assessment Regressivity and Property Taxation. *Journal of Real Estate Finance and Economics*, **60**(1), 155–169. 12, 1.1, 52, 54, 55

MOORE, WAYNE. 2005. Performance Comparison of Automated Valuation Models. *Journal of Property Tax Assessment and Administration*, **3**(1), 43–60. 32, B.2

MORRIS, CARL N. 1983. Parametric Empirical Bayes Inference: Theory and Applications. *Journal of the American Statistical Association*, **78**(381), 47–55. 63

MURALIDHARAN, KARTHIK, NIEHAUS, PAUL, & SUKHTANKAR, SANDIP. 2016. Building State Capacity: Evidence from Biometric Smartcards in India. *American Economic Review*, **106**(10), 2895–2929. 1

NIEHAUS, PAUL, ATANASSOVA, ANTONIA, BERTRAND, MARIANNE, & MULLAINATHAN, SENDHIL. 2013. Targeting with Agents. *American Economic Journal: Economic Policy*, **5**(1), 206–38. 1, 1.1

OKUNOGBE, OYEBOLA. 2021. Becoming Legible to the State. The Role of Detection and Enforcement Capacity in Tax Compliance. *World Bank Policy Research Working Paper 9852*. 1.1, 25

OKUNOGBE, OYEBOLA, & POULIQUEN, VICTOR. 2022. Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing. *American Economic Journal: Economic Policy*, **14**. 16

OKUNOGBE, OYEBOLA, & SANTORO, FABRIZIO. 2023. Increasing Tax Collection in African

Countries: The Role of Information Technology. *Journal of African Economies*, **32**(Supplement 1), 57–83. 1.1, 8

Plaisant, Odile. 2008. Validation par Analyse Factorielle du Big Five Inventory français (BFI-Fr) Analyse Convergente avec le NEO-PI-R. *Annales medio-psychologiques*. B.5

Pomeranz, Dina. 2015. No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, **105**(8), 2539–69. 1

Rogger, Daniel, & Somani, Ravi. 2023. Hierarchy and Information. *Journal of Public Economics*, **219**, 104823. 1, 1.1

Sadka, Joyce, Seira, Enrique, & Woodruff, Christopher. 2018. Information and Bargaining through Agents: Experimental Evidence from Mexico's Labor Courts. *NBER Working paper No. 251237*, October. 17

Scott, James. 1999. *Seeing Like a State: How Certain Schemes to Improve the Human Condition have Failed.* Yale University Press. 1

Stantcheva, Stefanie. 2021. Understanding Tax Policy: How do People Reason?*. *The Quarterly Journal of Economics*, **136**(4), 2309–2369. 1.1, 6

vom Hau, Matthias, Peres-Cajías, José Alejandro, & Soifer, Hillel David. 2023. No taxation without informational foundation: on the role of legibility in tax state development. *Journal of Institutional Economics*, **19**(3), 426–443. 1.1

Weidmann, Ben, & Deming, David. 2021. Team Players: How Social Skills Improve Team Performance. *Econometrica*, **89**(6), 2637–2657. B.5

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*. 1.1

Figures

FIGURE 1
EXAMPLES OF LOW, MEDIUM AND HIGH VALUE PROPERTIES IN STUDY AREAS

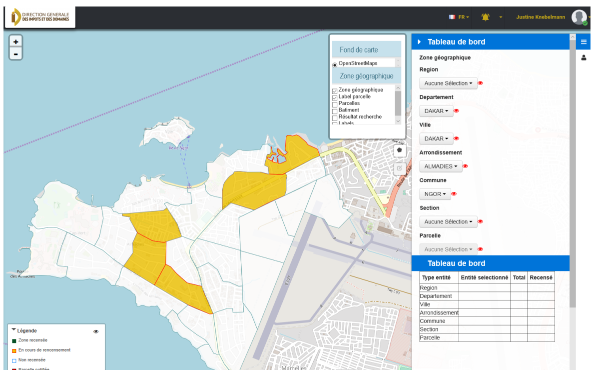
(A) LOW-VALUE PROPERTY

(B) MEDIUM VALUE PROPERTY

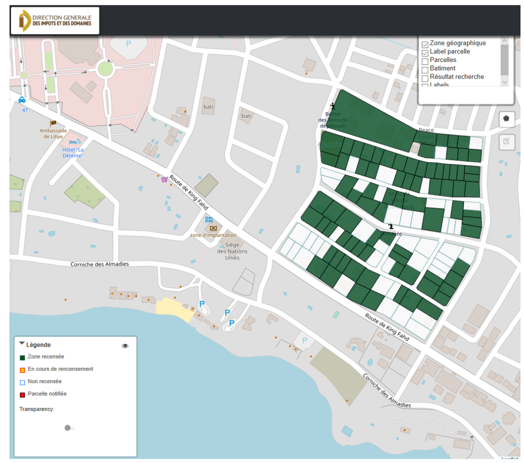
(C) HIGH VALUE PROPERTY

Notes: This Figure displays examples of properties from the study areas in the region of Dakar. Picture (A) shows a property from the bottom 10% of the property value distribution of our sample, Picture (B) shows a property with a value around the median of the distribution, and Picture (C) shows a property from the top 10% of the distribution. The monthly market rental values for each property are, respectively, 100,000 FCFA (163 USD); 520,000 FCFA (845 USD) and 2.3 million FCFA (3,740 USD). Source: valuations by licensed real estate assessor.

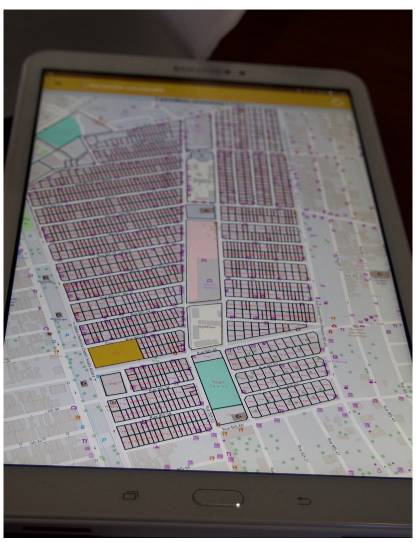
FIGURE 2
PROPERTY TAX CENSUS OPERATIONS WITH THE NEW DIGITAL TOOL



(A) WEB INTERFACE: ASSIGNMENT OF SECTIONS TO BUREAUCRATS



(B) WEB INTERFACE: PROGRESS WITHIN A SECTION

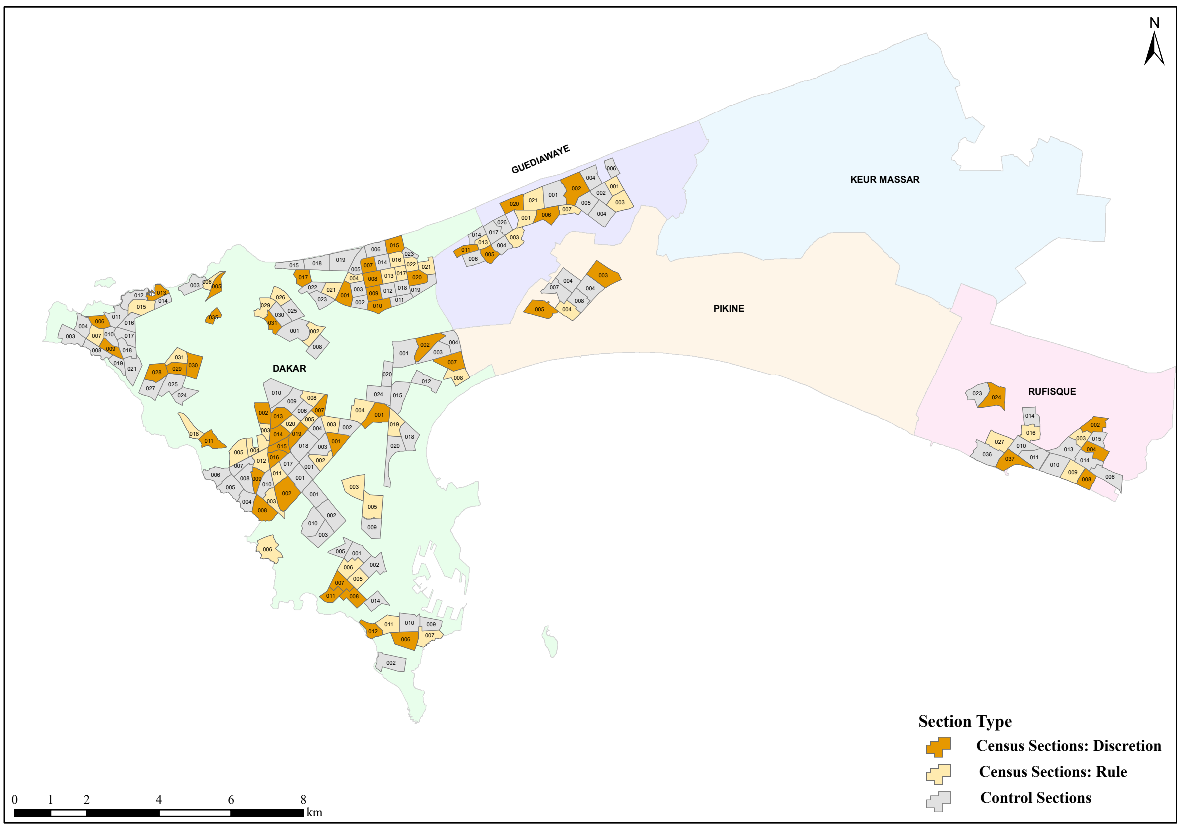


(C) TABLET INTERFACE: PRE-LOADED SECTION AND PLOT IDENTIFIERS

(D) BUREAUCRATS CONDUCTING PROPERTY TAX CENSUS

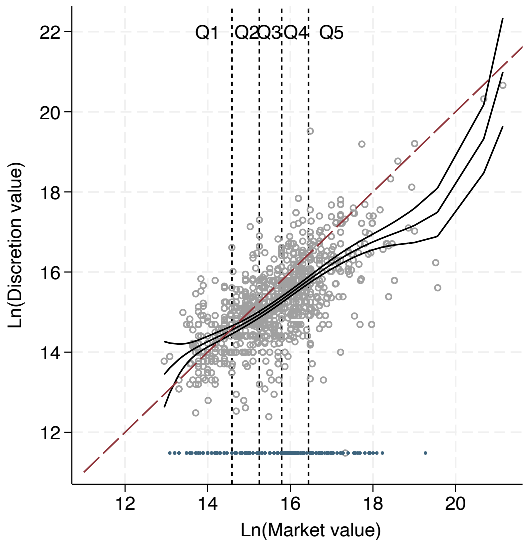
Notes: This Figure illustrates the property tax census operations using the new digital tool. The digital tool has a Web interface for tax office managers, and a tablet interface used by bureaucrats in the field.

FIGURE 3
GEOGRAPHICAL SCOPE AND EXPERIMENTAL DESIGN IN DAKAR REGION

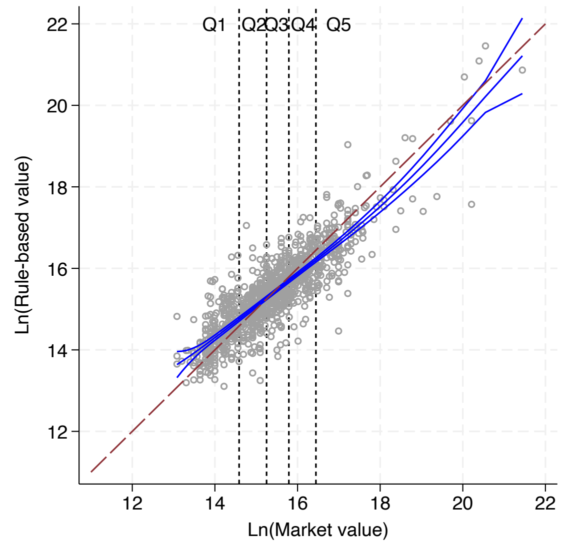


Notes: This Map shows the scope and design of our experiment in the region of Dakar, which comprises the cities of Dakar, Guediawaye, Pikine, Rufisque and Keur Massar. We randomly assigned 48 cadastral sections to receive the property tax census with the discretionary valuation method (represented in orange). Another 48 sections were assigned to receive the census with rule-based valuation (represented in yellow). Lastly, 97 sections (in gray) were randomly assigned to be a pure control group with no property tax census. We stratified by tax office, total number of plots, and share of plots eligible for the property tax. In total, the 193 sections comprise 83,360 plots, and the 96 treated sections span 42,423 plots.

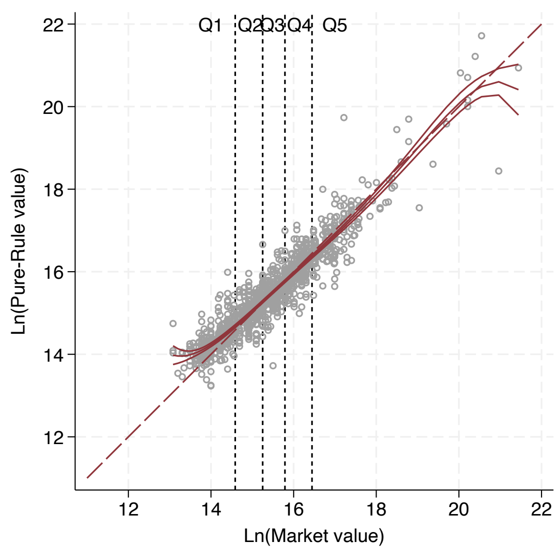
FIGURE 4
VALUATIONS FOR DIFFERENT DEGREES OF DISCRETION



(A) DISCRETION



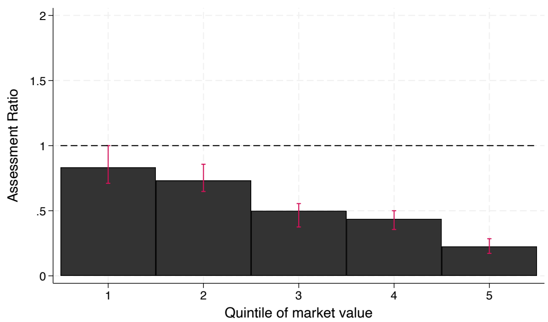
(B) RULE-BASED



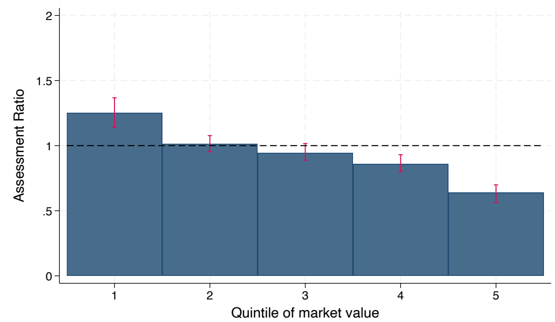
(C) PURE RULE

Notes: This Figure plots valuations of the discretionary arm (Panel (A)), the rule-based values in the rule arm (Panel (B)) and the pure rule values applied to the rule arm (Panel (C)). In each Panel, the x-axis plots $Ln(MarketValue)$ and the y-axis plots $Ln(TaxRollValue)$. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. The curve shows the 6th-degree polynomial fit between the two values, with its 95% confidence interval. The blue dots plot observations for which $TaxRollValue = 0$. The dashed red line is the 45-degree identity line. The black vertical lines indicate the quintiles of market values. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

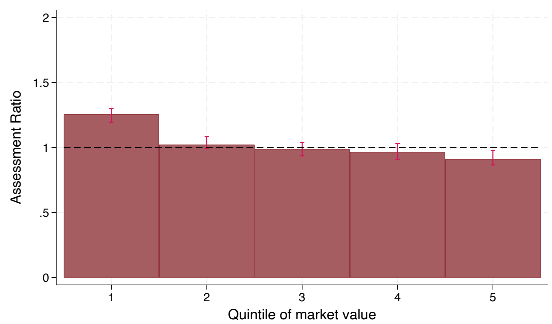
FIGURE 5
ASSESSMENT RATIO BY QUINTILE FOR DIFFERENT DEGREES OF DISCRETION



(A) DISCRETION



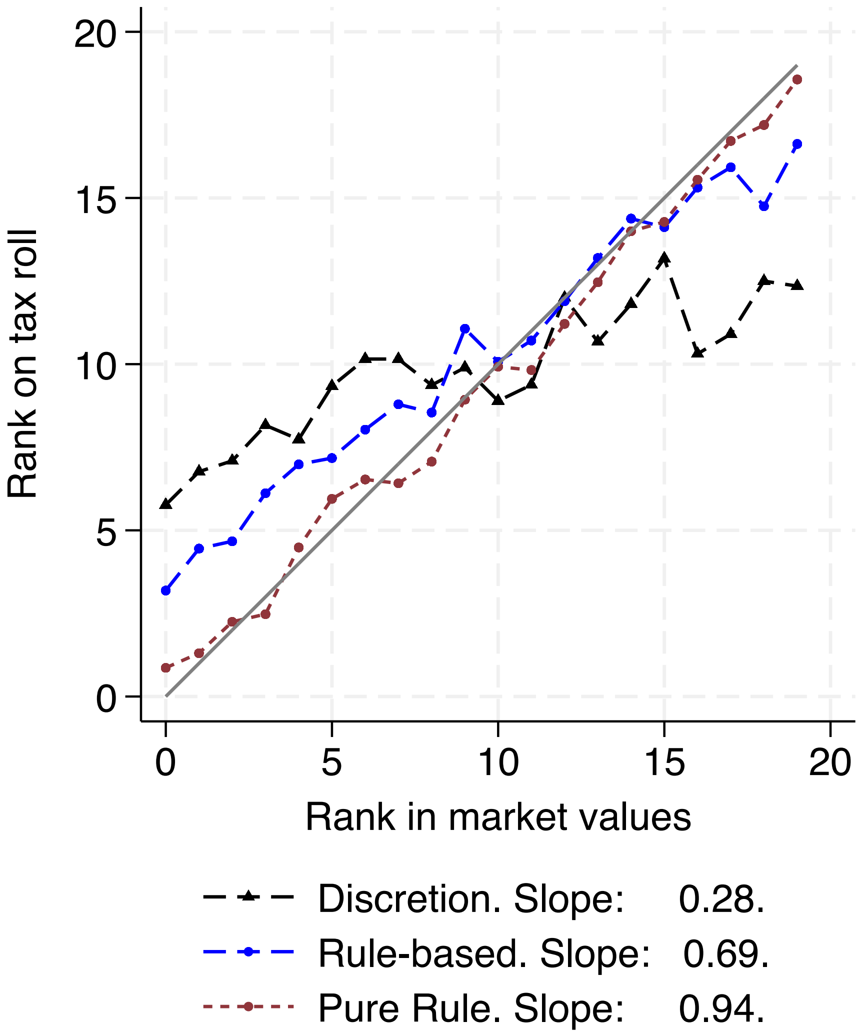
(B) RULE-BASED



(C) PURE RULE

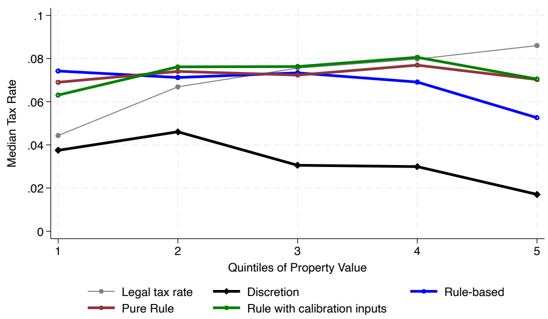
Notes: This Figure plots the median assessment ratio (tax roll value over market value) by quintile for the discretionary arm (Panel (A)), for rule-based values in the rule arm (Panel (B)), for pure rule values applied to the rule arm (Panel (C)). The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. Quintiles are based on market values. The red line shows the 95% confidence interval for the median. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

FIGURE 6
VERTICAL EQUITY: RANK-RANK CORRELATIONS

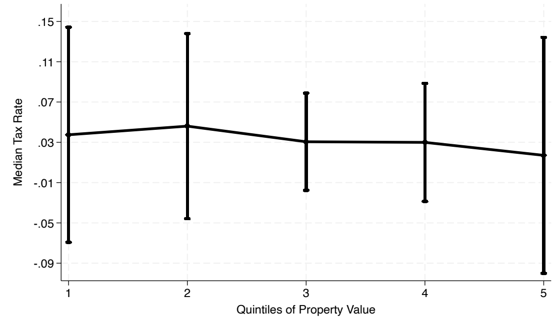


Notes: This Figure shows the rank-rank correlation between tax roll values and market values, separately for the discretionary arm (black line), rule-based values in the rule arm (blue line) and pure rule values applied to the rule arm (red line). The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. The x-axis shows a property's rank in market values, grouped in 20 bins. Ties are assigned the same rank. The y-axis shows the mean rank for the bin. We estimate the slope by regressing the binned tax roll rank on the binned market value rank. The gray line is the 45-degree identity line. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

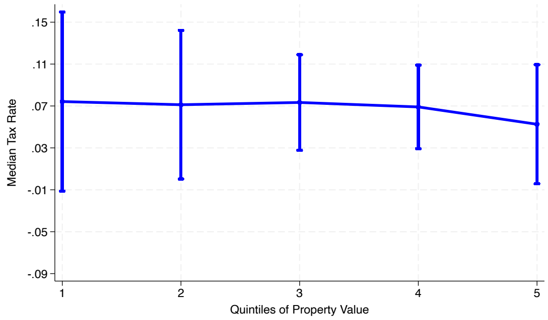
FIGURE 7
TAX RATES: REMOVING DISCRETION INCREASES ACCURACY AND EQUITY



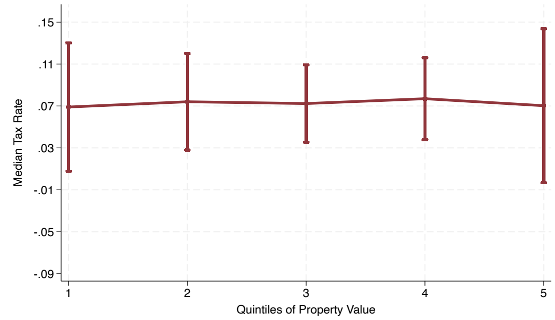
(A) MEDIAN TAX RATES BY QUINTILE



(B) DISCRETION



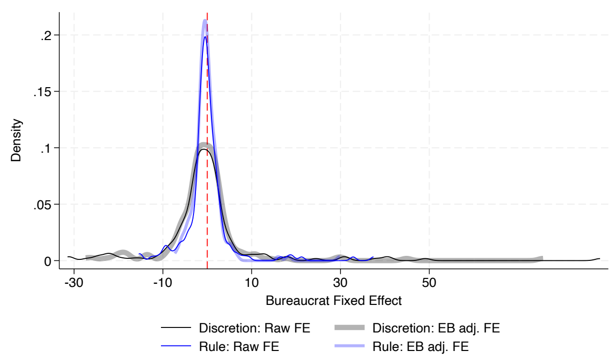
(C) RULE-BASED



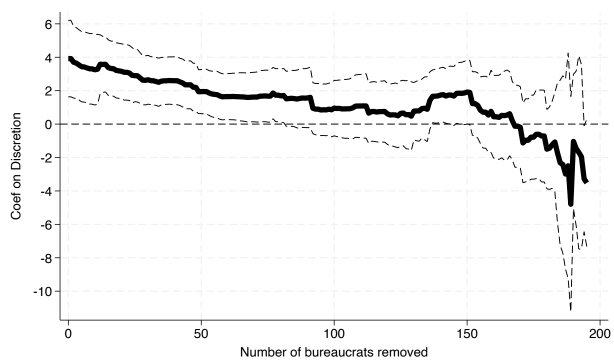
(D) PURE RULE

Notes: This Figure shows the median effective tax rate by quintile of market value for different degrees of bureaucrat discretion. A property's effective tax rate is computed as the tax liability based on the tax roll value over market value. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. In Panel (A), we show the tax profile for all methods. In Panels (B), (C), and (D), we show the tax profile separately adding the standard deviation of the tax rate in each quintile. The gray line in Panel (A) is the benchmark tax profile computed by applying the tax code rates directly to market values (the tax rate is 8.6 percent with an abatement if the property is owner-occupied which explains the progressive profile). The black lines in Panel (A) and Panel (B) are for the discretionary arm, the blue lines in Panel (A) and Panel (C) are for rule-based values in the rule arm, the red lines in Panel (A) and Panel (D) are for the pure rule values applied to the rule arm. Additionally in Panel (A) the green line plots the tax profile generated with the rule if using the characteristics from the calibration (assessors' dataset). Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

FIGURE 8
BUREAUCRAT FIXED-EFFECTS



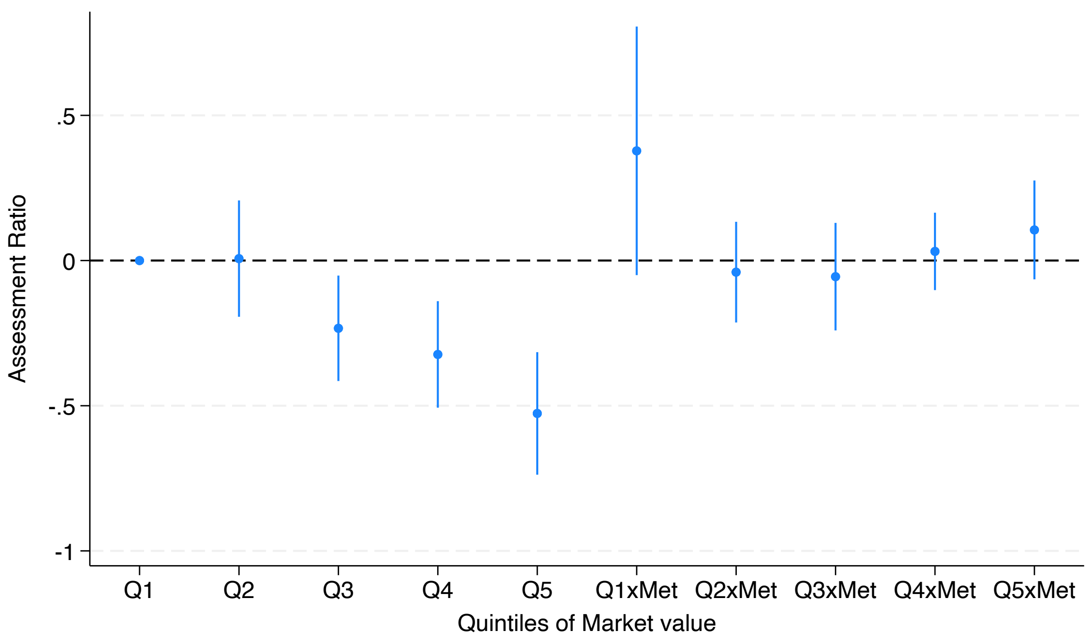
(A) ESTIMATED FIXED-EFFECTS



(B) SCREENING TO CLOSE THE GAP BETWEEN RULE-BASED AND DISCRETION

Notes: This Figure shows results from the bureaucrats fixed-effects estimation presented in Section 5.2. The α_b are estimated in specification 3 which we run separately for each arm: $|Gap|_{ijb} = \alpha_b + Val_j + \epsilon_{ijb}$. $|Gap|_{ijb}$ is the absolute value of the tax base gap, measured as tax roll value minus market value for property i of section j covered by bureaucrat b. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. We then apply an empirical Bayes adjustment to recover $\alpha_{b,EB}$. In Panel (A), we plot the kernel density estimate of the distribution of $\alpha_{b,EB}$ under discretion (in gray and black) and under the rule-based system (in light blue and blue). In Panel (B) we assess how much screening would be needed for the difference in tax base gap between rule and discretion to fade. We rank bureaucrats based on their discretionary $\alpha_{b,EB}$. Then, we run regression (2) 195 times on the discretionary arm, removing bureaucrats one by one starting by the worst one. The number of bureaucrats removed is indicated on the x-axis. The solid black line shows the estimated $\hat{\beta}$ coefficient on a dummy for discretionary sections. The dashed line indicates the the 95% confidence interval.

FIGURE 9
UNDER DISCRETION: THE UNDERVALUATION GRADIENT DOES NOT DIFFER WHEN THE OWNER IS MET



Notes: This Figure shows how the assessment ratio under discretion varies by quintile, and depending on whether the bureaucrat met the owner. We plot the β_n and γ_n coefficients from regression 5: $AR_i = \alpha + \sum_{n=1}^{5} \beta_n Q(n)_i + \sum_{n=1}^{5} \gamma_n M_i \cdot Q(n)_i + \epsilon_i$, where AR_i is the assessment ratio (tax roll value over market value) for property i, the $Q(n)$ are dummies for each quintile of the distribution of market values, M_i is a dummy taking value one if the bureaucrat reports meeting the owner. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. Errors are clustered at the section level. Sample: analysis sample, i.e., properties for which we also have market values, restricted to the discretionary arm ($N = 1,124$).

Tables

TABLE 1
BALANCE TABLE

Panel A: Section characteristics across treatment arm	Mean (SD)	$\hat{\beta}_{Discretion}$	P-value	N
Source: Cadastral data				
Number of plots	419 (177.61)	71.50 (37.40)	0.06	94
Built area (0/1)	0.93 (0.25)	0.01 (0.02)	0.51	41,609
Built area (m2)	282 (856.10)	−10.91 (27.45)	0.69	41,609
Source: Assessors				
Eligible (0/1)	0.87 (0.33)	0.02 (0.02)	0.39	2,844
Value (mil. winz)	10.71 (21.93)	1.91 (2.39)	0.43	2,469
Value per m2 (mil. winz)	0.02 (0.02)	0.00 (0.00)	0.24	2,410
Source: Baseline				
Non-response (0/1)	0.53 (0.50)	−0.03 (0.03)	0.32	2,537
Value (mil. winz)	3.95 (4.17)	−0.07 (0.32)	0.83	1,225
Rented (0/1)	0.30 (0.46)	0.03 (0.03)	0.30	1,238
Owner-occupied (0/1)	0.60 (0.49)	0.01 (0.03)	0.64	1,238
High household income (0/1)	0.22 (0.41)	0.00 (0.03)	0.94	1,238
In tax net (0/1)	0.18 (0.38)	0.00 (0.03)	0.93	1,238
Joint significance (1)			0.34	
Joint significance (2)			0.88	
Panel B: Bureaucrat characteristics across treatment arm	Mean (SD)	$\hat{\beta}_{Discretion}$	P-value	N
Age	31.51 (5.80)	0.15 (0.34)	0.66	1,288
Female	0.28 (0.45)	−0.01 (0.03)	0.61	1,288
Ever worked with tax adm.	0.18 (0.38)	0.03 (0.02)	0.20	1,288
From Dakar	0.48 (0.50)	−0.01 (0.03)	0.66	1,288
Commune of residence	0.06 (0.24)	0.00 (0.01)	0.94	1,266
Any higher education	0.98 (0.14)	−0.01 (0.01)	0.12	1,288
3 yrs higher education	0.40 (0.49)	−0.01 (0.03)	0.66	1,288
Ethnic group: Wolof (majority)	0.31 (0.46)	0.01 (0.03)	0.65	1,288
Religion: Tidjane (majority)	0.56 (0.50)	0.03 (0.03)	0.28	1,288
Public service motivation (index)	0.06 (0.85)	−0.02 (0.05)	0.73	1,288
In favor of government's role (index)	0.13 (0.97)	−0.07 (0.06)	0.21	1,268
In favor of widespread taxation (index)	0.01 (1.00)	−0.04 (0.06)	0.48	1,288
Joint significance			0.71	
Panel C: Bureaucrat characteristics and property values	Mean (SD)	$\hat{\beta}_{Ln(Value)}$	P-value	N
Age	31.53 (5.64)	0.03 (0.18)	0.86	2,236
Female	0.25 (0.43)	0.01 (0.01)	0.49	2,236
Ever worked with tax adm.	0.18 (0.39)	0.00 (0.01)	0.94	2,236
From Dakar	0.48 (0.50)	0.00 (0.02)	0.90	2,236
Commune of residence	0.09 (0.29)	0.00 (0.01)	0.56	2,187
Any higher education	0.99 (0.10)	0.00 (0.00)	0.79	2,236
3 yrs higher education	0.41 (0.49)	−0.04 (0.01)	0.01	2,236
Ethnic group: Wolof (majority)	0.32 (0.47)	0.00 (0.01)	0.92	2,236
Religion: Tidjane (majority)	0.57 (0.50)	−0.02 (0.02)	0.21	2,236
Public service motivation (index)	0.08 (0.80)	0.00 (0.02)	0.86	2,236
In favor of government's role (index)	0.08 (0.93)	0.02 (0.03)	0.55	2,214
In favor of widespread taxation (index)	−0.02 (0.99)	0.01 (0.03)	0.81	2,236
Joint significance			0.41	

Notes: This Table verifies that section and bureaucrat characteristics are balanced across treatment arms, and that bureaucrat characteristics do not correlate with market values. In Panel (A) we regress section characteristics on a dummy for discretionary sections. $\hat{\beta}_{Discretion}$ is the coefficient on the discretion dummy followed by its standard error. The variable sources are indicated and are either cadastral data, assessors' dataset or property owner baseline survey. All regressions are at the plot level except row one (at the section level). In Panel (B) we regress bureaucrat characteristics on a dummy for discretionary sections. Observations are at the bureaucrat X section level. In Panel (C), we regress bureaucrat characteristics on market values and report the coefficient of interest $\hat{\beta}_{Ln(Value)}$. The market value is the value obtained from licensed assessors. Observations are at the plot level. See Appendix Section B.5 for a detailed description of the variables from the bureaucrat surveys. In all regressions, we control for strata fixed-effects and we cluster errors at the section level. At the bottom of each Panel we report the p-value for an F-test of joint significance. In Panel (A) this is done separately for cadastral and assessor variables (1), and for baseline variables (2).

TABLE 2
THE UNDERVALUATION GRADIENT

	(1) Q1	(2) Q2	(3) Q3	(4) Q4	(5) Q5
Panel A: Discretion					
Median Ass. Ratio	0.83	0.73	0.50	0.44	0.23
$\hat{\beta}_n$	Ref.	-0.10	-0.33	-0.40	-0.61
P-value		0.31	0.00	0.00	0.00
P-value $\hat{\beta}_n \neq \hat{\beta}_{n+1}$		0.00	0.27	0.00	
Panel B: Rule-based					
Median Ass. Ratio	1.25	1.01	0.94	0.86	0.64
$\hat{\beta}_n$	Ref.	-0.24	-0.31	-0.39	-0.61
P-value		0.00	0.00	0.00	0.00
P-value $\hat{\beta}_n \neq \hat{\beta}_{n+1}$		0.15	0.10	0.00	
Panel C: Pure Rule					
Median Ass. Ratio	1.26	1.02	0.98	0.97	0.91
$\hat{\beta}_n$	Ref.	-0.23	-0.27	-0.29	-0.34
P-value		0.00	0.00	0.00	0.00
P-value $\hat{\beta}_n \neq \hat{\beta}_{n+1}$		0.34	0.61	0.14	

Notes: In this Table, we test whether the median assessment ratio changes significantly across quintiles of market values, under discretion (Panel (A)), under the rule-based system (Panel (B)), under the pure rule system applied to the rule arm (Panel (C)). We estimate $AR_i = \alpha + \sum_{n=1}^{5} \beta_n Q(n)_i + \epsilon_i$ where AR_i is the assessment ratio (tax roll value over market value) for property i, and the $Q(n)$ are dummies for each quintile of the distribution of market value. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. We run a quantile regression at the median. Standard errors are clustered at the section level. The second row of each Panel reports the $\hat{\beta}_n$ coefficients. Below, we report the P-value indicating whether $\hat{\beta}_n$ is significantly different from the reference ($Q(1)$). P-val $\hat{\beta}_n \neq \hat{\beta}_{n+1}$ indicates whether the coefficients for two subsequent quintiles are significantly different from each other. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

TABLE 3
REMOVING DISCRETION INCREASES ACCURACY

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) | Gap | *mil.FCFA*	(4) Ass. Ratio
Panel A: Discretion				
Mean[1] (sd)	−7.14 (17.69)	−2.41	8.92 (16.87)	0.71
Panel B: Rule-based				
Overall				
Mean[1] (sd)	−2.33 (12.78)	−0.16	4.67 (12.12)	1.06
$\hat{\beta}_{Discretion}$	−4.61***	−1.87***	3.88***	−0.35***
	(1.28)	(0.38)	(1.38)	(0.05)
Low Value				
Mean[1] (sd)	0.44 (2.06)	0.20	1.25 (1.70)	1.27
$\hat{\beta}_{Discretion}$	−0.53***	−0.61***	0.32**	−0.24**
	(0.19)	(0.13)	(0.15)	(0.09)
High Value				
Mean[1] (sd)	−4.83 (17.13)	−1.58	7.74 (16.02)	0.87
$\hat{\beta}_{Discretion}$	−6.52***	−4.41***	5.51***	−0.34***
	(1.77)	(0.68)	(1.89)	(0.05)
Panel C: Pure Rule				
Overall				
Mean[1] (sd)	−0.36 (7.64)	0.12	2.83 (7.11)	1.13
$\hat{\beta}_{Discretion}$	−5.37***	−2.42***	4.71***	−0.38***
	(0.90)	(0.44)	(0.93)	(0.04)
Low Value				
Mean[1] (sd)	0.39 (1.13)	0.25	0.73 (0.94)	1.24
$\hat{\beta}_{Discretion}$	−0.28	−0.57***	0.88***	−0.13
	(0.20)	(0.14)	(0.14)	(0.09)
High Value				
Mean[1] (sd)	−1.04 (10.43)	−0.33	4.72 (9.36)	1.03
$\hat{\beta}_{Discretion}$	−8.25***	−5.44***	6.87***	−0.47***
	(1.27)	(0.55)	(1.28)	(0.05)
N plots: 2290				
N Sections: 94				
Mean (sd) market value: 77.00 (15.80)				
Median market value: 5.60				

Notes: This Table shows the effect of discretion on the tax base gap. We run regression 2: $Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk}$ with four different outcomes. Y_{ijk} is the outcome for property i of section j and strata k, D is a dummy for discretionary sections and S_k is a strata fixed-effect. In column (1) the outcome variable is the tax base gap defined as tax roll value minus market value, column (2) uses the same outcome but with a quantile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio (tax roll value over market value). In Panel (B), values for the rule arm are the rule-based valuations incorporating bureaucrats' inputs. In Panel (C), values for the rule arm are the pure rule predictions based on remote covariates only. The first sub-panel of each Panel uses all properties of our analysis sample, the second is restricted to low value properties (quintile 1 and 2 of market values), the third is restricted to high value properties (quintiles 3 to 5 of market values). In each sub-panel, the first row displays descriptive statistics of the outcome variable in the rule-based arm; the second row shows the coefficient of interest and its standard error. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed-effects and errors are clustered at the section level. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm). [1]In column (2) the displayed value is the *median* of the tax base gap.

TABLE 4
ESTIMATING BUREAUCRAT FIXED-EFFECTS

	(1) Discretion	(2) Rule-based
N obs	1,055	1,063
N Bur FE	198	190
Mean of Outcome (mil. of FCFA)	8.00	3.81
Var of Outcome	217.09	103.65
R2 without Bur FE	0.38	0.28
R2 with Bur FE	0.52	0.39
Var(Bur FE)	141.71	33.11
Var(Shrinked Bur FE)	86.77	13.37
Share Variance	0.40	0.13
P-value of F test on Bur FE	0.00	0.00

Notes: This Table summarizes results from the estimation of bureaucrat fixed-effects presented in Section 5.2. The α_b are estimated in specification 3 which we run separately for each arm: $|Gap|_{ijb} = \alpha_b + Val_j + \epsilon_{ijb}$. $|Gap|_{ijb}$ is the absolute value of the tax base gap measured as tax roll value minus market value for property i of section j covered by bureaucrat b. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. We then apply an empirical Bayes adjustment to recover the shrinked fixed-effects $\alpha_{b,EB}$. Share of variance refers to the share of variance in the tax base gap accounted for by bureaucrat heterogeneity, it is computed as $Var(\alpha_{b,EB})$ over $Var(|Gap|)$.

TABLE 5
WHAT DO TOP BUREAUCRATS DO DIFFERENTLY?

Depedent Variable (0,1)	Positive value (1)	Owner Met (2)	Owner Details (3)	Contract (4)	Comment (5)	Conflict (6)	Bureaucrat estimate (7)
Top bureaucrat	0.129***	0.045**	0.056***	0.004	0.105***	-0.005	0.060**
	(0.026)	(0.018)	(0.016)	(0.005)	(0.038)	(0.011)	(0.024)
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	22314	22314	22314	20086	22314	10918	10918
R2	0.13	0.03	0.09	0.02	0.29	0.02	0.18
Adj. R2	0.13	0.03	0.09	0.01	0.29	0.02	0.18
Mean of dependent	0.58	0.23	0.31	0.03	0.49	0.09	0.18

Notes: In this Table, we report results from regressing plot level outcomes on a dummy taking value one if the plot is covered by a top bureaucrat. We use the bureaucrat fixed-effects estimated for the discretionary arm in Section 5.2 to define top bureaucrats as those with $\alpha_{b,EB} < 0$. All outcome variables are dummies. *Owner details* takes value one if the bureaucrat recovered the name and/or ID number of the owner. *Contract* takes value one if there is at least one rental contract reported for the plot. *Comment* takes value one if the bureaucrat left a comment. *Conflict* and *bureaucrat estimate* are conditional on leaving a comment: the former takes value one if the bureaucrat mentioned tensions or conflict with occupants, the latter takes value one if the bureaucrat explicitly states that she made an estimation without using information provided by occupants. We control by strata fixed-effects and errors are clustered at the section level. Sample: all plots of the discretionary arm.

TABLE 6
LAB-IN-THE-FIELD: BUREAUCRATS' LACK OF KNOWLEDGE AND BIASES

Depedent Variable	Ass. Ratio (1)	Ass. Ratio (2)	Ln(Value) (3)	Ln(Value) (4)
High value property	-1.419***	-1.491***		1.232***
	(0.112)	(0.109)		(0.091)
Info treatment		-0.085		
		(0.150)		
Info X High value		0.127		
		(0.150)		
Retired owner			-0.378***	-0.241***
			(0.091)	(0.091)
Retired X High value				0.090
				(0.133)
Strata FE	No	Yes	Yes	Yes
Bureaucrat FE	Yes	No	No	No
N	280	280	280	280
R2	0.83	0.62	0.07	0.57
Adj R2	0.60	0.61	0.01	0.54
Mean in reference	1.74	1.78	12.84	12.08

Notes: This Table shows results from the lab-in-the-field evaluations that we include in the bureaucrat endline survey. Bureaucrats are shown pictures of a high value property and a low value property, and are asked to provide an estimated rental value for each. Column (1) displays results from equation 4: $AR_{ib} = \alpha_b + \beta High_{ib} + \epsilon_{ib}$, where AR_{ib} is the assessment ratio computed as bureaucrat b's value over market value for property i. The market value is the value obtained from licensed assessors for these specific properties. We include bureaucrat fixed-effects in Column (1). In Column (2) we add the dummy $Info$ taking value one if the bureaucrat received a randomized information treatment on the true distribution of market values (shown in Figure A14). Columns (3) and (4) show results from regression 6: $Ln(Value_{ibk}) = \alpha + \beta_1 Retired_{ibk} + \beta_2 High_{ibk} + \beta_3 Retired_{ibk} \cdot High_{ibk} + A_k + \epsilon_{ibk}$. The dummy $Retired$ takes value one if the bureaucrat received the randomized information that the owner was retired (versus employed). The two randomizations were independent, and we control for bureaucrat randomization strata. Observations are at the bureaucrat X property level. Standard errors are clustered at the bureaucrat level.

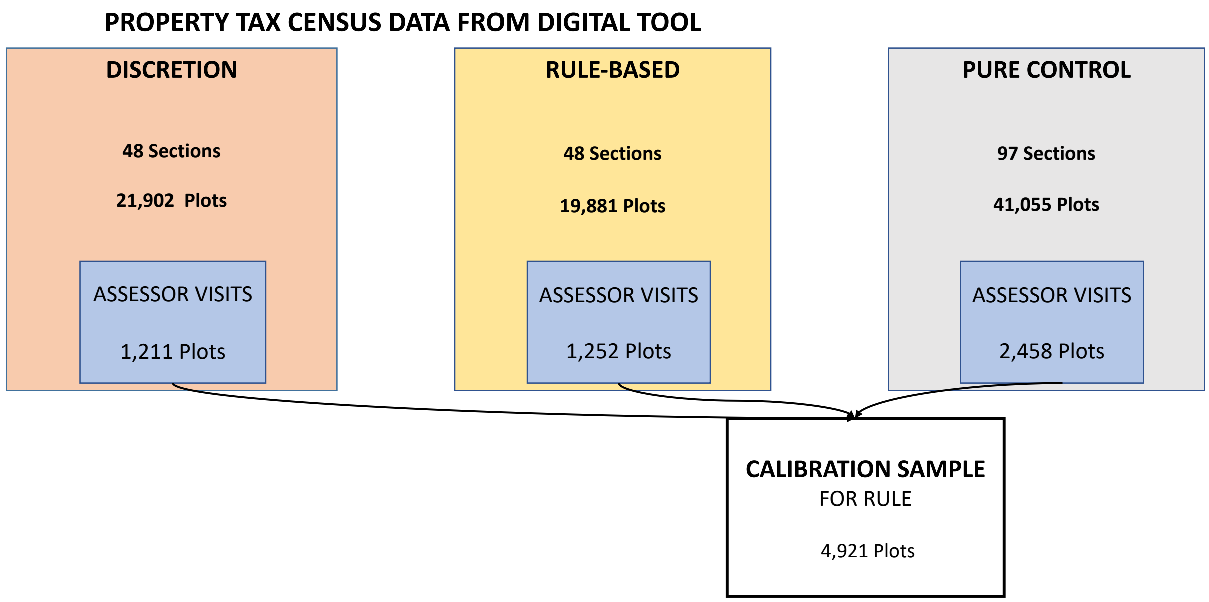
TABLE 7
RULE-BASED VS PURE RULE: MEASURING THE EFFECT OF A LIMITED DEGREE OF DISCRETION

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) |Gap| *mil.FCFA*	(4) Ass. Ratio
Panel A: Pure Rule				
Overall				
Mean[1] (sd)	−0.36 (7.64)	0.12	2.83 (7.11)	1.13
$\hat{\beta}_{RuleBur}$	−1.97***	−0.31**	1.84***	−0.07*
	(0.58)	(0.12)	(0.44)	(0.04)
Low Value				
Mean[1] (sd)	0.39 (1.13)	0.25	0.73 (0.94)	1.24
$\hat{\beta}_{RuleBur}$	0.06	−0.01	0.52***	0.03
	(0.12)	(0.07)	(0.10)	(0.05)
High Value				
Mean[1] (sd)	−1.04 (10.43)	−0.33	4.72 (9.36)	1.03
$\hat{\beta}_{RuleBur}$	−3.80***	−1.28***	3.03***	−0.17***
	(0.95)	(0.35)	(0.77)	(0.04)
Panel B: Rule with Assessor Inputs				
Overall				
Mean[1] (sd)	−0.40 (8.16)	0.14	2.78 (7.68)	1.10
$\hat{\beta}_{RuleBur}$	−1.94***	−0.33**	1.89***	−0.04
	(0.50)	(0.13)	(0.39)	(0.04)
Low Value				
Mean[1] (sd)	0.31 (0.90)	0.16	0.63 (0.72)	1.15
$\hat{\beta}_{RuleBur}$	0.13	0.01	0.62***	0.12**
	(0.12)	(0.07)	(0.10)	(0.05)
High Value				
Mean[1] (sd)	−1.03 (11.18)	−0.04	4.71 (10.19)	1.05
$\hat{\beta}_{RuleBur}$	−3.80***	−1.65***	3.03***	−0.19***
	(0.79)	(0.34)	(0.64)	(0.04)
N obs: 2331				
N plots: 1166				
N Sections: 47				
Mean (sd) market value: 86.00 (15.40)				
Median market value: 4.80				

Notes: This Table shows the effect on the tax base gap of a limited degree of discretion, when bureaucrats implement the rule-based system, compared to benchmark pure rules without any bureaucrat discretion. We run regression 7: $Y_{irjk} = \alpha + \beta RuleBur_{irjk} + S_k + \epsilon_{irjk}$ where Y_{itjk} is the outcome for plot i of section j and strata k under rule r, $RuleBur_{irjk}$ is a dummy taking value one if r is the rule implemented by bureaucrats, and zero if r is the benchmark pure rule with no bureaucrat discretion. In Panel (A), the benchmark rule is the pure rule with remote covariates. In Panel (B) the benchmark rule is the rule with all covariates but using the calibration inputs from the assessors' dataset. In column (1) the outcome variable is the tax base gap defined as tax roll value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as the tax roll value over the market value. Each panel is divided in three subpanels, the first one uses all properties from our analysis sample, the second is restricted to low value properties (quintile 1 and 2 of market values), the third one is restricted to high value properties (quintiles 3 to 5 of market values). In each sub-panel, the first row displays descriptive statistics of the outcome variable with the benchmark rule; the second row shows the coefficient of interest and its standard error. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed-effects and errors are clustered at the section level. Sample: analysis sample, i.e., properties for which we also have market values, restricted to the rule-based arm ($N = 1,166$). Each property appears twice in the regression sample. [1]In column (2) the displayed value is the *median* of the tax base gap.

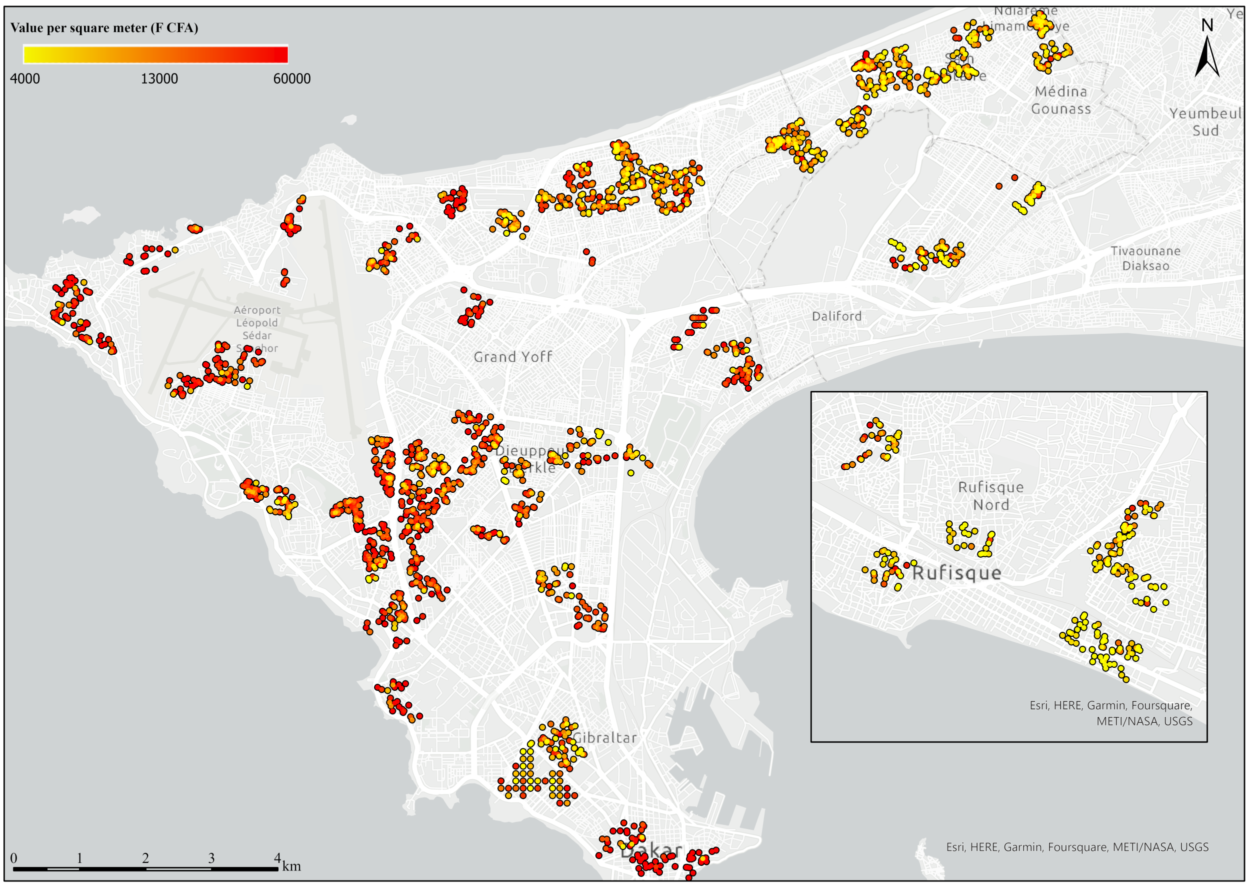
Additional Figures

FIGURE A1
OVERLAP BETWEEN EXPERIMENT ARMS AND ASSESSOR VALUATIONS



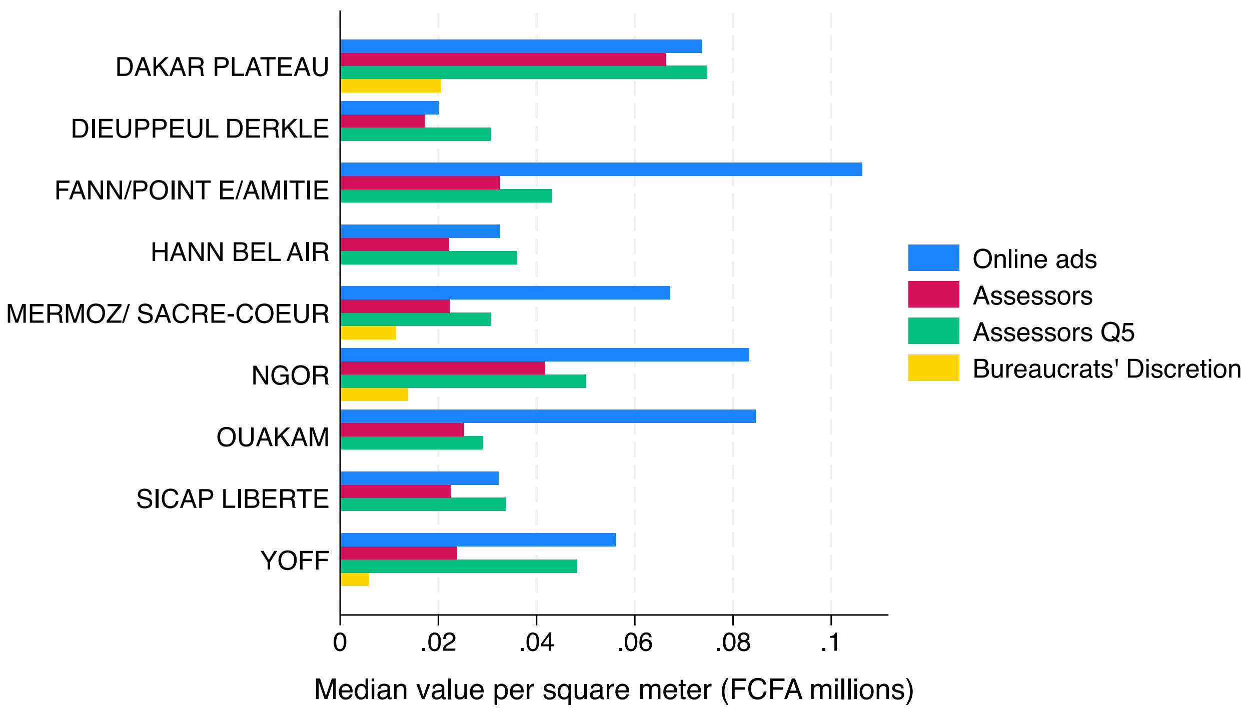
Notes: This Figure summarizes how our different samples overlap. Cadastral sections are randomly allocated either to the discretion arm, the rule arm, or the pure control arm. Plots visited by licensed assessors are drawn randomly from the three types of sections and constitute the calibration sample for the algorithm (see details in Appendix section B.1).

FIGURE A2
MARKET RENTAL VALUE PER SQUARE METER IN DAKAR



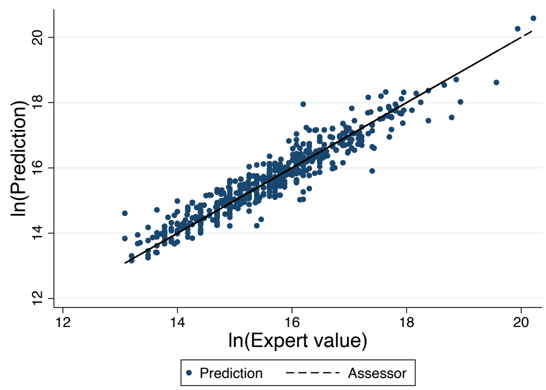
Notes: This Map shows market rental values in Dakar. We compute annual rental value per square meter using market values provided by licensed assessors. We use a different nuance of color for each quintile of the distribution. The legend indicates the mean values in FCFA for quintiles one, five and ten. The Map is restricted to the census sections spanning both arms.

FIGURE A3
COMPARISON WITH VALUES FROM ONLINE RENTAL LISTINGS

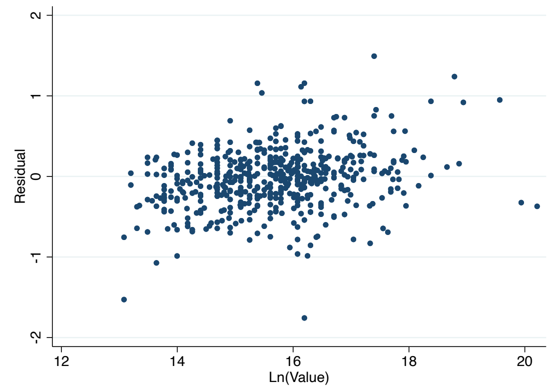


Notes: This Figure shows the median rental value per square meter (in millions of FCFA) by commune for different data sources. A commune comprises several sections, in total our experiment spans 22 communes. *Online ads* refers to rental values we scraped from listings posted on three Senegalese websites. Because these have very imprecise address information, we are only able to link them to a commune, and not to a section. We have 420 observations from online listings, situated in 9 communes. *Assessors* refers to market values provided by licensed assessors that we aggregate at the commune level. *Assessors Q5* refers to assessor values from the fifth (top) quintile only. *Bureaucrats' discretion* refers to bureaucrats' values in the discretionary arm.

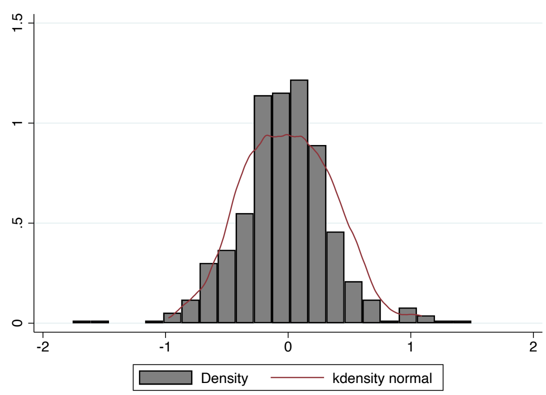
FIGURE A4
RESIDUALS IN THE PROPERTY VALUATION ALGORITHM



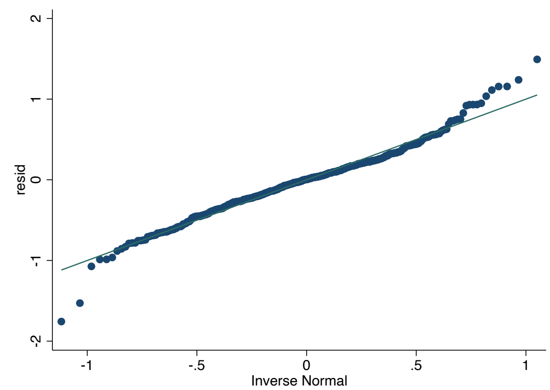
(A) PREDICTIONS OVER MARKET VALUES



(B) RESIDUALS OVER MARKET VALUES



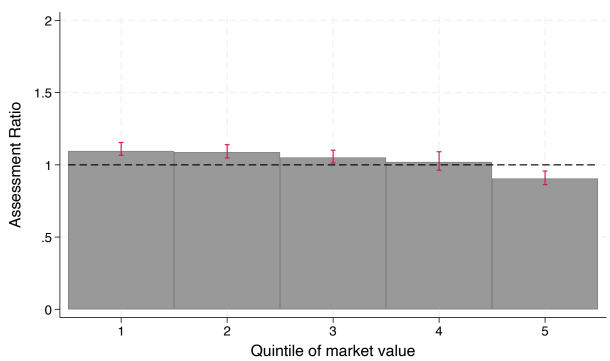
(C) DISTRIBUTION OF RESIDUALS



(D) DIAGNOSTIC: NORMALITY OF RESIDUALS

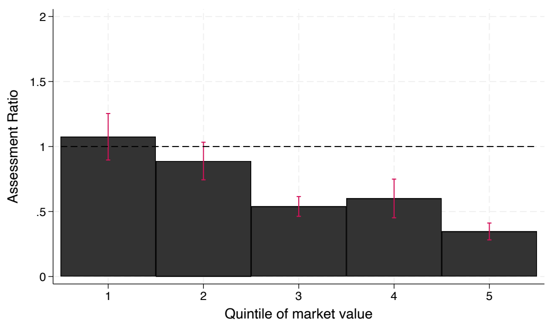
Notes: This Figure shows graphical results from the property valuation algorithm applied to the test sample. The algorithm is described in Sections 4.2 and B.2. In Panel (A) we plot predictions for the test sample $Ln(\widehat{Value})$ over assessor values $Ln(Value)$, in Panel (B) we plot residuals $(Ln(\widehat{Value}) - Ln(Value))$ over $Ln(Value)$, in Panel (C), the gray bars are the histogram of residuals, and as a comparison, we add the kernel density of a normal distribution with a mean of zero and with the same standard error as the distribution of residuals (in red). Panel (D) is a Q-Q diagnostic plot, where quantiles of the residual are plotted over the expected quantiles for a normal distribution.

FIGURE A5
ASSESSMENT RATIO BY QUINTILE FOR THE RULE WITH ASSESSORS' INPUTS

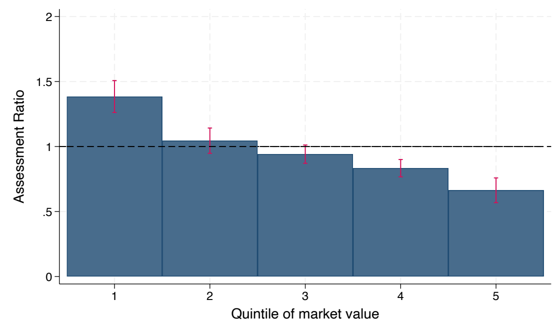


Notes: This Figure plots the median assessment ratio (tax roll value over market value) by quintile for the rule-based arm, using as inputs for the rule the assessor characteristics from the calibration dataset. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. Quintiles are based on market values. Sample: analysis sample, i.e., properties for which we also have market values, restricted to the rule-based arm ($N = 1,166$).

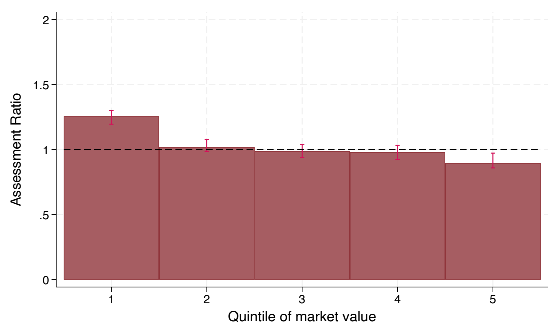
FIGURE A6
ASSESSMENT RATIO BY QUINTILE FOR DIFFERENT DEGREES OF DISCRETION (MEANS)



(A) DISCRETION



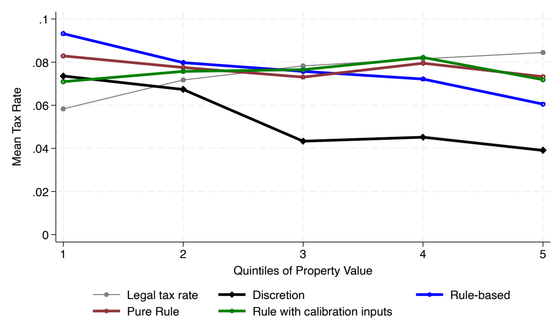
(B) RULE-BASED



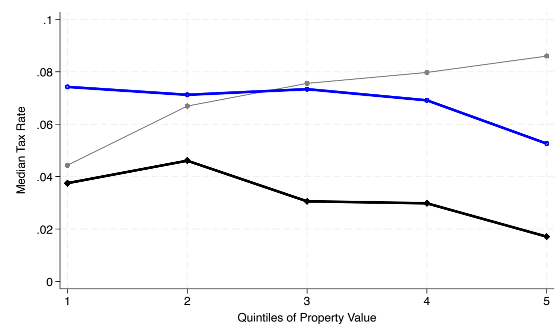
(C) PURE RULE

Notes: This Figure plots the mean assessment ratio (tax roll value over market value) by quintile for the discretionary arm (Panel (A)), for rule-based values in the rule arm (Panel (B)), for pure rule values applied to the rule arm (Panel (C)). The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. Quintiles are based on market values. The red line shows the 95% confidence interval for the mean. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

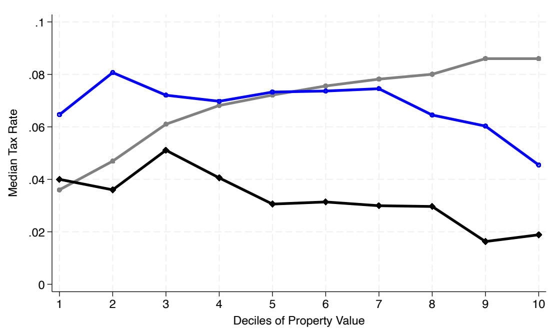
FIGURE A7
TAX RATES: ADDITIONAL RESULTS



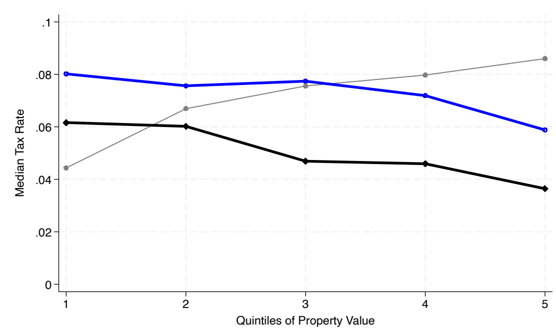
(A) MEAN TAX RATES



(B) ABATEMENT CALCULATED BASED ON OWNER-OCCUPIED ROOMS



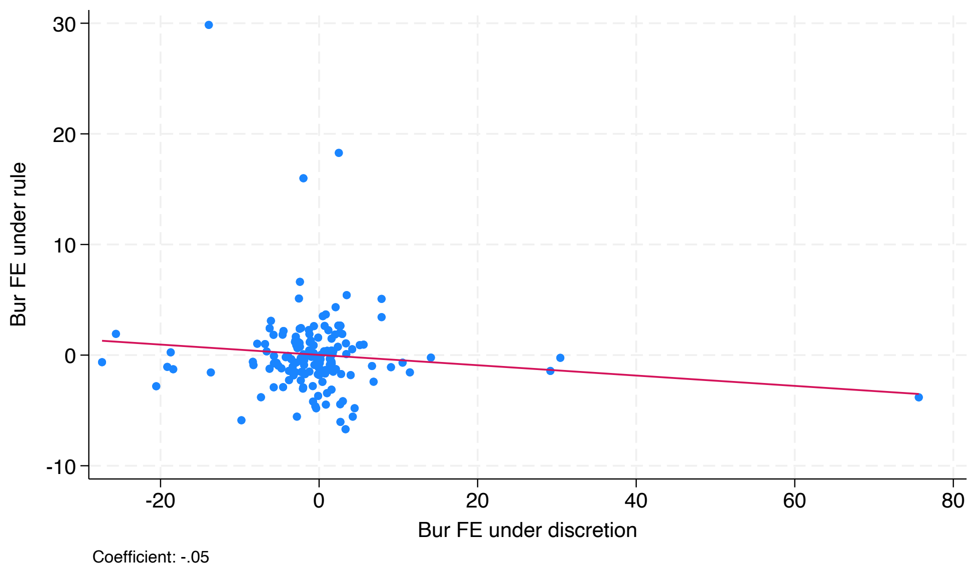
(C) TAX RATE BY DECILE



(D) RESTRICTING TO POSITIVE CASES

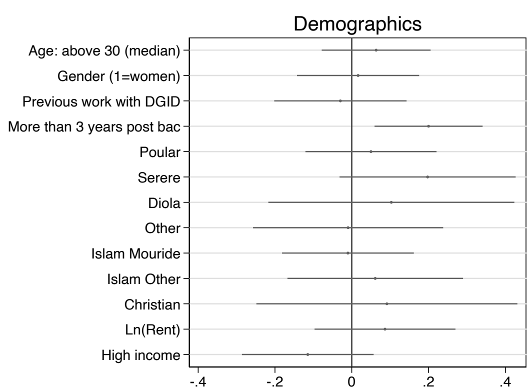
Notes: This Figure is a complement to Figure 7. In Panel (A), we plot the mean (instead of median) tax rate by quintile of market values. In Panel (B), the component of property value corresponding to the owner's main residence, to which the abatement is applied, is calculated identically in both arms. It is based on the number of rooms occupied by the owner – in Figure 7, for the discretionary arm, this component was determined using the values bureaucrats directly report for owner-occupied parts. In Panel (C), we show the median tax rate by decile of market value. In Panel (D), we restrict the sample to properties to which a positive value was assigned by bureaucrats. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

FIGURE A8
CORRELATION BETWEEN BUREAUCRAT FIXED-EFFECTS UNDER RULE AND DISCRETION

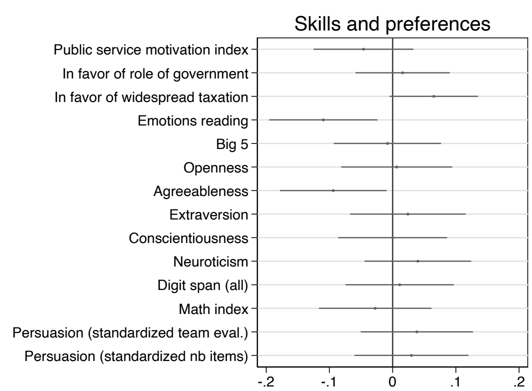


Notes: This Figure plots the correlation between the two fixed-effects $\alpha_{b,EB}$ estimated for a given bureaucrat, on the rule-based arm and on the discretionary arm. The estimation is presented in Section 5.2.

FIGURE A9
CORRELATES OF BEING A TOP BUREAUCRAT



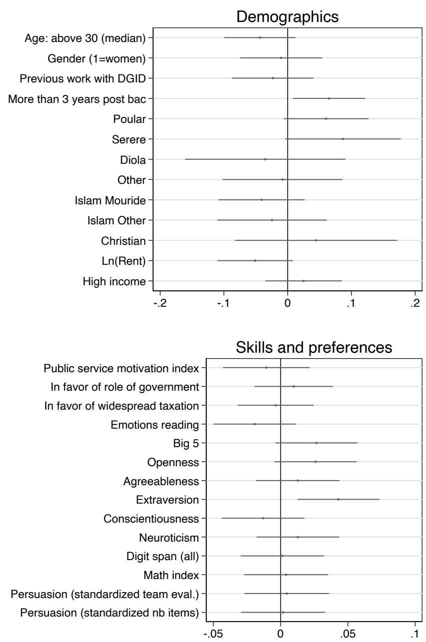
(A) DEMOGRAPHICS



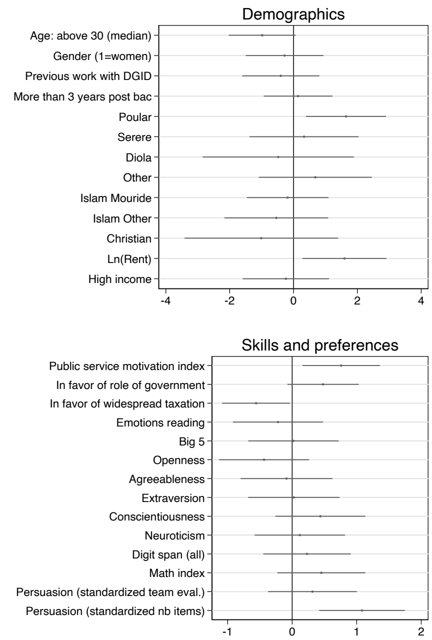
(B) SKILLS AND PREFERENCES

Notes: This Figure reports results from regressing an indicator for being a top bureaucrat on bureaucrat characteristics. We use the bureaucrat fixed-effects estimated for the discretionary arm in Section 5.2 to define top bureaucrats as those with $\alpha_{b,EB} < 0$. The source of the covariates are the bureaucrat baseline and endline surveys. The bar represents the 95% confidence interval on the coefficient of interest. See Appendix Section B.5 for a detailed description of the variables.

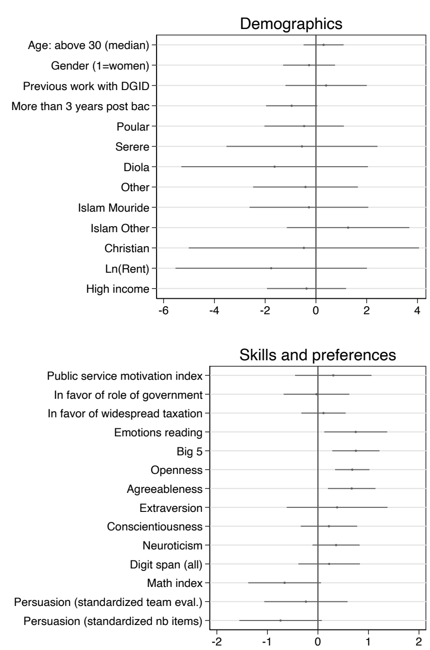
FIGURE A10
CORRELATES OF ALTERNATIVE MEASURES OF BUREAUCRAT PERFORMANCE



(A) SHARE OF PLOTS WITH OWNER DETAILS



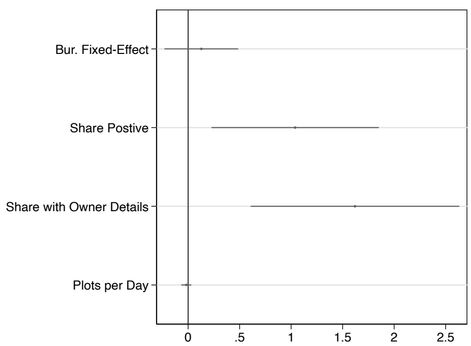
(B) PLOTS PER DAY



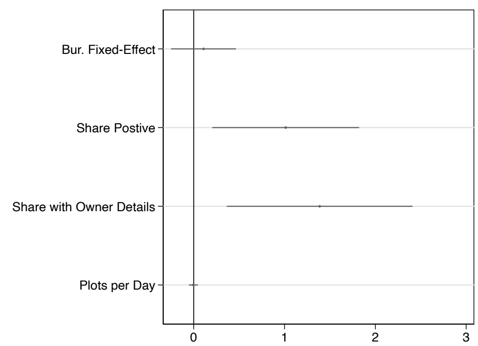
(C) ABILITY TO VALUE HIGH-END PROPERTIES IN FULL SAMPLE

Notes: This Figure reports results from regressing various measures of bureaucrat performance on bureaucrat characteristics. In Panel (A), the performance measure is the share of plots visited by the bureaucrat for which she recovered owner identification details (name and/or ID number). In Panel (B), the performance measure is the average number of plots visited by the bureaucrat in a day. The mean of the outcome variables are respectively 0.37 and 7. In Panel (C), the performance measure is a proxy for ability to value high-end properties. We compute the outcome variable as the absolute tax base gap between bureaucrat values and a predicted value, where the prediction relies on the pure rule with remote covariates. The sample is restricted to properties for which the predicted value is above median. The bar represents the 95% confidence interval on the coefficient of interest. See Appendix Section B.5 for a detailed description of the variables.

FIGURE A11
WHAT DO SUPERVISORS VALUE IN BUREAUCRATS' PERFORMANCE?



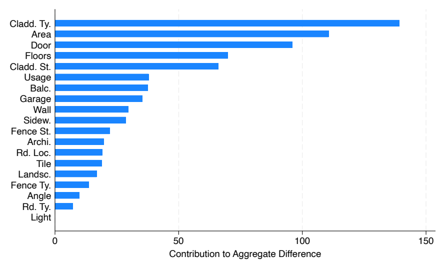
(A) OVERALL PERFORMANCE SCORE



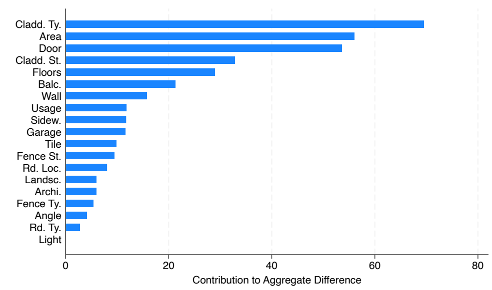
(B) SUM OF PERFORMANCE ITEMS

Notes: This Figure reports results from regressing supervisors' evaluation for a given bureaucrat on performance measures of the bureaucrat. In Panel (A), the outcome is the standardized global performance score a supervisor gave to a given bureaucrat. In Panel (B), the outcome is the standardized sum of scores for a list of performance items (social skills, fiscal knowledge, housing market knowledge, energy and stamina, negotiation skills, ease with technology, ease with reading maps). The performance measures are: a dummy for being a top bureaucrat ($\alpha_{b,EB} > 0$) that we define using the bureaucrat fixed-effects estimated in section 5.2; share of plots visited in the discretionary arm for which the bureaucrat provided a positive value; share of plots visited for which the bureaucrat recovered owner identification details (name and/or ID number); average number of plots per day covered by the bureaucrat. The bar represents the 95% confidence interval on the coefficient of interest. See Appendix Section B.5 for a detailed description of the variables.

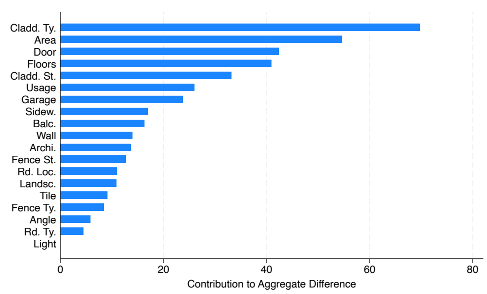
FIGURE A12
WHICH PROPERTY CHARACTERISTICS DRIVE THE EFFECT OF PARTIAL DISCRETION UNDER THE RULE-BASED SYSTEM



(A) ALL PROPERTIES



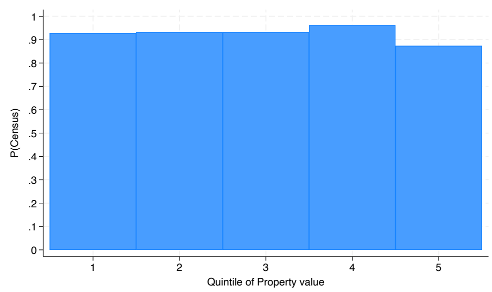
(B) LOW VALUE PROPERTIES



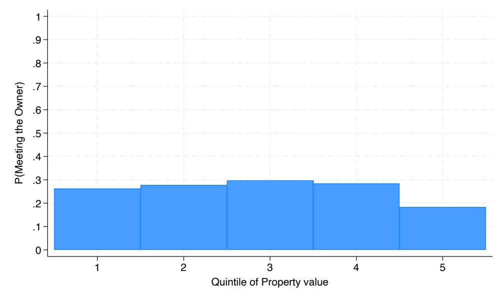
(C) HIGH VALUE PROPERTIES

Notes: This Figure shows the relative contribution of the observable property characteristic to the differences between the rule-based valuation implemented by bureaucrats and the benchmark rule using inputs by assessors. See Panel (B) of Table 7 for a measure of these differences. The contribution of characteristic X_k is calculated as the sum of the absolute difference in predicted values due to this characteristic, more precisely, $\sum_i |\gamma_k X_{k,RuleBur,i} - \gamma_k X_{k,RuleAss,i}|$ where the X_k are the 18 observable characteristics, subscript $RuleBur$ indicates the value taken by X_k when entered by bureaucrat and $RuleAss$ the value taken by X_k when entered by assessors for property i, γ_k is the coefficient for X_k in the algorithm (coefficients shown in Table A3). Panel (A) shows results for all properties from the analysis sample in the rule-based arm, in Panel (B) (resp. Panel (C)) we restrict to low value (resp. high value) properties.

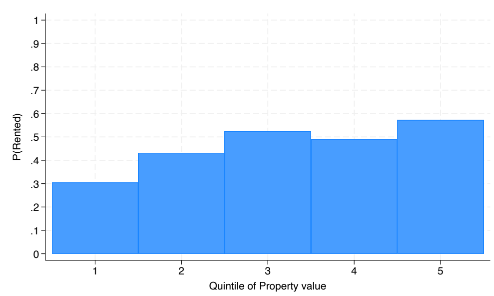
FIGURE A13
MODALITIES OF FIELD VISITS BY PROPERTY VALUE



(A) PROBABILITY OF PLOT BEING COVERED IN CENSUS



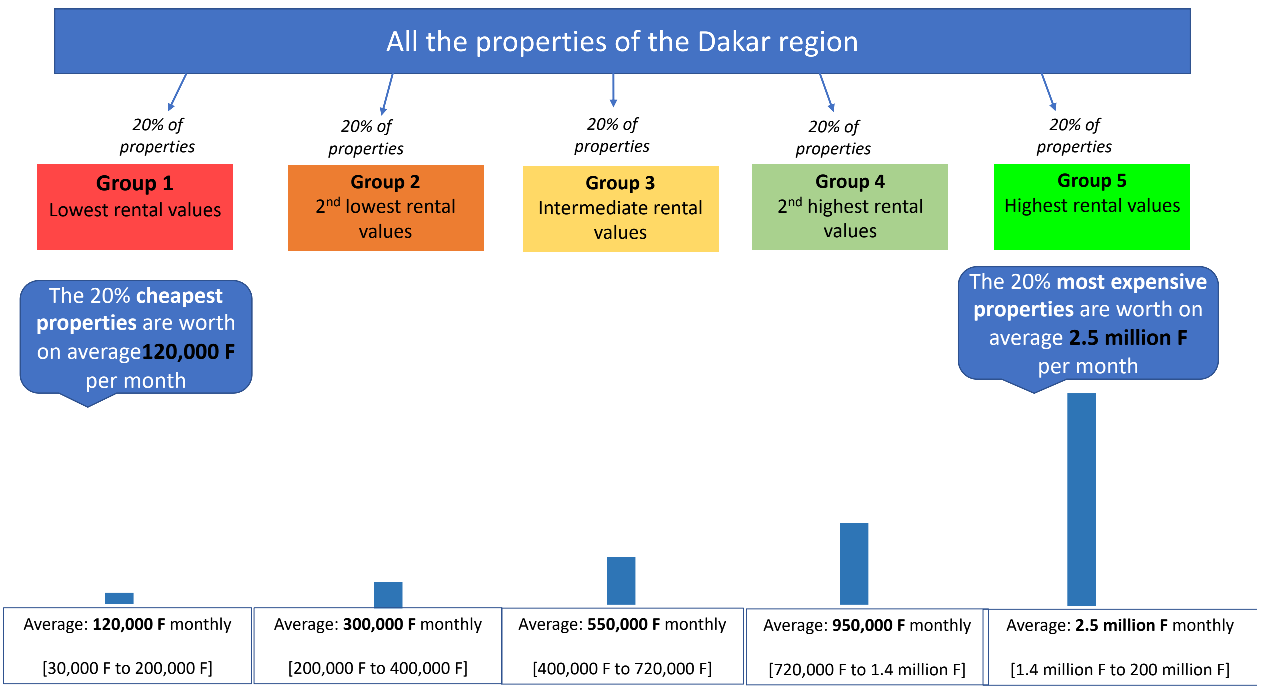
(B) PROBABILITY OF MEETING THE OWNER BY QUINTILE



(C) PROBABILITY OF PROPERTY BEING RENTED BY QUINTILE

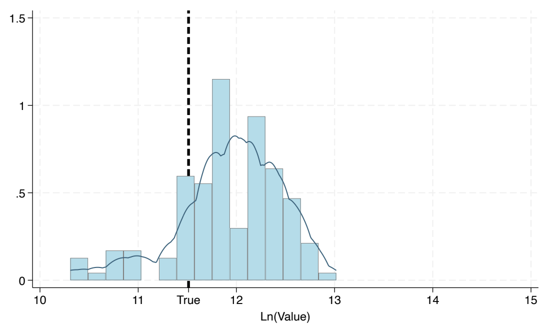
Notes: This Figure shows how three characteristics of property visits during the census vary by quintile of market value. In Panel (A), the underlying variable takes value one if the property is covered at all during the census. In Panel (B), the underlying variable takes value one if the bureaucrat reports meeting the owner. In Panel (C), the underlying variable takes value one if the bureaucrat classified the property as being rented at least in part. The market value is the value obtained from licensed assessors. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

Figure A14
Information Treatment for Lab-in-the-Field Valuations

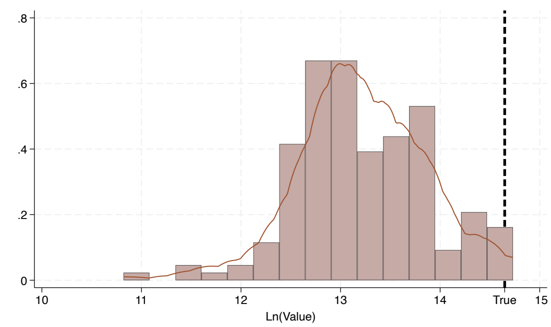


Notes: This Figure displays the information treatment shown to half of the respondents in the lab-in-the-field experiment included in the bureaucrat endline survey. The information is based on market values from licensed assessors. Bureaucrats were randomized into a treated and untreated group, stratifying by gender, accuracy rate observed in the census, and education level. The treated bureaucrats were asked two simple comprehension questions after seeing this chart, to make sure they had carefully looked at it and understood its content.

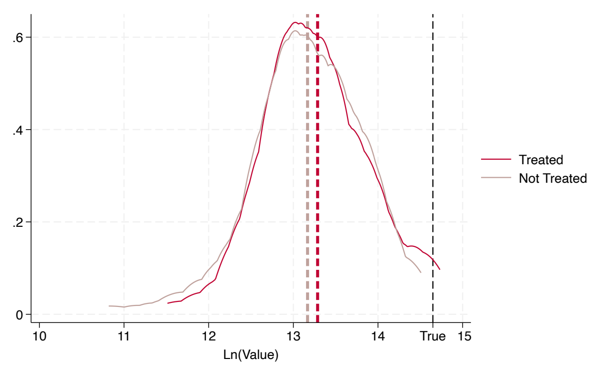
FIGURE A15
LAB-IN-THE-FIELD VALUATIONS: RESULTS



(A) LOW VALUE PROPERTY



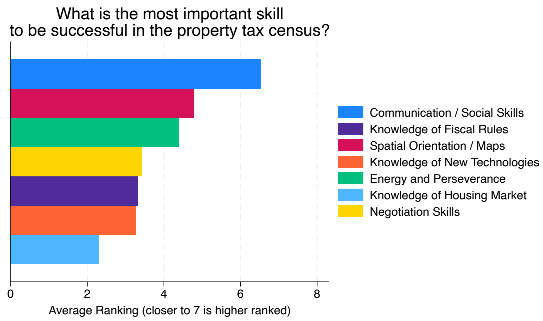
(B) HIGH VALUE PROPERTY



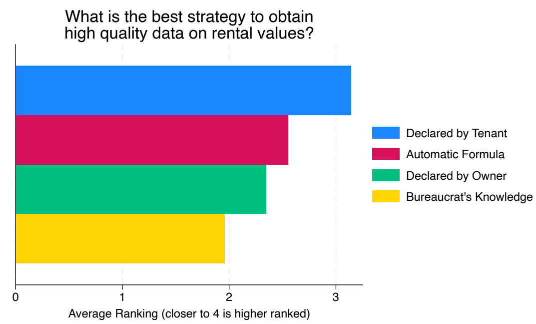
(C) HIGH VALUE PROPERTY - DISAGGREGATED BY INFORMATION TREATMENT

Notes: This Figure shows the distribution of responses from the lab-in-the-field experiment in which bureaucrats were asked to value two properties based on pictures. The low value (respectively high value) property is depicted in Panel (A) (resp. Panel (C)) of Figure 1. We plot the histogram of $Ln(Value)$ where value is the monthly rental value provided by the bureaucrats. The vertical line indicates the benchmark market value obtained from licensed assessors. In Panel (C), we disaggregate responses for the high value property by randomized information treatment status: a bureaucrat is treated if she saw the information chart shown in Figure A14.

FIGURE A16
BUREAUCRATS' OPINION ON CONDITIONS FOR SUCCESSFUL PROPERTY TAX CENSUS



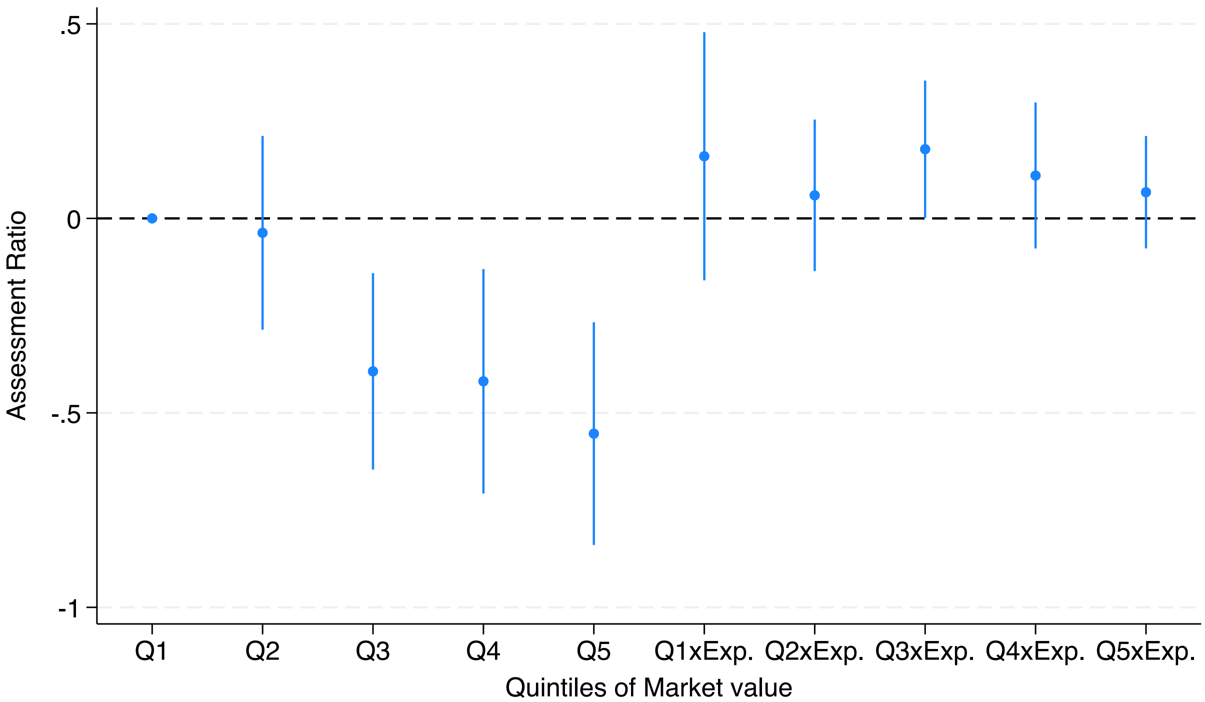
(A) MOST IMPORTANT SKILLS



(B) BEST STRATEGY TO RECOVER ACCURATE VALUES

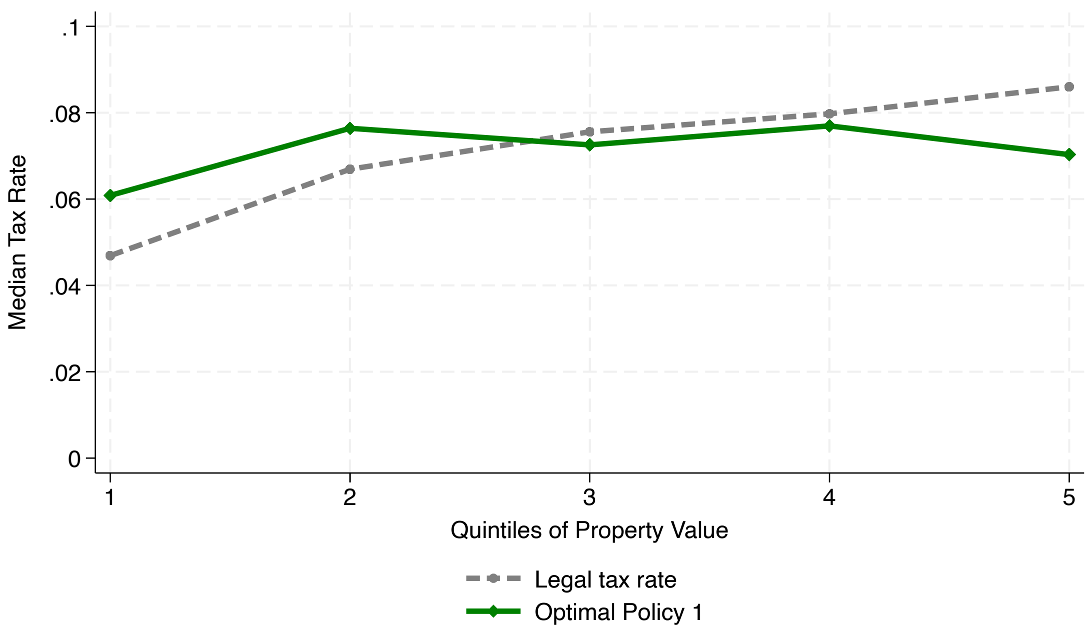
Notes: This Figure shows results from two survey questions from the bureaucrat endline survey. In Panel (A), respondents were asked: "According to you, what are the most important skills to be successful in the property tax census?". They were then shown a list of seven skills that they needed to rank from most to least important. We display the average ranking obtained by each skill, a higher value means the skill was ranked as more important on average. In Panel (B), the question is: "According to you, what is the best strategy for the DGID to recover high quality data on rental values in the region?". Respondents were then shown a list of four strategies that they needed to rank from most to least important. We display the average ranking obtained by each strategy, a higher value means it was ranked as more important on average.

FIGURE A17
HETEROGENEITY IN UNDERVALUATION GRADIENT BY EXPOSURE TO THE RULE



Notes: This Figure shows how the assessment ratio under discretion varies by quintile and depending on whether the bureaucrat was ever exposed to the rule yet. We plot the β_n and γ_n coefficients from the regression: $AR_i = \alpha + \sum_{n=1}^{5} \beta_n Q(n)_i + \sum_{n=1}^{5} \gamma_n Exposed_i \cdot Q(n)_i + \epsilon_i$, where AR_i is the assessment ratio (tax roll value over market value) for property i, the $Q(n)$ are dummies for each quintile of the distribution of market values, $Exposed_i$ is a dummy taking value one if the bureaucrat has already been exposed to the rule previously. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. Errors are clustered at the section level. Sample: analysis sample, i.e., properties for which we also have market values, restricted to the discretionary arm ($N = 1,124$).

FIGURE A18
TAX RATES UNDER OPTIMAL POLICY



Notes: This Figure shows the median effective tax rate by quintile of market value under the optimal policy described in Section 7. We use section fixed-effect and built area to predict whether a property belongs to the lowest quintile. If $Pred(Q1) = 1$, discretionary valuation is used. If $Pred(Q1) = 0$, the pure rule is applied. A property's effective tax rate is computed as tax liability based on tax roll value over market value. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. The gray line is the benchmark tax profile computed by applying the tax code rates directly to market values (the tax rate is 8.6 percent with an abatement if the property is owner-occupied which explains the progressive profile). Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm).

Additional Tables

TABLE A1
CREATION OF THE MARKET VALUES DATASET BY REAL ESTATE ASSESSORS

Panel A: Details on assessors' field work	
N sections	193
- per assessor	24.1
Info. from Office (%)	97.4
Info. from Agencies (%)	55.4
Info. from Occupants (%)	67.4
Panel B: Correlation with other sources of rental values	
Owner survey	0.39 (N=1,310)
Owner survey (rented)	0.49 (N=394)
Owner survey (fully rented)	0.62 (N=52)
Census (fully rented, met tenant)	0.50 (N=212)
Census (fully rented, met tenant, contract)	0.59 (N=48)
Census (full contract)	0.72 (N=63)
Census (full contract, met tenant)	0.83 (N=19)

Notes: In this Table, we provide additional details on the creation of the market values dataset with licensed real estate assessors. In Panel (A), we report the share of sections covered by the assessors for which they report having used information respectively from: previous work done by their office; real estate agencies they contacted; occupants (tenants). In Panel (B), we show how the market values provided by assessors correlate with values from other sources, at the plot level. *Owner survey* refers to our property owner baseline survey, *rented* means the property is rented at least in part, and *fully rented* means the whole property is rented. *Census* refers to the data from the property tax census. *Met tenant* means the bureaucrat reported the name of the tenant and did not report meeting the owner. *Contract* means the bureaucrat reported the existence of at least one rental contract on the property, *full contract* means the bureaucrat reported the existence of a rental contract (or multiple contracts) covering the whole property.

TABLE A2
PERFORMANCE STATISTICS: PROPERTY VALUATION ALGORITHM USING ALL COVARIATES

Estimated via 5x cross-validation	
R2	0.91
Adjusted R2	0.90
RMSE	0.36
Estimated on test sample	
MAPE	33.8
MAPE Q1	41.4
MAPE Q2	29.9
MAPE Q3	27.4
MAPE Q4	38.9
MAPE Q5	29.8
Freddie Mac 30%	59.6

Notes: This Table reports performance statistics for the property valuation algorithm, using all covariates listed in Table A3, and following specification 1. We calibrate the algorithm on the sample of market values and characteristics collected by assessors ($N = 4,921$), using an elastic-net regression and 5-fold cross-validation, as explained in Section 4.2. We hold out a test sample of 521 observations to test model fit. Results are reported in the lower panel. MAPE refers to the mean absolute percentage error (it is computed using monetary amounts as the outcome variable, *not* on the ln() transformation). We report the MAPE within each quintile. Freddie Mac refers to the share of predictions that fall within 30% of the true value.

TABLE A3
COEFFICIENTS: PROPERTY VALUATION ALGORITHM USING ALL COVARIATES

Ln(BuiltArea)	0.57
Floors	.178
Residential	
Commercial	.195
Mixed	.119
Quality Doors and Windows	
Very Good	.116
Average	
Bad	-.199
Landscape	-.082
Architecture	.044
Garage	
Simple	.074
Double	.148
None	
Balcony	.164
On Main Road	.043
Near Main Road	.012
Off Main Road	
Road Type	
Tarmac	.007
Pavements	.028
Gravel	.077
Sand	
None	0
Sidewalk	.059
Angle	.111
Street Lights	0

Fence Type	
None	
Metal	-.167
Wall	.023
Wall w. wrought iron	.01
Fence State	
Very Good	.043
Average	
Bad	-.064
Cement	.141
Cladding Type	
Wis	-.074
Plain	-.118
Paint	
Tiles	0
Stone	.073
None	-.229
Cladding State	
Very Good	.077
Average	
Bad	-.124
Tiles	.049
Cons	11.712
Section FEs	
N zero	22/193
mean	.134
max	2.317
sd	.52

Notes: This Table reports the coefficients from the property valuation algorithm summarized in Table A2, following specification 1. We calibrate the algorithm on the sample of market values and characteristics collected by assessors ($N = 4,921$). In the implementation of the program in the rule-based arm, built area and section fixed-effects are pre-loaded in the application, while the other characteristics are entered in the field.

TABLE A4
PROPERTY VALUATION ALGORITHM USING REMOTE COVARIATES

Estimated via 5x cross-validation	
R2	0.88
Adjusted R2	0.87
RMSE	0.43
Estimated out of sample	
MAPE	41.4
MAPE Q1	65.7
MAPE Q2	33.7
MAPE Q3	25.7
MAPE Q4	42.7
MAPE Q5	34.0
Freddie Mac 30%	54.2

(A) PERFORMANCE STATISTICS

Ln(BuiltArea)	0.68
Floors	.214
Cons	11.317
Section FEs	
N zero	11/193
mean	.179
max	2.261

(B) COEFFICIENTS

Notes: This Table reports performance statistics (Panel (A)) and coefficients (Panel (B)) of the pure rule property valuation algorithm, which uses only remote covariates (section fixed-effects, built area, number of floors). We calibrate the algorithm on the sample of market values by assessors ($N = 4,921$), using an elastic-net regression and 5-fold cross-validation, as explained in Section 4.2. We hold out a test sample of 521 observations to test model fit. MAPE refers to the mean absolute percentage error (it is computed using monetary amounts as the outcome variable, *not* on the ln() transformation). We report the MAPE within each quintile. Freddie Mac refers to the share of predictions that fall within 30% of the true value.

TABLE A5
THE UNDERVALUATION GRADIENT: ROBUSTNESS RESULTS UNDER DISCRETION

	(1) Q1	(2) Q2	(3) Q3	(4) Q4	(5) Q5
Panel A: Discretion					
Mean Ass. Ratio	1.15	0.92	0.57	0.56	0.47
$\hat{\beta}_n$	Ref.	-0.23	-0.58	-0.59	-0.69
P-value		0.05	0.00	0.00	0.00
P-value $\hat{\beta}_n \neq \hat{\beta}_{n+1}$		0.00	0.88	0.39	
Panel B: Values per m^2					
Median Ass. Ratio	0.89	0.69	0.54	0.50	0.23
$\hat{\beta}_n$	Ref.	-0.19	-0.35	-0.39	-0.66
P-value		0.04	0.00	0.00	0.00
P-value $\hat{\beta}_n \neq \hat{\beta}_{n+1}$		0.04	0.52	0.00	
Panel C: Quintiles computed out-of-sample					
Median Ass. Ratio	0.75	0.60	0.39	0.44	0.35
$\hat{\beta}_n$	Ref.	-0.15	-0.36	-0.31	-0.40
P-value		0.04	0.00	0.00	0.00
P-value $\hat{\beta}_n \neq \hat{\beta}_{n+1}$		0.00	0.54	0.24	

Notes: This Table shows robustness results as a complement to Table 2, plotting coefficients β_n from regression: $AR_{ij} = \alpha + \sum_{n=1}^{5} \beta_n Q(n)_{ij} + \epsilon_{ij}$ where AR_{ij} is the assessment ratio (tax roll value over market value) for property i of section j, and the $Q(n)$ are dummies for each quintile of the distribution of market value. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. Errors are clustered at the section level. In Panel (A) we run a standard OLS regression. In Panel (B), we run the quantile regression at the median, but the five quintiles used as regressors are defined using market property value *per square meter*. In Panel (C), quintiles are defined using *predicted* property values using a prediction calibrated on pure control areas only. Sample: analysis sample, i.e., properties for which we also have market values, restricted to the discretionary arm ($N = 1,124$).

TABLE A6
REMOVING DISCRETION INCREASES ACCURACY: ROBUSTNESS

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) ∣Gap∣ *mil.FCFA*	(4) Ass. Ratio
Panel A: Rule-based (with bur. FEs)				
Mean[1] (sd)	-1.25 (10.80)	0.00	3.81 (10.18)	1.16
$\hat{\beta}_{Discretion}$	-5.55^{***}	-2.41^{***}	4.49^{***}	-0.39^{***}
	(1.29)	(0.50)	(1.24)	(0.06)
Mean[1] (sd) Discretion	-6.11 (15.62)	-2.40	8.00 (14.73)	0.76
N plots: 2118				
N Bureaucrats: 234				
Panel B: Rule-based (calibrated on owner survey)				
Mean[1] (sd)	-6.56 (18.05)	-1.72	6.98 (17.89)	0.76
$\hat{\beta}_{Discretion}$	-0.48	-0.75^{**}	1.45	-0.04
	(1.92)	(0.35)	(1.92)	(0.05)
Mean[1] (sd) Discretion	-7.54 (18.98)	-2.41	8.96 (18.35)	0.71
N plots: 2290				
Panel C: Pure rule (calibrated on owner survey)				
Mean[1] (sd)	-6.66 (18.05)	-1.77	7.10 (17.88)	0.75
$\hat{\beta}_{Discretion}$	-0.39	-0.73^{**}	1.34	-0.03
	(1.92)	(0.36)	(1.91)	(0.05)
Mean[1] (sd) Discretion	-7.54 (18.98)	-2.41	8.96 (18.35)	0.71
N plots: 2290				

Notes: This Table shows robustness results on the effect of discretion on the tax base gap, as a complement to Table 3. In Panel (A), we show the effect of discretion within bureaucrat: we control for bureaucrat fixed-effects. In Panel (B), for the rule-based arm, we use a rule calibrated on the self-reported values from our baseline property owner survey. In Panel (C), for the rule-based arm, we use a pure rule (only built area and section fixed-effects) calibrated on the self-reported values from our baseline property owner survey. We run regression 2: $Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk}$ with four different outcomes. Y_{ijk} is the outcome for property i of section j and strata k, D is a dummy for discretionary sections and S_k is a strata fixed-effect. In column (1) the outcome variable is the tax base gap defined as tax roll value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as tax roll value over the market value. In Panel (A), the number of observations drops to $2,191$, since plots that are not covered by the census are not assigned a bureaucrat identifier. In each sub-panel, the first row displays descriptive statistics of the outcome variable in the rule-based arm; the second row shows the coefficient of interest and its standard error. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed-effects and errors are clustered at the section level. Sample: analysis sample, i.e., properties from both arms for which we also have market values ($N = 1,124$ in the discretionary arm and $N = 1,166$ in the rule-based arm). [1]In column (2) the displayed value is the $median$ of the tax base gap.

TABLE A7
REMOVING DISCRETION INCREASES ACCURACY: INTENSIVE MARGIN

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) ∣Gap∣ *mil.FCFA*	(4) Ass. Ratio
Panel A: Discretion				
Mean[1] (sd)	−4.45 (15.39)	−1.32	6.89 (14.46)	0.97
Panel B: Rule-based				
Mean[1] (sd)	−1.25 (10.80)	0.00	3.81 (10.18)	1.16
$\hat{\beta}_{Discretion}$	-2.98^{***}	-1.11^{***}	2.65^{**}	-0.21^{***}
	(1.08)	(0.26)	(1.13)	(0.05)
Panel C: Pure Rule				
Mean[1] (sd)	−0.13 (6.77)	0.12	2.62 (6.24)	1.13
$\hat{\beta}_{Discretion}$	-3.78^{***}	-1.34^{***}	3.71^{***}	-0.18^{***}
	(0.88)	(0.30)	(0.89)	(0.05)
Panel D: Lee bounds				
Lower bound	−5.83		2.07	−0.37
Upper bound	−1.42		5.73	0.10
CI for $\hat{\beta}_{Discretion}$	[−6.74;−0.32]		[1.00;6.56]	[−0.45;0.17]
N plots: 1885				
N Sections: 94				
Mean (sd) market value: 73.90 (14.40)				
Median market value: 5.40				

Notes: This Table shows the effect of discretion on the tax base gap, similarly to Table 3, except that in this case we focus on the intensive margin: we drop plots for which the tax roll value is zero ($N = 490$). This can occur because the plot was not visited, or because the plot was visited but the bureaucrat did not assign a value (the latter is only arises under discretion). We run regression 2: $Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk}$ with four different outcomes. Y_{ijk} is the outcome for property i of section j and strata k, D is a dummy for discretionary sections and S_k is a strata fixed-effect. In column (1) the outcome variable is the tax base gap defined as tax roll value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as tax roll value over the market value. In Panel (A), we display summary statistics for the discretionary arm. In Panel (B), values for the rule arm are the rule-based valuations incorporating bureaucrats' inputs. In Panel (C), values for the rule arm are the pure rule predictions based on remote covariates only. The first rows of Panels (B) and (C) display descriptive statistics of the outcome variable in the rule-based arm; the second rows show the coefficient of interest and its standard error. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed-effects and errors are clustered at the section level. In Panel (D), we conduct a Lee bound estimation to correct for non-random attrition. Rows 1 and 2 display the estimated lower and upper bound of the treatment effect, and row 3 shows the final confidence interval of the effect, accounting for uncertainty coming both from non-random attrition and from sampling error. Sample: analysis sample, i.e., properties from both arms for which we also have market values, restricted to positive tax roll values ($N = 1,871$). [1]In column (2) the displayed value is the *median* of the tax base gap.

TABLE A8
HORIZONTAL AND VERTICAL EQUITY STATISTICS

	(1) Market Values	(2) Discretion	(3) Rule-based	(4) Pure Rule	(5) Rule (calibration inputs)
Mean Value	15.80	7.21	12.71	17.09	15.72
SD	76.98	37.52	88.00	112.13	92.67
Median Value	5.55	2.40	4.27	4.86	5.02
Mean Ass. Ratio		0.71	1.06	1.13	1.10
Median Ass. Ratio		0.50	0.95	1.04	1.05
Share Accurate		0.22	0.45	0.63	0.68
PRD		1.60	1.28	1.02	1.02
COD		114.75	52.86	31.05	25.50
Q1		103.34	54.91	29.02	25.17
Q2		86.07	45.67	27.40	23.05
Q3		84.50	41.90	25.37	22.71
Q4		99.60	37.78	26.25	22.19
Q5		188.53	64.51	40.74	35.84

Notes: This Table reports summary statistics on valuations with each method, as well as specific indicators measuring the horizontal and vertical equity of valuations. Column (1) provides statistics on market values obtained from licensed assessors. Column (2) provides statistics on values from the discretionary arm. Column (3) provides statistics on valuations in the rule-based arm. Column (4) provides statistics on valuations in the rule arm when applying the pure rule. Column (5) reports on valuations in the rule arm when using predictions based on the inputs used for the calibration, from the assessors' dataset (instead of bureaucrats' inputs). Values are annual property rental values in millions of FCFA. Rows four and five show the mean and median assessment ratio (AR), computed as tax roll value over market value. Row six shows the share of properties accurately valued, meaning that the tax roll value is within 30% of the market value. The PRD is the Price Related Differential, a measure of vertical equity. The PRD is calculated as the mean AR divided by the weighted mean AR (weighted by market value). A PRD lower than one indicates that valuations are progressive, while the higher above one the value is, the more regressive valuations are. The COD is the coefficient of dispersion, a measure of horizontal equity (more precisely, uniformity of valuations). It is measured as average percentage dispersion of AR values around the median AR. Finally, we report the COD computed within each quintile of market values.

TABLE A9
EFFECT OF DISCRETION ON EXTENSIVE MARGIN

(A) FULL SAMPLE

Dependent Variable	Plots per day	Covered	Eligible	Valued	Rented	Main Res.	Rent value	Main Res. value	Owner Met	Conflict
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Discretionary	1.156	0.002	-0.031	-0.228***	-0.006	-0.051**	-0.069***	-0.347***	0.002	-0.013*
	(1.306)	(0.019)	(0.021)	(0.023)	(0.017)	(0.021)	(0.020)	(0.032)	(0.014)	(0.007)
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2309	41609	41609	41609	41609	41609	41609	41609	41609	41609
N sections	94	94	94	94	94	94	94	94	94	94
Adj R2	0.05	0.03	0.06	0.12	0.05	0.04	0.05	0.19	0.02	0.02
Mean of dep.	16.56	0.92	0.79	0.71	0.38	0.46	0.36	0.57	0.24	0.05

(B) SAMPLE WITH MARKET VALUES

Dependent Variable	Plots per day	Covered	Eligible	Valued	Rented	Main Res.	Rent value	Main Res. value	Owner Met	Conflict
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Discretionary	-0.011	0.027	0.031	-0.178***	0.019	-0.020	0.001	-0.310***	0.001	-0.008
	(0.087)	(0.021)	(0.022)	(0.030)	(0.023)	(0.028)	(0.024)	(0.040)	(0.025)	(0.010)
Strata FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1104	2409	2409	2290	2290	2290	2290	2290	2290	2290
N sections	94	94	94	94	94	94	94	94	94	94
Adj R2	0.03	0.03	0.03	0.10	0.05	0.04	0.05	0.18	0.02	0.02
Mean of dep.	2.03	0.93	0.88	0.82	0.43	0.54	0.40	0.67	0.26	0.05

Notes: This Table shows the effect of discretion on extensive margin outcomes. Panel (A) shows results for the whole sample while Panel (B) is restricted to plots for which we have market values. The coefficient of interest is the one on an indicator for discretionary sections. In column (1) the outcome is the number of plots covered in a day and observations are at the day X section level. In columns (2) to (10) regressions are at the plot level and the outcome variables are dummies taking value one, respectively, if: the plot is covered in the census, it is classified as eligible for the tax, there is a positive property value, it is classified as rented (at least in part), it is classified as main residence, there is a positive value for rented parts, there is a positive value for owner-occupied parts, the bureaucrat reports having met the owner, the bureaucrat reports that there has been tensions or conflict (scraped from text comments). In all regressions we control for strata fixed-effects and errors are clustered at the section level.

TABLE A10
BUREAUCRATS' IMPLICIT ALGORITHM

Ln(BuiltArea)	0.41
Floors	.166
Residential	
Commercial	.012
Mixed	.086
Quality Doors and Windows	
Very Good	0
Average	
Bad	-.124
Landscape	-.057
Architecture	0
Garage	
Simple	.096
Double	.06
None	
Balcony	0
On Main Road	.07
Near Main Road	.06
Off Main Road	
Road Type	
Tarmac	0
Pavements	0
Gravel	.338
Sand	
None	0
Sidewalk	.002
Angle	.093
Street Lights	.099

Fence Type	
None	
Metal	0
Wall	0
Wall w. wrought iron	0
Fence State	
Very Good	0
Average	
Bad	0
Cement	0
Cladding Type	
Wis	-.298
Plain	-.096
Paint	
Tiles	0
Stone	0
None	-.084
Cladding State	
Very Good	0
Average	
Bad	-.018
Tiles	0
Cons	12.549
Section FEs	
N zero	16/48
mean	.095
max	1.249
sd	.349

Notes: This Table reports coefficients for bureaucrats' implicit valuation algorithm. We follow the exact same methodology as described in Section 4.2, but using bureaucrats' discretionary values as the outcome variable, on the sample of properties from the discretionary arm. We use the observable characteristics reported by the assessors as regressors, since bureaucrats do not report observable characteristics in the discretionary arm. Sample: $N = 1,203$.

Table A11
Rule-based valuation: observable characteristics reported by bureaucrats vs assessors

Characteristic	Share Identical
Easy/objective	
Floors	0.71
when $\neq$ % where bur. = ass. $+1$	0.56
Usage	0.75
Wall (dummy)	0.79
Tiles (dummy)	0.65
Balcony (dummy)	0.78
Angle (dummy)	0.92
Street lights (dummy)	0.79
Garage	0.62
Road type	0.71
Complex/subjective	
Fence type	0.55
Fence state	0.51
Cladding type	0.14
Cladding state	0.48
Quality doors and windows	0.49
Landscape (dummy)	0.81
Architecture (dummy)	0.66
Road (main)	0.50
Sidewalk (dummy)	0.65

Notes: In this Table, we compare observable characteristics reported by assessors with those reported by bureaucrats. For the number of floors: additionally we indicate the percentage of cases for which the bureaucrat reported one floor more than the assessor, among cases where there is a mismatch. Sample: analysis sample, i.e., properties for which we also have market values, restricted to the rule-based arm ($N = 1,166$)

TABLE A12
LEARNING OVER TIME

Dependent Variable	Discretion |Gap|	 |Gap|	 Value	 |Gap|	Rule |Gap|	 Value
	(1)	(2)	(3)	(4)	(5)	(6)
Numb. properties	0.050*			-0.024*		
	(0.026)			(0.014)		
(Numb. properties)^2	-0.000			0.000*		
	(0.000)			(0.000)		
Numb. days		0.342**	0.011		-0.208**	-0.030
		(0.152)	(0.021)		(0.085)	(0.027)
(Numb. days)^2		-0.004	-0.001***		0.002**	0.000
		(0.003)	(0.000)		(0.001)	(0.000)
Section control for Market Value	Yes	Yes	Yes	Yes	Yes	Yes
Bureaucrat FE	Yes	Yes	Yes	Yes	Yes	Yes
N	1055	1055	20086	1063	1063	17458
Mean	8.00	8.00	3.97	3.81	3.81	5.52
R2	0.51	0.51	0.15	0.44	0.44	0.33
Adj R2	0.39	0.39	0.14	0.31	0.31	0.32

Notes: In this Table, we assess whether there is any learning by bureaucrats over the course of the property tax census. We do so by analyzing whether the tax base gap changes over time or with the number of properties covered. Columns (1) to (3) show results for the discretionary arm while columns (4) to (6) show results for the rule-based arm. On average, a bureaucrat worked 32 days, and covered 142 properties. In columns (1), (2), (4) and (5), the outcome is the absolute value of the tax base gap defined as tax roll value minus market value. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll. The sample is restricted to the analysis sample, i.e., properties for which we also have market values. In columns (3) and (6) the outcome is tax roll value and the regression includes all properties of a given arm. All outcome variables are in millions of FCFA and winsorized at the 1% level. The regressors of interest are number of properties visited by a given bureaucrat and its squared value (columns (1) and (4)), and total number of days worked by a given bureaucrat and its squared value. We control for bureaucrat fixed-effects and a section level decile of market values, errors are clustered at the section level. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively.

TABLE A13
VALUATIONS AND OWNER CHARACTERISTICS

Dependent Variable	Value (winz.)	Ln(Value)	Value (winz.)	Ln(Value)
	(1)	(2)	(3)	(4)
Ln(remote pred.)	3.986***	0.705***	3.117***	0.662***
	(0.340)	(0.028)	(0.652)	(0.087)
Bureaucrat Met Owner	-0.600***	-0.159***		
	(0.197)	(0.024)		
Deceased	-0.927***	-0.278***		
	(0.288)	(0.040)		
Female	0.180	-0.059**		
	(0.185)	(0.023)		
Male	0.000	0.000		
	(.)	(.)		
Unknown	-1.677***	-0.045		
	(0.350)	(0.042)		
Multiple	6.511***	0.343**		
	(1.860)	(0.159)		
Owner is retired	-0.041	-0.070***		
	(0.166)	(0.024)		
Agents' Commune	1.294**	0.050		
	(0.563)	(0.089)		
Same ethnic group			0.087	0.113
			(0.370)	(0.071)
Missing			-0.591	-0.012
			(0.676)	(0.129)
Poor			0.000	0.000
			(.)	(.)
Rich			-1.265	-0.065
			(0.758)	(0.113)
Not employed			-0.303	-0.139
			(0.627)	(0.091)
Employed			0.000	0.000
			(.)	(.)
Retired			-0.257	-0.242**
			(0.690)	(0.105)
Female bur and owner			-0.508	0.158
			(0.979)	(0.112)
Female bur and male owner			0.651	0.048
			(0.587)	(0.124)
Male bur and female owner			0.999	0.255**
			(0.791)	(0.118)
Rented			0.756	0.083
			(0.467)	(0.094)
In Tax Net			0.278	0.001
			(0.595)	(0.099)
Strata FE	Yes	Yes	Yes	Yes
N	14451	11242	505	413
R2	0.31	0.50	0.41	0.51
Adj R2	0.31	0.50	0.35	0.46

Notes: In this Table, we test whether owner characteristics correlate with tax roll values in the discretionary arm. Columns (1) and (2) include all plots from the discretionary arm. The owner characteristics come from the census. In columns (3) and (4), we restrict the sample to properties covered by the property owner baseline survey, which is the source of the owner characteristics. We control for $Ln(remotepred)$, which is the predicted property value based solely on remote covariates. In all columns, we control for randomization strata. Errors are clustered at the section level.

TABLE A14
HETEROGENEITY DEPENDING ON HOW THE FIELD VISIT GOES

	No Bureaucrat FE			with Bureaucrat FE		
	(1)	(2)	(3)	(4)	(5)	(6)
	Value	|Gap|	Ass. Ratio	Value	|Gap|	Ass. Ratio
Panel A: Owner Met						
$\hat{\beta}_{Discretion}$	-2.76^{***}	1.93	-0.47^{***}	-3.69^{***}	1.24	-0.43^{***}
	(0.78)	(1.37)	(0.07)	(1.09)	(1.32)	(0.10)
Mean (sd) in Rule	7.21 (7.27)	4.07 (13.29)	1.17 (0.77)	7.21 (7.27)	4.07 (13.29)	1.17 (0.77)
N	597	597	340	597	597	340
Low Value						
$\hat{\beta}_{Discretion}$	-0.58	0.27	-0.31^{**}	-0.08	1.01^{*}	0.17
	(0.39)	(0.30)	(0.14)	(0.68)	(0.52)	(0.24)
Mean (sd) in Rule	3.00 (1.58)	1.07 (1.10)	1.38	3.00 (1.58)	1.07 (1.10)	1.38
N	257	257	257	257	257	257
High Value						
$\hat{\beta}_{Discretion}$	-4.70^{***}	2.98	-0.49^{***}	-5.16^{***}	4.01^{***}	-0.54^{***}
	(1.18)	(2.26)	(0.07)	(1.51)	(1.51)	(0.08)
Mean (sd) in Rule	10.80 (8.24)	6.64 (17.68)	0.99	10.80 (8.24)	6.64 (17.68)	0.99
N	340	340	340	340	340	340
Panel C: Rented						
$\hat{\beta}_{Discretion}$	-1.19	3.35^{**}	-0.16^{**}	-1.70	4.75^{**}	-0.20^{*}
Mean (sd) in Rule	8.98 (8.80)	4.33 (10.45)	1.18 (0.87)	8.98 (8.80)	4.33 (10.45)	1.18 (0.87)
N	977	977	977	977	977	977
Low Value						
$\hat{\beta}_{Discretion}$	0.36	0.72^{*}	0.13	0.94	1.18	0.51
	(0.49)	(0.39)	(0.19)	(1.00)	(0.80)	(0.32)
Mean (sd) in Rule	3.61 (2.82)	1.44 (2.30)	1.52	3.61 (2.82)	1.44 (2.30)	1.52
N	327	327	327	327	327	327
High Value						
$\hat{\beta}_{Discretion}$	-2.21^{*}	4.61^{**}	-0.23^{***}	-3.26^{*}	6.72^{**}	-0.31^{***}
	(1.15)	(2.13)	(0.06)	(1.70)	(2.96)	(0.10)
Mean (sd) in Rule	12.16 (9.56)	6.05 (12.78)	0.97	12.16 (9.56)	6.05 (12.78)	0.97
N	650	650	650	650	650	650
Panel C: Bureaucrats' estimate						
$\hat{\beta}_{Discretion}$	-2.15	2.41	-0.26^{***}	-1.65	4.00^{**}	-0.24^{**}
Mean (sd) in Rule	7.50 (8.21)	3.81 (10.18)	1.16 (0.77)	7.50 (8.21)	3.81 (10.18)	1.16 (0.77)
N	1,193	1,193	1,193	1,193	1,193	1,193
Panel D: Conflict						
$\hat{\beta}_{Discretion}$	-3.04	0.06	-0.55^{***}	-5.61	-2.79	-0.64^{***}
Mean (sd) in Rule	7.91 (8.16)	5.70 (15.87)	0.98 (0.55)	7.91 (8.16)	5.70 (15.87)	0.98 (0.55)
N	110	110	110	110	110	110

Notes: This Table reports results from regression 2: $Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk}$, restricting the sample to specific subsets of observations. Y_{ijk} is the outcome for property i of section j and strata k, D is a dummy for discretionary sections and S_k is a strata fixed-effect. In Panel (A), the discretionary sample is restricted to cases where the bureaucrat reports having determined the value herself (scraped from text comments); in Panel (B), for both arms, we restrict the samples to cases where the bureaucrat reports meeting the owner; in Panel (C), for both arms, we restrict the sample to cases where the bureaucrat indicated that the property was rented at least in part; in Panel (D), for both arms, we restrict the sample to cases where the bureaucrat reports that there was some tensions or conflict (scraped from text comments). In columns (1) and (4) the outcome variable is tax roll value, in columns (2) and (5), the outcome variable is the absolute value of the tax base gap defined as tax roll value minus market value. All amounts are in millions of FCFA and winsorized at the 1% level. In columns (3) and (6) the outcome is the assessment ratio, computed as tax roll value over the market value. The market value is the value obtained from licensed assessors. Tax roll value is the value from the census that ends up on the tax roll.

TABLE A15
COST-BENEFIT ANALYSIS

	Costs	Tax liabilities	Ratio
Assessors	503.3 mn	38.5 bn	x76
Discretion	118.1 mn	15.7 bn	x133
Rule	134.7 mn	22.4 bn	x166
Pure Rule	16.6 mn	40 bn	x2416
Optimal Policy	17.4 mn	45 bn	x2588

Notes: This Table displays costs, total tax liabilities, and the liabilities-to-cost ratio under each method. Costs correspond to field costs and algorithm calibration costs, and excluded program costs that are invariant across methods (such as software development). Tax liabilities are total liabilities assuming full compliance. Costs and liabilities are calculated assuming each method is in turn applied to all eligible plots of the 96 study sections ($N = 32,677$).

A Details on Institutional Context and Program

A.1 Institutional Context

The creation of the property tax valuation roll and the tax bills is the responsability of the national tax administration, *Direction Générale des Impôts et Domaines du Sénégal* or *DGID*. Once emitted, the tax bills are physically distributed by the national treasury, and owners pay in a treasury office. Revenues accrue to municipalities at the end of the year. The region includes four cities: Dakar, Guediawaye, Pikine, Rufisque. Each city is divided in municipalities or *Communes*. Each municipality is divided into cadastral sections. In our paper, we call *the property tax* the combination of the property tax and the garbage tax. Indeed, both are managed through the same tax bills and have the same base. Their exact denominations are *Contribution Foncière des Propriétés Bâties* or *CFPB* and *Taxe d'Enlèvement des Ordures Ménagères* or *TEOM*. The garbage tax has a rate of 3.6 percent. The property tax has a 5 percent rate, and if the property is the owner's main residence, there is a reduction of the tax base by 1.5 million CFAF. The discretionary valuation method which is the status quo before the program is called the *comparative method* in the tax code (meaning that bureaucrats should compare the property with similar ones for which they have values in mind). The tax code also provides for a *cadastral valuation method* where experts from the cadastral division of the administration conduct in depth visits and measurements to value a property. Due to length and cost, in practice, these inspections barely ever occur. Resource constraints are the main reason why property census operations were extremely rare before the program. The staff available to conduct this field work is limited and also works on other taxes, the registrations were extremely time consuming since all the information was collected on paper and needed to be typed into the system once back into the office. Several factors explain why there is now this push by the government to expand the property tax net. First, in 2013, a decentralisation bill was passed reshaping the responsibilities of local governments and requiring that they assist the national tax administration in enforcing local taxation. These decentralisation policies have also made it more crucial to grow the budgets of local governments, now responsible for more services. These incentives are also built in Senegal's relationship with international donors: starting 2019, a budget support agreement with the World Bank and bilateral donors set objectives in terms of local tax revenues and municipalities' capacity, disbursements are conditional on increased property tax registrations. Furthermore, the government is interested in expanding the tax net in general, and due to its intrinsic geographical component, the property tax is considered as a promising entry point to build better information sets on taxpayers' wealth, income, and address (none of the existing tax registers have precise addresses). As an illustration, in 2021, the tax administration launched both its *Yaatal* (or "expansion") program, aiming to double the number of registered taxpayers, and a national property census project (*Recensement National des Propriétés Bâties*), with the objective for all properties of the national territory to be covered by property tax census operations in the near future. Finally, in this context, some resource constraints were lifted thanks to the support of the African Development Bank encouraging the modernization of public administration, these funds were instrumental in the financing of the program.

A.2 Program

The application was developed starting 2017 by the administration, a private Senegalese company, and the research team. It has a Web components (to assign tasks, visualize and validate information, monitor advancement on maps, create tax bills), and an android component allowing to conduct the census on tablets in the field with pre-loaded maps. The section and plot identification system integrated in the software were already used by the administration since 2012 (introduction of the unique cadastral plot identification number *NICAD*), however it is the first time this mapping was digitized and incorporated in an interactive application. The bureaucrats working on the census were hired through several channels: some had already done similar tasks for the administration in the past, some had been suggested by the municipalities, finally, many were recruited through an online job advertisement on a public employment platform. They receive a four day training delivered by the tax administration, covering local public finance concepts, the utilization of the application, reading maps, property characteristics, and interactions with occupants. In the field, they are equipped with caps, shirts and badges showing their affiliation to the tax administration. They are paid a monthly fee and a bonus based on the number of plots they cover, and the share of plots for which they recover the name and identification details of the owner. Thus the incentives are exactly the same across arms. There is a supervisor for every fourteen bureaucrats on average. The supervisors were hired through the same channels. They need to validated the forms submitted by the bureaucrats before these are sent to the server. Based on responses in the bureaucrat endline survey, 61 percent of bureaucrats completely agree and 35 percent agree with the statement *the administration closely verifies the data I collect*.

B Data: Additional Details

B.1 Market Values by Assessors

The assessors we hired are licensed real estate experts, who work in the private sector and are affiliated to the Senegalese National Order of Experts.[92] Their usual job is to provide certified market valuations of properties for insurance purposes, before a sale or a inheritance, in relation to construction projects, etc. To design the data collection, we relied on discussions with practitioners to build on methodologies used in more established property tax systems such as the United Kingdom and South Africa. Before starting the field work in a given section, the assessors were asked to gather location specific information from their office as well as from real estate agencies and brokers they are in close contact with, we show the frequency with which this occurs in Appendix Table A1. For each property, we pre-load the built area measurement in the questionnaire on tablet to help the assessors in their valuations.

The sampling of properties to be visited by assessors was done in way to allow partial overlap between assessor valuations and (i) the baseline property owner survey; (ii) plots covered by the census for which a rental contract is noted (this is very rare, 2.6 percent of census observations). The rationale for doing so is to be able to check correlations between different sources of rental values. We first draw 26 plots randomly in each section. If less than 13 plots also included in the owner baseline were drawn, we do some replacements to reach 13/26 overlapping plots (or the maximum number possible if there are fewer baseline plots in the section). If less than 2 plots covered by the census and with a rental contract were drawn, we do some replacements to reach 2/26 of these (or the maximum number possible). Then we add a random draw of replacement plots for each section. As a result, among the 5,806 plots sampled for assessor valuation, 1,383 were covered in the property owner survey, and 138 were covered by the tax census and a rental contract was reported.

B.2 Property Valuation Algorithm

Selection and coding of characteristics. We selected the observable characteristics to be used in the property valuation algorithm by drawing from on existing methodologies of the cadaster department, and through work sessions bringing together members of the admiistration, the research team, and international practitioners. As much as possible, the phrasing of the characteristics and of their different modalities were preserved from pre-existing forms. The retained characteristics are: usage (residential, commercial or mixed), type of fence (four options), state of the fence (very good, average, bad), type of cladding (six options), state of the cladding (very good, average, bad), cement ('hard') wall (yes or no), presence of decorative tiles (yes or no), quality of doors and windows (very good, average, bad), landscape improvement (yes or no), architectural improvement (yes or no), garage (simple, double or none), balcony (yes or no), location with respect to main road (on, near, off), type of road (five options), presence of sidewalk (yes or no), whether the

[92] See: https://www.experts-ones.com/

property is at an angle (yes or no), presence of street lights (yes or no). The characteristics that have 'yes' or 'no' answers are coded as dummy variables. The characteristics that have multiple choice answers are all coded as categorical variables (using dummies for each modality), including characteristics related to state or quality. This enables us to be agnostic on the relative importance of each modality.

Calibration. For the functional form as well as the calibration details, we follow recent recommendations from the property valuation literature (Davis *et al.*, 2012; McCluskey *et al.*, 2013; Franzsen & McCluskey, 2017; Ali *et al.*, 2018; Fish, 2018; Guan *et al.*, 2011; Moore, 2005; International Association of Assessing Officers, 2022).[93] Out of our sample of 4,921 plots with market values, we randomly assign 10 percent of observations to the test sample (we draw within section to have observations from each section and recover section fixed-effects). Using cross-validation, we run an elastic net regression five times, and recover the median value of each coefficient. The elastic net is retained because we find that it performs better out-of-sample than the simple OLS and the Lasso regression. Next, we apply the resulting coefficients to the dataset to obtain predicted values, and identify outliers. Outliers are defined as predictions for which the residual is more than three standard deviations away from the mean value of residuals. These observations ($N = 55$) are dropped (following McCluskey *et al.* (2013)). We repeat the calibration of the elastic net regression with five-fold cross-validation. We recover the median value for each coefficient, these are our final coefficients. The R^2 and RMSE values reported in the performance statistics in Table A2 are the mean value of each statistics over the 5 iterations. The additional performance statistics reported in the lower panel of Table A2 are estimated on the test sample. See Figure A4 for a graphical analysis of predictions and residuals.

Computation of predicted values in monetary amounts. The algorithm predicts $\widehat{Ln(Value)}$. To compute predicted property value $\widehat{Value}$, a correction term needs to be applied to $exp(\widehat{Ln(Value)})$.[94] The corrected predicted value can be 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 prediction model is normally distributed, it can be shown that predicted values should be computed with $\alpha_c = exp(\frac{\hat{\sigma^2}}{2})$ 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. Based on the distribution of residuals (Panels (C) and (D) of Figure A4), we assume that the error term is normally distributed. We compute $\hat{\sigma^2}$ using the RMSE and find $\alpha_c = 1.07$. This is the correction term we use to recover predicted values in monetary amounts.

Computation of the sub-components of property value. The predicted value is at the plot level. We are interested in dividing this value into different sub-components in two instances: (i) when there are multiple owners on the plot – this is a rare feature in

[93]We had the opportunity to interact directly with some of these practitioners and experts between 2017 and 2022.

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

the Dakar real estate, less than 0.7 percent of plots in the census data – since each owner needs to receive a tax bill; (ii) when the property is partly rented and partly occupied by the owner (18.8 percent of plots in the census data) since in these cases the abatement only applies to owner-occupied parts. The administration has no information at all on built areas at a more precise level than the plot, therefore we rely on the number of rooms to divide plot value into its different components. If the property is partly rented and partly occupied by the owner: the corresponding values are computed as a share of total plot value, based on the number of rooms allocated to each usage. If there are multiple owners on the plot, each owner's value is computed based on her number of rooms out of total number of rooms. When we compare tax liabilities across arms, we provide a robustness check where the share of value subject to the abatement is computed in the exact same way in both arms, using number of rooms (see Figure A7 Panel (B)).

Implementation in the new digital tool. The implementation of the algorithm is integrated into the software. The integration of geocoded plot details and cadastral data into the software allows to automatically recover built area measurement and location fixed-effects. The software administrator can modify the coefficients associated to each characteristic, and also add, remove or modify observable characteristics used in the computation and that appear on the tablets in the field. This flexibility is key for the sustainable adoption of the digital tool by the administration. For this pilot phase, the calculation was done by the research team, in order to include the most recent updates in the algorithm and in the GIS dataset. The prediction will be automatically computed in the software starting 2024.

B.3 Data from the Property Tax Census

The information recorded in the field on tablets by the bureaucrats is automatically sent on a data server hosted by the tax administration. We receive regular data extractions from the server and use this to compile our dataset for the analysis.

Extractions. We mainly use two extractions, one recovered on January 31st, 2023 (for census operations carried out between 2019 and January 2023), and one recovered on May 5th, 2023 (for census operations between January and May 2023). We add to these extractions a subset of plots recovered from earlier extractions and that were deleted from the server due to a technical problem. There is a very small number (0.7% of plots) of plots covered twice. In these cases, we keep the observation from the most recent visit.

Sections. Initially, instead of 96, there were 97 treated sections. However for one treated section the administrative borders were modified before the census started, the section 'disappeared' and its plots were redistributed into two control sections. The section which was dissolved was in the Rufisque Ouest Commune, attached to the Rufisque tax office, the plots were redistributed into one section in the same Commune, and one in Rufisque Est, also linked to the same tax office. In our main analysis we control by randomization strata, and tax office is one of the stratification variables, thus we do not expect this to be an issue for the causal interpretation of the comparison between rule and discretion. Out of the 96 targeted sections for the census, the census was interrupted in two of them at its very start, because of pre-existing tensions between the local population and the tax administration on another topic, property titles. Only 11 out of 89 and 16 out of 404 plots were covered before the interruption, and they will not be exploited to prepare tax bills. We drop these two sections from the analysis. One was in the discretion arm, one in the rule arm, both in the Commune of Yoff (Ngor-Almadies tax office).

Creating the plot level dataset. The raw dataset is at the owner X tenant level. For our analysis, we build a plot level dataset. For 99.32% of the 38,417 plots covered in the census, there is only one owner per plot, making the plot level analysis relevant for this context. We use 'plot' and 'property' indifferently. The raw data needs to be treated differently for the discretion and rule arms. We are interested in total property value, and share occupied as main residence, since this determines the abatement for tax liabilities. In the discretion arm: we add up all rent values to generate total rent value for the plot, if there are several values for owner-occupied parts we also add these up, and thus generate the plot level total property value, which is the sum of rents and owner-occupied values. We categorize the plot as main residence if the bureaucrat ticks main residence, and as rented if the bureaucrat ticks rented. Note that this can occur even if the bureaucrat does not provide the associated values. Under the rule, we first compute the plot level predicted value by applying the coefficients from the property valuation formula, using the built area measurement we have from our GIS cadastral dataset and the section fixed-effect based on each plot's location. If the property is fully rented, or fully occupied by the owner, the predicted plot level value allows to generate the tax liability. If there are multiple owners on the property, and/or if the property is partly rented and partly occupied by the owner, we assign values proportionally to the number of rooms allocated to each owner or usage, as described in Section B.2. The number of rooms is recorded by bureaucrats. When we compare tax liabilities across arms, we provide a robustness check where the share of value subject to the abatement is computed in the exact same way in both arms, using number of rooms (see Figure A7 Panel (B)).

Corrections: values entered manually. We correct two types of entries which we assume are due to typos during the field work. First, we identify cases were property values are negative. We replace these by the absolute value. There are only 2 cases. Second, we identify cases where a monthly value is too small to be realistic, below 10,000 FCFA (16 USD). We replace these by 0. There are 610 cases out of 20,079 plots in the discretion arm (30 out of the 1,237 in the sample with market values).

Corrections: observable characteristics. There were several occurrences of technical challenges leading to the temporary absence of some or all of entries for observable characteristics on the tablets in the field. 192 properties out of the 18,148 of the rule arm have all observable characteristics missing. 8,744 have one characteristics or more missing, most often these four: Architecture, Sidewalk, Quality of Doors and Windows, Presence of a Cement Wall. This involves 569 out of the 1,166 plots of the sample with market values. All the characteristics are categorical variables entered as dummy variables in the property valuation algorithm. Therefore our replacement strategy is as follows: we replace each missing characteristic by its mean value in a given section. If there are no occurrences of the characteristic in the section, we replace the missing characteristic by its mean value overall.[95]

Comments. Bureaucrats can leave comments associated to each plot they visit. There is a comment in 54% of cases. Since these are open ended text entries they are not immediately usable. We process comments relying on key words to identify three relevant cases for our analysis: the bureaucrat reports there is conflict or tensions between her and the occupants (6.4% of observations); the bureaucrat reports having estimated the value herself without relying on any information from occupants (12.8% of observations); the bureaucrat reports that the property is managed by a real estate agency (6.7% of cases).

B.4 Verifications using Photos

We carry out some verifications based on pictures on a random sample of 100 properties that were visited both during the property tax census and by the real estate assessor. Among cases for which we have a picture from the census and from the assessors, the property matches for 92.5% of observations. The picture is missing or not interpretable for 20 cases in the census. Among cases where the picture matches, the same number of floors was entered in 68.9% of cases. When the floor numbers were different, the assessor is correct in 68% of cases. The most frequent situation is one in which the bureaucrat erroneously counted the ground floor as '1' instead of '0' (69.6% of cases where the floor number mismatches).

B.5 Description of Variables from the Bureaucrat Surveys

Baseline bureaucrat survey

Ever worked with the tax administration: Takes value one if the bureaucrat did any work with the tax administration before the program, either as a civil servant or a temporary employee.

From Dakar: Takes value one if the hometown of the bureaucrat is Dakar or one of the suburbs.

Any higher education: Takes value one if the bureaucrat completed high school and studied in a higher education institution (including vocational training).

Three years or more of higher education: Takes value one if the bureaucrat completed a three year degree or more in higher education.

Ethnic group: Self-reported with a possibility not to answer: Wolof or Lebou, Poular, Serere, Diola, Other.

Religion: Islam Tidjane, Islam Mouride, Islam Other, Christian.

Public service motivation: Standardized score computed from the sum of Likert-scale responses to the following questions: It is important for me to work in the public sector;

[95] We have these characteristics as recovered by the assessors, but we prefer not using this as our main replacement strategy since we are also measuring the gap between predictions based on bureaucrats versus assessor inputs.

I would not mind doing the same job in the private sector; It is not necessarily important for me that my work is useful for the community; I do not hesitate to devote all my energy to work.

In favor of government's role: Standardized score computed from the sum of Likert-scale responses to the following questions: According to me, the government can do a lot to make society more fair; the state should have the responsibility to satisfy everyone's basic needs (versus individuals taking care of their own needs).

In favor of widespread taxation: Standardized score computed from the sum of Likert-scale responses to the following questions: It is fair for a retired person to pay taxes if (s)he owns property; Only the richest people should pay taxes.

Endline bureaucrat survey

Rent: Amount of rent paid by the bureaucrat if (s)he is a tenant in current dwelling.

Income: Monthly income before the job with the tax administration for this program.

Emotions reading: Standardized score from a multiracial version of the Read the Mind in the Eyes Test (Weidmann & Deming, 2021; Dodell-Feder *et al.*, 2020).

Big five score: Standardized score from a 10 questions francophone version of the big five personality traits test (Plaisant, 2008). We also use separately the standardized score for each subcomponent: *Openness, Agreeableness, Extraversion, Conscientiousness, Neuroticism.*

Digit span: Standardized score for the combination of a forward and backward digit span test.

Math index: Standardized score obtained for six math questions: $10 + 5$; $27 - 4$; $32 - 13$; 7×6; 150000×4; 70000×12.

Persuasion evaluation: Standardized score given by members of the research team to bureaucrats for a verbal exercise, where the bureaucrat is told: Consider you are doing the census, and the owner of a property refuses to cooperate and provide information. What would you tell them? The graders were told to grade according to how persuasive the bureaucrats' discourse was. This exercise builds on (Chioda *et al.*, 2021).

Persuasion sum of items: Standardized sum of eight potential persuasion arguments that the bureaucrat included in his or her reply in the verbal exercise. The items are: giving more explanations on how the information will be used; suggesting to change languages or make polite salutations adapted to the owners' profile; suggesting to break the ice with some small talk or jokes; mentioning the public services the tax revenues will fund; reassuring the owner by saying that all owners of the neighborhood are getting the census; suggesting to have the supervisor of the tax office intervene; threatening the owner by mentioning possible prosecutions; offering to plan a meeting for later.

Endline supervisor survey

Overall performance score: Standardized score given by supervisors when asked: How would you grade the overall performance of bureaucrat [Name]? for each bureaucrat they supervise.

Performance items: Standardized sum of scores given by supervisors when grading each bureaucrat they supervise on the following items: social skills, fiscal knowledge, housing market knowledge, energy and stamina, negotiation skills, ease with technology, ease with reading maps.

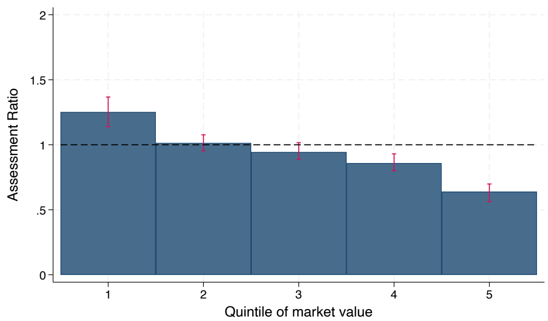
C Robustness: Missing Property Characteristics

In the sample of 1,166 plots of the rule arm with market value, there is at least one of the observable property characteristics missing from the census data in 595 observations. There are four characteristics (architecture, sidewalk, quality of doors and windows and wall) missing in 569 cases. This is due to a technical problem which occurred during some weeks of the census: these variables were simply absent from the form on bureaucrats' tablets. In the main analysis, we carry out some replacements as explained in Appendix B.3. In this Appendix, we use two strategies as robustness checks: first, we drop the plots for which at least one observable characteristic is missing from bureaucrats' inputs; second, we re-calibrate our property valuation formula excluding the four characteristics that are missing for a large number of cases.[96]

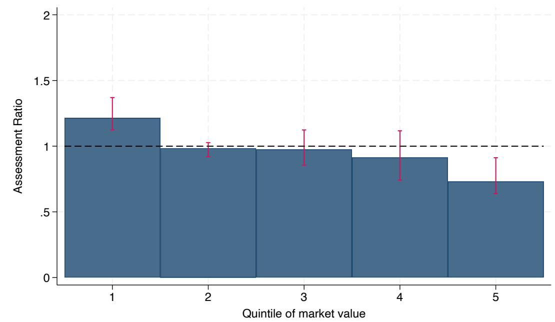
Our main results are unchanged: the rule outperforms discretion, and this is mainly driven by the upper part of the distribution, the pure rule outperforms the rule implemented by bureaucrats. Compared to when the rule is calculated using assessors' inputs, the rule implemented by bureaucrats is more regressive since it undervalues more strongly high-end properties.

[96]This rule performs as well as our main rule, the R^2 is 0.89, the MAPE is 32.6.

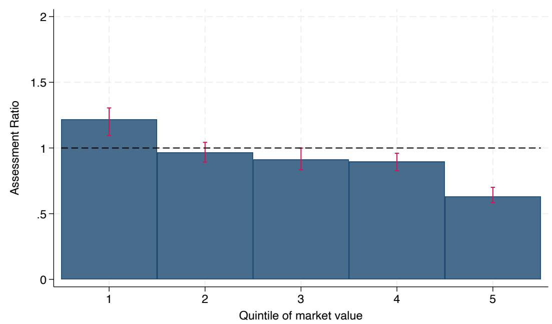
FIGURE A19
ASSESSMENT RATIO BY QUINTILE



(A) RULE-BASED - MAIN RESULTS



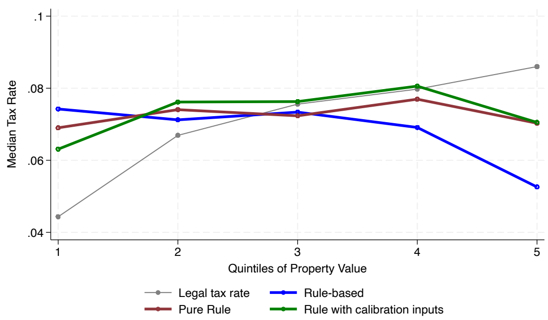
(B) ROBUSTNESS: DROPPING MISSING



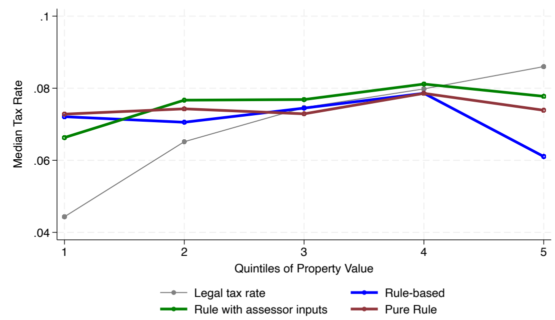
(C) ROBUSTNESS: RECALIBRATING RULE

Notes: This Figure plots the median assessment ratio by quintile, in Panel (A) we reproduce our main results for the rule arm from Figure 5. In Panel (B) we show results after dropping plots with at least one observable characteristic missing. In Panel (C) we show results using the rule re-calibrated without the four often missing characteristics. The assessment ratio is computed as assessed value for a given property divided by the market value of the same property. The red line shows the confidence interval for the median value of each quintile. Quintiles are based on market values.

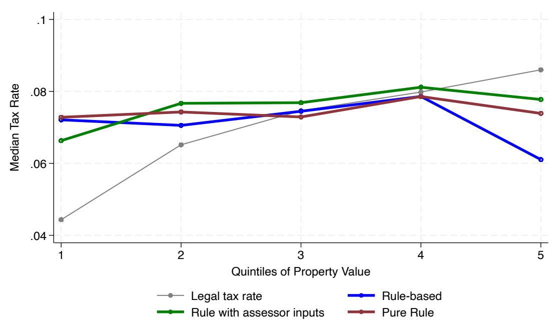
FIGURE A20
TAX RATES UNDER THE RULE



(A) RULE-BASED - MAIN RESULTS



(B) ROBUSTNESS: DROPPING MISSING



(C) ROBUSTNESS: RECALIBRATING RULE

Notes: This Figure shows the median tax rate by quintile in the rule arm. Panel (A) reproduces our main results from Figure 7. In Panel (B) we show results after dropping plots with at least one observable characteristic missing. In Panel (C) we show results using the rule re-calibrated without the four often missing characteristics. The gray line is the benchmark tax profile computed directly on market values. The blue line is the tax profile with the rule, the red line is the tax profile using the pure rule with remote covariates only. Additionally, the green line shows the tax profile generated with the rule if using the characteristics from the calibration (assessors' dataset). The tax rate is 8.6 percent with an abatement if the property is owner-occupied which explains the progressive profile.

TABLE A16
REMOVING DISCRETION INCREASES ACCURACY: DROPPING MISSING

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) |Gap| *mil.FCFA*	(4) Ass. Ratio
Panel A: Discretion				
Mean[1] (sd)	−7.14 (17.69)	−2.41	8.92 (16.87)	0.71
Panel B: Rule-based				
Overall				
Mean[1] (sd)	−2.58 (11.58)	−0.18	4.29 (11.06)	1.00
$\hat{\beta}_{Discretion}$	−2.85*	−0.91**	2.53	−0.18**
	(1.55)	(0.38)	(1.58)	(0.07)
Low Value				
Mean[1] (sd)	0.19 (2.08)	0.01	1.21 (1.70)	1.18
$\hat{\beta}_{Discretion}$	−0.06	−0.16	0.14	−0.04
	(0.22)	(0.17)	(0.19)	(0.12)
High Value				
Mean[1] (sd)	−6.96 (17.58)	−2.68	9.16 (16.54)	0.71
$\hat{\beta}_{Discretion}$	−3.79	−3.38***	3.18	−0.19**
	(2.69)	(1.26)	(2.73)	(0.08)
Panel C: Pure Rule				
Overall				
Mean[1] (sd)	−0.29 (6.38)	0.11	2.03 (6.06)	1.12
$\hat{\beta}_{Discretion}$	−4.62***	−1.79***	4.24***	−0.32***
	(1.10)	(0.51)	(0.99)	(0.06)
Low Value				
Mean[1] (sd)	0.23 (0.86)	0.14	0.61 (0.66)	1.19
$\hat{\beta}_{Discretion}$	0.00	−0.43***	0.91***	−0.01
	(0.22)	(0.14)	(0.16)	(0.11)
High Value				
Mean[1] (sd)	−1.12 (10.15)	−0.15	4.27 (9.28)	1.01
$\hat{\beta}_{Discretion}$	−8.1414.1	−5.77***	6.62***	−0.49***
	(1.75)	(0.76)	(1.55)	(0.05)
N plots: 1695				
N Sections: 85				
Mean (sd) market value: 64.10 (14.10)				
Median market value: 5.40				

Notes: This Table shows the effect of discretion on the tax base gap, similar to our Table 3, except that we drop plots with at least one observable characteristic missing. We run regression 2: $Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk}$ with four different outcomes. In column (1) the outcome variable is the tax base gap defined as bureaucrat value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as the bureaucrat's value over the market value. In Panel (A), we display summary statistics for the discretion arm. In Panel (B), values for the rule arm are the rule-based valuations based on bureaucrats' inputs. In Panel (C), values for the rule arm are the pure rule predictions based on remote covariates. Each panel is divided in three subpanels, the first one uses the full sample of properties, the second is restricted to low value properties (quintile 1 and 2 of market values), the third one is restricted to high value properties (quintiles 3 to 5 of market values). In each subpanel, the first rows display descriptive statistics of the outcome variable in the rule arm; the second rows show the coefficient of interest on the dummy for Discretion. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed effects and errors are clustered at the section level. [1]In column (2) the displayed value is the $median$ of the tax base gap.

TABLE A17
REMOVING DISCRETION INCREASES ACCURACY: RECALIBRATING RULE

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) |Gap| *mil.FCFA*	(4) Ass. Ratio
Panel A: Discretion				
Mean[1] (sd)	−7.12 (17.72)	−2.41	8.94 (16.88)	0.71
Panel B: Rule-based				
Overall				
Mean[1] (sd)	−2.17 (12.93)	−0.21	4.80 (12.20)	1.05
$\hat{\beta}_{Discretion}$	−4.77***	−1.93***	3.76***	−0.35***
	(1.27)	(0.42)	(1.38)	(0.05)
Low Value				
Mean[1] (sd)	0.38 (2.09)	0.11	1.27 (1.70)	1.23
$\hat{\beta}_{Discretion}$	−0.45**	−0.52***	0.32**	−0.21**
	(0.20)	(0.13)	(0.16)	(0.10)
High Value				
Mean[1] (sd)	−4.47 (17.39)	−1.31	7.97 (16.09)	0.89
$\hat{\beta}_{Discretion}$	−6.86***	−4.50***	5.32***	−0.37***
	(1.77)	(0.67)	(1.89)	(0.05)
Panel C: Pure Rule				
Overall				
Mean[1] (sd)	−0.36 (7.64)	0.12	2.83 (7.11)	1.13
$\hat{\beta}_{Discretion}$	−5.37***	−2.42***	4.71***	−0.38***
	(0.90)	(0.44)	(0.93)	(0.04)
Low Value				
Mean[1] (sd)	0.39 (1.13)	0.25	0.73 (0.94)	1.24
$\hat{\beta}_{Discretion}$	−0.28	−0.57***	0.88***	−0.13
	(0.20)	(0.14)	(0.14)	(0.09)
High Value				
Mean[1] (sd)	−1.04 (10.43)	−0.33	4.72 (9.36)	1.03
$\hat{\beta}_{Discretion}$	−8.2515.8	−5.44***	6.87***	−0.47***
	(1.27)	(0.55)	(1.28)	(0.05)

N plots: 2290
N Sections: 94
Mean (sd) market value: 77.00 (15.80)
Median market value: 5.60

Notes: This Table shows the effect of discretion on the tax base gap, similar to our Table 3, except we use the rule re-calibrated without the four often missing property characteristics. We run regression 2: $Y_{ijk} = \alpha + \beta D_{jk} + S_k + \epsilon_{ijk}$ with four different outcomes. In column (1) the outcome variable is the tax base gap defined as bureaucrat value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as the bureaucrat's value over the market value. In Panel (A), we display summary statistics for the discretion arm. In Panel (B), values for the rule arm are the rule-based valuations based on bureaucrats' inputs. In Panel (C), values for the rule arm are the pure rule predictions based on remote covariates. Each panel is divided in three subpanels, the first one uses the full sample of properties, the second is restricted to low value properties (quintile 1 and 2 of market values), the third one is restricted to high value properties (quintiles 3 to 5 of market values). In each subpanel, the first rows display descriptive statistics of the outcome variable in the rule arm; the second rows show the coefficient of interest on the dummy for Discretion. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed effects and errors are clustered at the section level. [1]In column (2) the displayed value is the *median* of the tax base gap.

TABLE A18
RULE-BASED VS PURE RULE: DROPPING MISSING

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) |Gap| *mil.FCFA*	(4) Ass. Ratio
Panel A: Pure Rule				
Overall				
Mean[1] (sd)	−0.29 (6.38)	0.11	2.03 (6.06)	1.12
$\hat{\beta}_{RuleBur}$	−2.29***	−0.33	2.27***	−0.13*
	(0.77)	(0.00)	(0.59)	(0.07)
Low Value				
Mean[1] (sd)	0.23 (0.86)	0.14	0.61 (0.66)	1.19
$\hat{\beta}_{RuleBur}$	−0.04	−0.19	0.61***	−0.02
	(0.15)	(0.00)	(0.13)	(0.07)
High Value				
Mean[1] (sd)	−1.12 (10.15)	−0.15	4.27 (9.28)	1.01
$\hat{\beta}_{RuleBur}$	−5.84***	−2.17	4.89***	−0.30***
	(1.52)	(0.00)	(1.24)	(0.09)
Panel B: Rule with Assessor Inputs				
Overall				
Mean[1] (sd)	−0.27 (6.71)	0.11	1.97 (6.42)	1.09
$\hat{\beta}_{RuleBur}$	−2.31***	−0.34*	2.33***	−0.10
	(0.72)	(0.17)	(0.58)	(0.07)
Low Value				
Mean[1] (sd)	0.22 (0.79)	0.11	0.54 (0.61)	1.12
$\hat{\beta}_{RuleBur}$	−0.03	−0.13	0.67***	0.05
	(0.16)	(0.00)	(0.14)	(0.07)
High Value				
Mean[1] (sd)	−1.04 (10.70)	−0.04	4.22 (9.88)	1.04
$\hat{\beta}_{RuleBur}$	−5.91***	−2.20**	4.93***	−0.33***
	(1.40)	(0.96)	(1.19)	(0.09)

N obs: 1141
N plots: 571
N Sections: 38
Mean (sd) market value: 59.20 (9.80)
Median market value: 3.50

Notes: This Table shows the effect on the tax base gap of the limited degree of discretion which remains under the rule based process implemented by bureaucrats, compared to benchmark rules without any bureaucrat discretion. It is similar to our Table 7, except that we drop plots with at least one observable characteristic missing. We run regression 7: $Y_{irjk} = \alpha + \beta RuleBur_{irjk} + S_k + \epsilon_{irjk}$ where Y_{itjk} is the outcome for plot i of section j and strata k under rule r, $RuleBur_{irjk}$ is a dummy taking value one if r is the rule as implemented by bureaucrats, and zero if r is the benchmark rule with no bureaucrat discretion. In Panel (A), the benchmark rule is the pure rule with remote covariates, in Panel (B) the benchmark rule is the rule calculated using the observable characteristics recovered in the assessors' dataset. In column (1) the outcome variable is the tax base gap defined as rule value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as the rule value over the market value. Each panel is divided in three subpanels, the first one uses the full sample of properties, the second is restricted to low value properties (quintile 1 and 2 of market values), the third one is restricted to high value properties (quintiles 3 to 5 of market values). The first row of each subpanel display descriptive statistics of the outcome variable under the benchmark rule; the second row show the coefficient of interest on the dummy for $RuleBur$. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed effects and errors are clustered at the section level. Each property appears twice in the regression sample. [1]In column (2) the displayed value is the *median* of the tax base gap.

TABLE A19
RULE-BASED VS PURE RULE: RECALIBRATING RULE

	(1) Gap *mil.FCFA*	(2) Gap (median) *mil.FCFA*	(3) ∣Gap∣ *mil.FCFA*	(4) Ass. Ratio
Panel A: Pure Rule				
Overall				
Mean[1] (sd)	−0.36 (7.64)	0.12	2.83 (7.11)	1.13
$\hat{\beta}_{RuleBur}$	-1.81^{***}	-0.31^{**}	1.97^{***}	-0.08^{**}
	(0.57)	(0.12)	(0.46)	(0.04)
Low Value				
Mean[1] (sd)	0.39 (1.13)	0.25	0.73 (0.94)	1.24
$\hat{\beta}_{RuleBur}$	0.00	−0.09	0.54^{***}	0.00
	(0.12)	(0.08)	(0.10)	(0.05)
High Value				
Mean[1] (sd)	−1.04 (10.43)	−0.33	4.72 (9.36)	1.03
$\hat{\beta}_{RuleBur}$	-3.43^{***}	-1.19^{***}	3.25^{***}	-0.15^{***}
	(0.94)	(0.35)	(0.78)	(0.04)
Panel B: Rule with Assessor Inputs				
Overall				
Mean[1] (sd)	−0.62 (8.24)	0.06	2.76 (7.79)	1.08
$\hat{\beta}_{RuleBur}$	-1.55^{***}	-0.30^{**}	2.04^{***}	−0.03
	(0.45)	(0.13)	(0.40)	(0.03)
Low Value				
Mean[1] (sd)	0.26 (0.90)	0.14	0.61 (0.71)	1.13
$\hat{\beta}_{RuleBur}$	0.12	−0.02	0.66^{***}	0.10^{**}
	(0.12)	(0.08)	(0.10)	(0.05)
High Value				
Mean[1] (sd)	−1.42 (11.26)	−0.18	4.69 (10.34)	1.03
$\hat{\beta}_{RuleBur}$	-3.05^{***}	-1.24^{***}	3.28^{***}	-0.14^{***}
	(0.74)	(0.32)	(0.65)	(0.04)

N obs: 2331
N plots: 1166
N Sections: 47
Mean (sd) market value: 86.00 (15.40)
Median market value: 4.80

Notes: This Table shows the effect on the tax base gap of the limited degree of discretion which remains under the rule based process implemented by bureaucrats, compared to benchmark rules without any bureaucrat discretion. It is similar to our Table 7, except that the rule used here is the one re-calibrated without the four often missing characteristics. We run regression 7: $Y_{irjk} = \alpha + \beta RuleBur_{irjk} + S_k + \epsilon_{irjk}$ where Y_{itjk} is the outcome for plot i of section j and strata k under rule r, $RuleBur_{irjk}$ is a dummy taking value one if r is the rule as implemented by bureaucrats, and zero if r is the benchmark rule with no bureaucrat discretion. In Panel (A), the benchmark rule is the pure rule with remote covariates, in Panel (B) the benchmark rule is the rule calculated using the observable characteristics recovered in the assessors' dataset. In column (1) the outcome variable is the tax base gap defined as rule value minus market value, column (2) uses the same outcome but with a quintile regression at the median. In column (3) the outcome is the absolute value of the tax base gap. All amounts are in millions of FCFA and winsorized at the 1% level. In column (4) the outcome is the assessment ratio, computed as the rule value over the market value. Each panel is divided in three subpanels, the first one uses the full sample of properties, the second is restricted to low value properties (quintile 1 and 2 of market values), the third one is restricted to high value properties (quintiles 3 to 5 of market values). The first row of each subpanel display descriptive statistics of the outcome variable under the benchmark rule; the second row show the coefficient of interest on the dummy for $RuleBur$. *,** and *** indicate statistical significance at the 10,5 and 1% level respectively. We control for strata fixed effects and errors are clustered at the section level. Each property appears twice in the regression sample. [1]In column (2) the displayed value is the $median$ of the tax base gap.

D Screening Bureaucrats: Additional Results

D.1 Bureaucrat skills and measures of performance

In Figure A10, we show correlations between bureaucrat characteristics and other measures of performance: the share of plots for which the bureaucrat recovered owners' identification details, the number of plots per day, and the ability to value high end properties in the full sample of the discretion arm, where we rely on predicted values as a benchmark. Social skills such as openness and extraversion correlate with the ability to recover owner details. Social skills are also deemed crucial by bureaucrats themselves, as shown in Figure A16. Persuasion correlates with the number of plots covered per day, and with a lower absolute tax base gap for above median properties in the full sample. Supervisor evaluations correlate with some measures of performance as shown in Figure A11, but not with the bureaucrat fixed-effects.

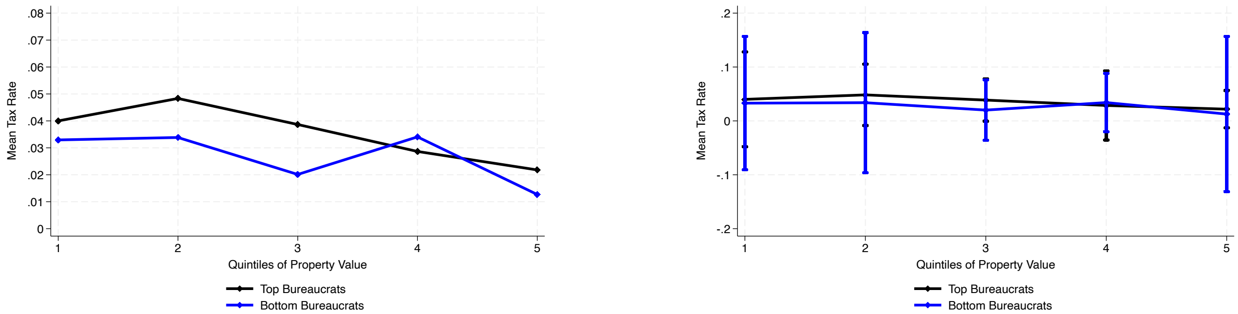
D.2 k-means clustering

We use 18 continuous bureaucrat characteristics, and the k-means clustering procedure constitutes two groups with the maximum Euclidean distance across all these characteristics (see Table A20 for the mean values of each characteristic across clusters). One caveat is that due to missing observations for some bureaucrat covariates we are able to assign a cluster only to 67 bureaucrats. Yet, we find that the two groups distinguish bureaucrats of different performance: bureaucrats from cluster 2 are 25 percentage points more likely to be a top bureaucrat (p-value of 0.07), and in a regression on the discretion arm only, we find that the absolute tax base gap is 2.64 millions FCFA smaller for a bureaucrat from cluster 2 compared to a bureaucrat from cluster 1 (p-value of 0.11).

D.3 Tax profiles

When splitting bureaucrats into different groups, it is important to consider two dimensions to assess whether one group is better at generating the tax roll: accuracy of the tax profile (tax rates and how they vary through the distribution); and dispersion. First, in Figure A21 we show the tax profiles for top and bottom bureaucrats as estimated in our fixed-effect analysis of Section 5.2, where top bureaucrats are those for which $\alpha_{b,EB} < 0$. The tax profile generated by top bureaucrats is clearly preferable, which is mechanical since bureaucrat types are defined based on their relative tax base gap. Next, in Panels (A) and (B) of Figure A22, we assess whether similar results could be obtain if screening on three years or more of higher education, which is the main demographic and easily screenable variable that bureaucrat type correlates with. The difference in tax rates is not as clear-cut as when using bureaucrat type directly, but still, the tax profile for higher educated bureaucrats is preferable, and they also generate much less dispersion than bureaucrats without this level of education. Finally, in Panels (C) and (D) of Figure A22, we sort bureaucrats based on the k-means clustering exercise. There is a slight difference in the tax profile across the two clusters, but more striking is again the difference in terms of dispersion. Bureaucrats from cluster 2 generate more horizontal equity throughout most of the distribution.

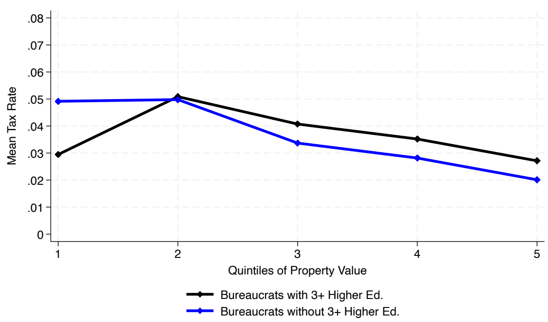
FIGURE A21
TAX RATES BY BUREAUCRAT TYPE



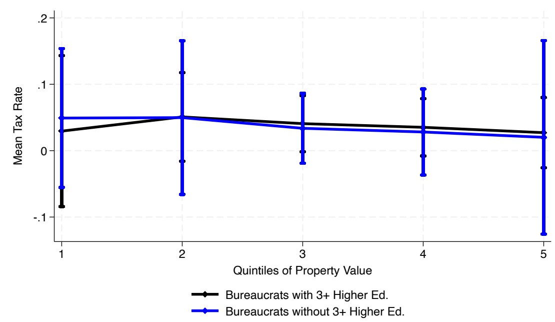
(A) TAX RATES **(B) DISPERSION**

Notes: This Figure shows the median tax rate by quintile in the discretion arm, splitting the sample into plots covered by top versus bottom bureaucrats. Bureaucrat type is defined based on the fixed-effects estimated in Section 5.2: a top bureaucrat is one for which $\alpha_{b,EB} < 0$.

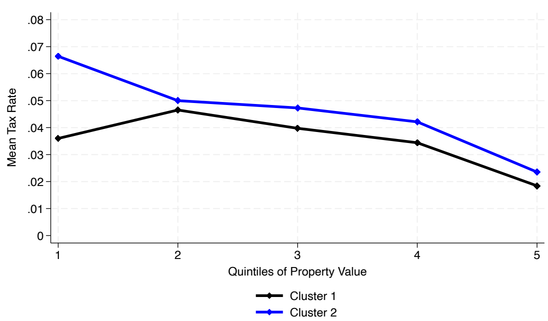
FIGURE A22
TAX RATES WHEN SCREENING BUREAUCRATS



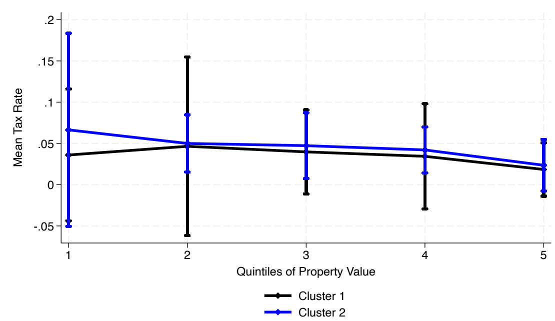
(A) BY EDUCATION: TAX RATES



(B) BY EDUCATION: DISPERSION



(C) BY CLUSTER: TAX RATES



(D) BY CLUSTER: DISPERSION

Notes: This Figure shows the median tax rate by quintile in the discretion arm when screening bureaucrats. In Panels (A) and (B): bureaucrats are sorted into with (81) or without (112) 3 years or more of higher education. There are $N = 1,087$ plots in the Figures. In Panels (C) and (D), bureaucrats are sorted into two clusters obtained by k-means clustering. Only 67 bureaucrats are assigned a cluster, 47 in cluster 1 and 20 in cluster 2 (there are $N = 408$ plots in the Figures).

TABLE A20
K-MEANS CLUSTERING RESULTS

	Cluster 1	Cluster 2
Age	28.33	38.00
Education Level	11.02	11.26
Index for PSM	0.05	−0.14
Index for Govt	0.02	0.00
Index for Tax morale	0.10	0.37
Emotions	0.00	0.00
Big Five	−0.16	0.12
Digit span	−0.10	0.05
Math	−0.05	−0.12
Persuasion score	−0.13	−0.03
Persuasion sum	−0.24	0.33
Supervisor Eval.	0.07	0.26
Gap High End Prop.	−0.71	−0.65
Gap Low End Prop.	0.56	0.77
Tax Magnitude score	−0.25	−0.07
Tax Questions score	0.06	0.03
Fair to Tax Retired	0.60	0.58
Only Rich should Pay	0.15	0.11
N	48	19

Notes: In this Table we report the mean value of each characteristic used in the k-means clustering of bureaucrats, column (1) reports the mean for the first cluster and column (2) reports the mean for the second cluster. The k-means clustering procedure constitutes two groups with the maximum Euclidean distance across all characteristic. We only include continuous variables. Education Level is a scale between 1 (secondary education completed) and 8 (university beyond bachelors); PSM is a public motivation score, Govt measures the perception of the role of government, and Tax morale perceptions that taxation should be widespread, all three are measured at baseline. Emotions is the Read the Mind in the Eyes test score, Persuasion score and Persuasion sum are scores from a verbal exercise graded by the research team, Supervisor Eval. is the score provided by the supervisor for this given bureaucrat. The Gap and Tax score variables are from the endline survey questions where we measure bureaucrats knowledge of property values and tax rules. Fair to Tax Retired and Only Rich should Pay are perceptions measured at endline. All scores are standardized.