Cuadernos de historia económica

Smithian State Formation

By: Diana Ricciulli- Marin
Mateo Uribe-Castro

No. 68
August, 2026

Centro de Estudios Económicos Regionales (CEER) - Cartagena

Smithian State Formation*

Diana Ricciulli-Marin[†1] and Mateo Uribe-Castro[‡2]

[1]Banco de la República
[2]Universidad de los Andes

The series **Cuadernos de Historia Económica** is a publication of Banco de la República in Cartagena. The opinions contained in this document are the authors' sole responsibility and do not commit Banco de la República or its Board of Directors.

Abstract

We examine the relationship between land inequality and state capacity, measured with tax collection. Contrary to traditional models that emphasize the redistributive role of taxation, we focus on its role in providing public goods that help develop markets, increasing productivity (e.g. enforcement of property rights, coordination, roads, electricity). We build a simple model of public good provision where landowners decide whether or not to comply with property taxes taking into account: (i) that government expenditure increases property values, and (ii) their expectation of the punishment when evading taxes. We validate the model empirically, using data from Colombian municipalities between 1923 and 1960, in two ways. First, we use detailed land values' data from cadastres available for a subsample of municipalities and the model's structure to predict tax revenues and compare them with actual revenues. Second, we provide empirical evidence supporting the model's main prediction: land concentration is positively correlated with tax revenues per capita, a relationship that is robust to controlling for potential sources of omitted variable bias. Two additional empirical findings further support the mechanisms underlying the model: the positive relationship is stronger in places with greater potential gains from market development, and land concentration reduces the average fiscal cost of collecting one peso of revenue.

JEL Classifications: H26, H41, H71

Keywords: wealth inequality, tax compliance, local public goods

*We thank Kevin Milligan, Kory Kroft, Antonella Bandiera, Leopoldo Fergusson, Claudio Ferraz, Patrick Francois, Viktoriya Hnatkovska, Terry Moon, Nathan Nunn, John Wallis, Juliana Jaramillo, Jaime Bonet, Jhorland Ayala, and Jonathan Weigel, and participants at EGAP 29 - Open Stream, RIDGE Political Economy Workshop, Uniandes CEDE Seminar, and Stanford HSS Workshop for helpful comments and suggestions. The opinions expressed in this paper belong to the authors' and do not necessarily reflect the views of the Colombian Central Bank or its Board of Directors. Financial support from the Stone Center of Wealth and Income Inequality at UBC is gratefully acknowledged.

[†]email: driccima@banrep.gov.co

[‡]email: m.uribecastro@uniandes.edu.co

Desigualdad de la tierra y formación del Estado*

Diana Ricciulli-Marin† Mateo Uribe-Castro‡

Banco de la República Universidad de los Andes

La serie **Cuadernos de Historia Económica** es una publicación del Banco de la República en Cartagena. Las opiniones contenidas en este documento son de exclusiva responsabilidad de los autores y no comprometen al Banco de la República ni a su Junta Directiva.

Resumen

Estudiamos la relación entre la desigualdad de la tierra y la capacidad estatal, medida como recaudo tributario. A diferencia de los modelos tradicionales que enfatizan el papel redistributivo de los impuestos, este artículo estudia su rol en la provisión de bienes públicos que favorecen el desarrollo de los mercados e incrementan la productividad (por ejemplo, la protección de los derechos de propiedad, la coordinación, las carreteras y la electrificación). Construimos un modelo simple de provisión de bienes públicos en el que los terratenientes deciden si cumplir o no con el impuesto predial, teniendo en cuenta: i) que el gasto público incrementa el valor de sus propiedades, y ii) sus expectativas sobre las sanciones asociadas a la evasión fiscal. Validamos el modelo empíricamente, usando datos de municipios colombianos entre 1923 y 1960. Primero, empleamos datos de catastros disponibles para una submuestra de municipios y la estructura del modelo para predecir los ingresos tributarios y compararlos con los ingresos observados. Segundo, presentamos evidencia que respalda la predicción central del modelo: la concentración de la tierra está positivamente correlacionada con los ingresos fiscales per cápita, relación que es robusta al controlar por posibles fuentes de sesgo por variable omitida. Dos hallazgos adicionales refuerzan los mecanismos subyacentes del modelo: la relación positiva es más pronunciada en municipios con mayor potencial de desarrollo de mercados, y la concentración de la tierra reduce el costo fiscal promedio de recaudar un peso de ingresos.

Clasificación JEL: H26, H41, H71

Palabras clave: desigualdad de la tierra, cumplimiento tributario, bienes públicos locales

*Agradecemos a Antonella Bandiera, Leopoldo Fergusson, Claudio Ferraz, Patrick Francois, Viktoriya Hnatkovska, Juliana Jaramillo, Kory Kroft, Kevin Milligan, Terry Moon, Nathan Nunn, John Wallis, Jaime Bonet, Jhorland Ayala y Jonathan Weigel, y a los participantes del EGAP 29 - Open Stream, RIDGE Political Economy Workshop, Uniandes CEDE Seminar, y Stanford HSS Workshop por sus comentarios y sugerencias. Las opiniones expresadas en este artículo son de los autores y no reflejan necesariamente las del Banco de la República de Colombia ni las de su Junta Directiva. Se agradece el apoyo financiero del Stone Center of Wealth and Income Inequality en la Universidad de British Columbia.

†correo: driccima@banrep.gov.co

‡correo: m.uribecastro@uniandes.edu.co

1 Introduction

Prominent explanations for why elites pay taxes assume that they do so only under political pressure, as the state uses taxes for redistribution. For instance, Acemoglu and Robinson (2006) argue that landowning elites agree to redistribution as a strategic response to the threat of revolution, extending the franchise to commit credibly to future transfers. Their account builds on the logic of Meltzer and Richard (1981), who show that rising inequality increases support for redistribution, under the assumption that all individuals comply with the tax schedule, a logic that was also alluded to by de Tocqueville (1835). These perspectives understand taxation as the outcome of redistributive conflict, where asset-owners comply only to avoid more costly outcomes. However, recent work by Besley (2020) emphasizes that state capacity depends not only on coercion, but also on reciprocity: when citizens believe they benefit from government action, they are more willing to comply.

This alternative perspective emphasizes coordination rather than redistribution. Historically, many states expanded by providing public goods that reduced uncertainty, and transaction and transportation costs, allowing easier exchange. Land and capital owners benefit from the state's provision of public goods (e.g. property rights protection, infrastructure construction, and contract enforcement) since asset values increase when markets work better. However, these public goods are not amenable to private provision, precisely because the state's role as coordinator is necessary.[4] Robert Bates (1987) referred to this idea as "Smithian state formation:" elites support taxation not because they are coerced, but because they reap the benefits from state action.[5] When public expenditure raises asset values, taxation can be seen as an investment rather than a burden.

In this paper, we formalize this idea in a simple model of public good provision with heterogeneous landowners that we test using historical data from Colombian municipalities. The model builds on the tax compliance framework of Allingham and Sandmo (1972), adapting it to capture how the distribution of land ownership shapes the willingness of elites to pay taxes. Landowners observe a fiscal policy composed of a tax rate, a penalty rate for non-compliance, and a fiscal multiplier that maps public expenditure into private land values. Each landowner chooses whether to comply, balancing the expected benefits of taxation with the expected cost of evasion. The model predicts that when land inequality is high, landowners have both more to gain from public goods and fewer peers to hide behind. As a result, a larger share of them comply, and tax revenues increase.

The model rests on two key assumptions. First, we assume that public expenditure increases land values proportionally. This link is indirect: property rights protection in-

[4] Smith lays out this argument on Book V of The Wealth of Nations, referring specifically to "Public works and institutions for facilitating the commerce of society."

[5] "One of the basic arguments linking political centralization with economic reward rests upon the desire of people to benefit from the gains in welfare which can be reaped from markets.(...) The origins of the state, then, lie in the welfare gains that can be reaped through the promotion of markets." Essays on the Political Economy of Rural Africa.

creases expected returns on investment, transportation infrastructure enhances access to markets, and public services such as electricity and sanitation improve amenities and productivity, all of which raise land values over time (Besley, 1998; Donaldson and Hornbeck, 2016; Coury et al., 2022). Second, we assume that enforcement is more effective when land is concentrated. Specifically, we model enforcement probability as a function of the number of landowners rather than the full land distribution, reflecting a situation in which the state does not observe the complete distribution of landholdings. This assumption simplifies the enforcement mechanism and reflects administrative capacity constraints.

These assumptions offer analytical leverage but also raise important questions. The model treats landowners as policy takers: they do not set tax rates, influence enforcement, or shape expenditure decisions. Their only choice is whether to comply. This abstraction rules out elite capture through lobbying, tax bargaining, or informal arrangements, which are common features in weakly institutionalized states. While this is a simplification, it isolates the logic of voluntary compliance and allows us to study coordination incentives under minimal political structure.

Moreover, the model adopts linear compliance costs, consistent with the deterrence literature, where paying taxes is modeled as a monetary transfer with no increasing marginal cost (Becker, 1968; Allingham and Sandmo, 1972). This distinguishes our approach from collective action models where contributions to public goods involve increasing participation or mobilization costs (Olson, 1965; Banerjee et al., 2007). Our focus is not on effort allocation or participation thresholds, but on how inequality shapes the incentives to comply with taxation when public goods increase private returns.

We test the model using newly digitized municipal-level fiscal data from Colombia between 1923 and 1960, a period in which tax rates were set nationally but municipalities retained substantial autonomy over tax collection and public spending. We calibrate the model using property tax and penalty rates consistent with Colombia's 20th century tax policy, and use the number of landowners as a share of agricultural workers from the 1938 Population Census to calibrate the enforcement probability. For municipalities in Cundinamarca, we additionally observe the full distribution of land ownership in 1890 from detailed cadastral records (Acemoglu et al., 2007), allowing us to predict tax revenues by determining which landowners would comply based on their land value shares and the model's structure. The model predicts that municipalities with higher land concentration should collect higher tax revenues, a prediction that holds along both dimensions of inequality: a higher land Gini coefficient (intensive margin) and a lower share of landowners in the agricultural population (extensive margin) are both associated with greater tax revenues. Moreover, the model's predicted revenues are strongly correlated with observed revenues, a notable result given the model's parsimonious structure and the measurement error inherent in this exercise.

To evaluate whether our results generalize beyond the Cundinamarca sample, we extend the analysis to a broader sample of 702 municipalities. In this case, we focus on land concentration at the extensive margin, namely the share of landowners in the agri-

cultural population in 1938, as it is the only inequality measure available for the full sample. This measure is strongly correlated with Cundinamarca's 1890 land Gini coefficients, suggesting both long-run persistence in land concentration and a strong relationship between both inequality margins. Consistent with the model's predictions, municipalities with higher land concentration (a lower share of landowners) exhibit higher tax revenues per capita. This relationship is robust to alternative measures, such as tax revenues per hectare, and remains stable throughout the sample period despite institutional and economic changes that may have affected municipalities differently according to their initial level of inequality.

While we are cautious about interpreting the correlations as causal, our model does argue for a direct role of inequality in facilitating higher tax revenues. The main empirical challenge is that land inequality may be correlated with other factors that also affect fiscal capacity or taxpayers' willingness to pay. Therefore, we evaluate whether key sources of endogeneity could be biasing the results. We consider four potential threats.

First, municipalities with greater inequality may benefit from economies of scale, and their higher level of economic activity may lead to more revenues. To account for this, we include controls for the share of employment in manufacturing and services in 1912 and a dummy for departmental capitals. These variables proxy for persistent differences in economic development. The coefficient on landowner share remains negative and statistically significant, although smaller in magnitude.

To further understand whether the observed relationship between inequality and tax revenue reflects underlying economic activity or a more direct mechanism, we analyze the composition of municipal tax revenue. We distinguish between direct taxes (mainly property taxes) and indirect taxes (sales taxes). If the correlation were driven primarily by income or consumption levels in richer municipalities, one would expect a stronger association between inequality and indirect taxes. Yet, we find the opposite. The negative correlation between landowner share and direct tax revenue is nearly twice as large as that for indirect tax revenue. This distinction is meaningful because direct taxes are more closely tied to property ownership and compliance, while indirect taxes are a function of broader economic activity. The stronger result for direct taxation supports the model's central mechanism: when public goods raise land values, landowners in more unequal municipalities have stronger incentives to comply with property taxes.

Second, differences in agrarian specialization may jointly influence land inequality and tax capacity. Certain crops or agricultural systems, like cattle ranching, are associated with large-scale landholdings and more unequal distributions. At the same time, these land-intensive systems might shape fiscal legibility, i.e., how property values are assessed or how easily revenue is collected (Garfias and Sellars, 2026). To address this possibility, we control for two relevant indicators: the number of coffee trees per capita and the number of cattle slaughtered per capita in 1948. These proxies capture the intensity and type of agricultural activity, particularly export orientation and land use scale. While cattle production is associated with higher tax revenues, adding these controls does not alter

the core finding: municipalities with higher land inequality still tend to raise more tax revenue per capita.

Third, historical legacies of state-building institutions may jointly shape land inequality and fiscal capacity. To address this concern, we include controls for the presence of *encomiendas* in 1550 and the presence of indigenous settlements in 1535. *Encomiendas*, early fiscal institutions used by the Spanish Crown to distribute the rents from Conquest, created concentrated claims on local resources that influenced both early inequality and state formation (Faguet et al., 2024). These institutions laid the groundwork for local fiscal capacity that persisted into the 20th century. Including these historical controls does not attenuate the negative association between landowner share and per capita tax revenue, suggesting that the relationship is not driven by persistent colonial differences in fiscal infrastructure.

Finally, the positive correlation between land concentration and tax revenues may reflect landowners' strategic response to the threat of peasant rebellion rather than voluntary compliance motivated by the benefits of public goods. To assess this alternative mechanism, we control for peasant–landowner legal disputes between 1900 and 1931 using data from LeGrand (1986), who compiled historical records of land conflicts during this period. Controlling for historical land conflict does not meaningfully change the estimated effect of land concentration, suggesting that the relationship is not driven by landowners' responses to the threat of peasant unrest.

To bound the overall threat from unobserved confounding, we report the Oster (2019) δ statistic. This measure asks how important unobservables would need to be, relative to the observed controls, to fully explain away the estimated effect. A value of $\delta > 1$ indicates that unobservables would need to account for more variation than the entire set of included controls for the coefficient to be driven to zero, which is considered a demanding threshold. The δ statistic for our preferred specification is 1.077, implying that unobservables would need to be at least as important as the full set of observed controls to drive the estimated effect to zero, a degree of confounding that is unlikely given the richness of our control set, lending further credibility to the estimated relationship.

We further examine whether the data support the two key mechanisms in the model that increase landowners' willingness to comply in more unequal societies: higher returns from public goods and more effective enforcement. First, we test whether the relationship between land inequality and tax revenues is stronger in places where public investment is more likely to increase land values. In our framework, tax compliance becomes more attractive when public goods enhance the value of landholdings, a mechanism that should be particularly relevant in municipalities more integrated into markets. In Colombia, coffee cultivation between 1920 and 1970 connected rural areas to national and international markets and required long-term investment, seasonal labor, and access to credit, all of which depended on functioning market institutions. Consistent with the model, we find that the negative relationship between landowner share and tax revenues is significantly stronger in coffee-growing municipalities, where the fiscal return to public investment

was likely higher.

Second, we test whether land inequality affects enforcement costs by examining tax collection efficiency, measured as the ratio of tax collection expenditures to total revenues. Municipalities with a higher landowner share, where land is more equally distributed, spend significantly more to collect each peso of tax revenue. This finding suggests that tax enforcement is more costly when land ownership is dispersed. In contrast, municipalities with high land concentration face lower enforcement costs because the government monitors fewer landowners and large landholders have stronger incentives to comply. Together, these results provide evidence that inequality can improve tax compliance by increasing the private returns to public investment and reducing the cost of enforcement.

Our findings offer a reinterpretation of the relationship between inequality and state capacity. Rather than obstructing taxation, land concentration may facilitate it when tax revenues are used for promoting markets. Our argument complements work emphasizing the importance of elite-state bargains in historical fiscal development (e.g., Hoffman and Norberg, 2002; North and Weingast, 1989; Besley and Persson, 2009), and stands in contrast to findings in more modern settings where the redistribution framework may be more relevant. For instance, Hollenbach and Silva (2019) argue that inequality reduces tax compliance in modern-day Brazil.

In highlighting the role of inequality in enabling elite coordination around market-supporting public goods, we aim to recover a political logic of state formation that is distinct from both redistribution and repression. In Colombia, during a time of limited democracy but expanding fiscal capacity, taxation was not always a threat to elites. Sometimes, it was a bargain, as it was in the African countries where Robert Bates developed this insight.

Related Literature

This paper contributes to a literature on the political economy of public goods by exploring how elite coordination facilitates state investment. Wallis (2003) shows that in mid 19th-century Indiana, property taxation aligned the interests of landowners with infrastructure provision by directly tying public investment to asset appreciation. Property taxes acted as a coordinating device, enabling investments in roads and canals that would have been difficult to achieve through voluntary contributions alone. Dell (2010) makes a complementary argument in the Peruvian context: she finds that the absence of large landowners in regions historically subject to the mita system is related to lower public investment and access to markets. However, Dell does not develop the precise mechanism linking inequality to public good provision. The argument in this paper provides a clear story linking both via tax compliance. Together, these studies highlight how land concentration can shape state formation, not only through its distributive consequences but also by determining whether the incentives of elites align with public investment.

Second, we contribute to the literature on redistribution and political conflict. Classic

models (Meltzer and Richard, 1981; Acemoglu and Robinson, 2006) emphasize elite resistance to taxation in unequal societies. While our argument produces similar empirical predictions, it relies on a different logic: elites pay taxes not to avoid redistribution, but because public expenditure increases the value of their assets.

Third, this work contributes to the literature on state capacity by highlighting the conditions under which economic elites choose to support taxation voluntarily. Earlier work emphasized supply-side investments, such as building bureaucratic or legal capacity (Besley and Persson, 2013; Dincecco and Katz, 2016). More recent accounts, including Besley (2020), recognize the importance of reciprocity: state capacity is more likely to emerge when taxpayers see direct benefits from public investment. Our contribution lies in formalizing a demand-side mechanism: when taxation improves the value of elite-owned assets, fiscal capacity can arise even in settings with limited enforcement. By demonstrating how land inequality can facilitate this process, we provide a complementary perspective to traditional accounts that focus primarily on state strength and coercion.

Fourth, our model and empirical findings contribute to the public economics literature on voluntary public good provision. Bergstrom et al. (1986) show that, in a Nash equilibrium, wealthier individuals contribute more because they derive greater benefits from the public good, a mechanism that mirrors the proportional-benefit logic in our model. Similarly, Baland and Platteau (1997, 1999) study collective action over common-property resources and find that, although wealthier agents contribute more individually, the overall effect of inequality on collective outcomes is theoretically ambiguous because free-riding by smaller agents and distributional conflicts may offset these contributions. Our framework provides new insights into this literature in the context of property taxation, where public investments raise land values and tax liabilities are relatively easy to enforce. This setting is particularly relevant in early stages of development, when wealth is concentrated in land, an observable and immobile asset that governments can monitor at relatively low cost, especially when concentrated in the hands of a few.

Finally, we contribute to Latin American economic history by examining the role of land inequality in shaping fiscal institutions. Engerman and Sokoloff (2002) argue that factor endowments in the Americas shaped differences in inequality, which led to the formation of institutions that privileged elites and limited broad-based investment in public goods. Building on this, Sokoloff and Zolt (2007) show that countries with these colonial legacies often relied on indirect taxes rather than direct taxation of income or property, reflecting elite resistance to fiscal burdens. In contrast, we study landowners' decisions for a given tax system and show that municipalities in Colombia with more concentrated land ownership collected more property taxes per capita. This result departs from the cross-country pattern documented by Sokoloff and Zolt and suggests that inequality can, under specific conditions, facilitate local state capacity. When public spending raises land values, as in our framework, fiscal contributions become aligned with landowners' interests. Our findings offer a new perspective on how inequality shaped the incentives for

public good provision at the subnational level in Latin America.

2 Theoretical Framework

We build a simple theoretical model of public goods provision and tax compliance to explore the relationship between inequality and taxation. In particular, we adapt the Allingham and Sandmo (1972) framework to show how taxes can serve as a coordination device between landowners and the state to provide public goods. Landowners weigh the benefits of paying taxes vs. free-riding, against the expected cost of being found evading and charged a penalty rate. Our model is built around two main assumptions: (i) with tax revenues, the state funds public goods that increase land values proportionally, and (ii) the probability of the state discovering a landowner evading taxes is inversely proportional to the number of landowners in the municipality.

Assumption 1 A central assumption of the model is that local public spending increases land values proportionally. While stylized, this assumption reflects a well-documented empirical relationship between market-enabling public goods and asset values. These goods—such as property rights protection, transportation infrastructure, and legal enforcement—facilitate participation in markets by reducing uncertainty and transaction costs. Because they cannot be privately provided at scale, landowners have an incentive to support their provision.

There is substantial evidence that such public goods raise land values by increasing the returns to economic activity. Property rights enforcement, for instance, has been shown to increase investment and agricultural productivity (Besley, 1998; Hornbeck, 2010). Land titling boosts investment and housing improvements in urban settings (Galiani and Schargrodsky, 2010). Infrastructure investment raises rural land values by expanding access to markets (Donaldson and Hornbeck, 2016; Wallis, 2003). Similarly, access to utilities such as electricity and piped water raises urban and rural land values (Lewis and Severnini, 2020; Kitchens and Fishback, 2013; Coury et al., 2022). Complementing this, Boehm (2022) finds that weak legal enforcement distorts input use across sectors, reducing aggregate productivity—underscoring the broader gains from contract enforcement institutions. Taken together, these findings support the view that when the state provides public goods that facilitate exchange, the value of land increases. This underpins our modeling choice: taxation can be viewed not as redistribution but as a means to fund investments that benefit landowners directly through higher asset values.

Assumption 2 We assume that the probability of detecting tax evasion is inversely related to the number of landowners in a municipality. This reflects a setting in which the state has limited administrative capacity and imperfect information about the full

land distribution—particularly when cadastral records are outdated or incomplete. Under such conditions, it is more plausible that authorities can identify landowners when they are few in number. This assumption is consistent with the idea that property taxes are easier to enforce when land is concentrated: land is immobile and visible, and large landowners tend to be politically and economically prominent, making them easier to monitor. Empirically, Klepper and Nagin (1989) find that perceptions of audit probability increase with the amount of tax at stake, and that tax authorities tend to focus enforcement efforts on large taxpayers. In our context, this suggests that enforcement is more credible when a small number of actors own a large share of taxable land.

2.1 Model Setup

Landowners and Land Holdings Consider a municipality m with total population normalized to 1, of which $L_m \in (0, 1)$ are landowners. Each landowner i holds a plot of size l_i, and total land value is normalized to 1:

$$\sum_{i=1}^{L_m} l_i = 1.$$

A vector $(l_1, \ldots, l_{L_m})$ fully describes the land distribution.

Fiscal Policy Each landowner faces a property tax at rate τ on their land value l_i. If a landowner complies, they pay τl_i. If they evade, they pay nothing up front but risk detection with probability $\pi(L_m)$, in which case they must pay $(\tau + p)l_i$. An evader still benefits from the public good provided by others' tax payments, though they do not contribute themselves.

We assume the probability that evasion is detected, $\pi(L_m)$, is higher when L_m is small (fewer landowners). For simplicity, we specify a linear function:

$$\pi(L_m) = 1 - L_m$$

The model could be extended to allow for a function with increasing enforcement costs, like $(1 - L_m)^\beta$.

Public Goods Given a set of compliers C, tax revenues are given by:

$$T_m = \sum_{i \in \mathrm{C}} \tau l_i$$

They finance a public good that raises each landowners' land value to $s(T_m)l_i$. The function $s(\cdot)$ is increasing in total tax revenue T_m, capturing the idea that public goods and services (roads, security of property rights, market infrastructure) increase land productivity. Examples of $s(\cdot)$ include:

$$s(T) = \theta T \quad \text{or} \quad s(T) = \theta T^{\gamma}$$

In this formulation, θ represents a fiscal multiplier that summarizes the ability of government expenditure to increase land values.

2.2 Threshold Rule and Solution

Payoff Comparison Let C be the set of compliers. If landowner i complies, they pay τl_i and enjoy:

$$U_i(\text{comply}) = s(T_m^+)l_i - \tau l_i,$$

where $T_m^+ = \sum_{j \in C \cup \{i\}} \tau l_j$. If they evade, they pay nothing immediately but risk penalty with probability $\pi(L_m)$. Their expected payoff is:

$$U_i(\text{evade}) = s(T_m^-)l_i - \pi(L_m)(\tau + p)l_i,$$

where $T_m^- = \sum_{j \in C \setminus \{i\}} \tau l_j$.

Indifference Condition and Threshold A threshold strategy emerges: there exists some α such that all landowners with $l_i \geq \alpha$ comply, while those with $l_i < \alpha$ evade. For the marginal landowner at $l_i = \alpha$:

$$s(T_m^- + \tau\, l_i)\, l_i - \tau l_i = s(T_m^-)\, l_i - \pi(L_m)(\tau + p)l_i.$$

Canceling l_i and rearranging gives

$$s\big(T_m^- + \tau\, \alpha\big) - s\big(T_m^-\big) = \tau - \pi(L_m)\,(\tau + p) \tag{1}$$

This equation implicitly defines the cutoff α. When $s(\cdot)$ is linear, $s(T) = \theta\, T$, one can solve for α in closed form. Assuming $\theta = 1$ and substituting into equation (1) gives $\tau\alpha = \tau - \pi(L_m)(\tau + p)$. With $\pi(L_m) = 1 - L_m$:

$$\alpha^* = 1 - (1 - L_m)\left(1 + \frac{p}{\tau}\right). \tag{2}$$

A smaller number of landowners (L_m) or a higher penalty rate (p) lowers the compliance threshold (α^*), increasing the number of landowners who choose to pay taxes. By contrast, a higher tax rate (τ) raises the threshold and reduces compliance.

In this linear version of the model, each landowner has a strictly dominant strategy: comply if ($l_i \geq \alpha^*$) and evade otherwise. The equilibrium is therefore characterized by all landowners with ($l_i \geq \alpha^*$) choosing to comply.

2.3 Comparative Statics

We now examine how tax revenues are affected by changes in the number of landowners L_m and the distribution of l_i using the simplest version of our model, where $s(T_m) = T_m$ (i.e., $\theta = 1$) and $\pi(L_m) = 1 - L_m$.

We first study how T_m changes with L_m, assuming each landowner owns the same amount of land $1/L_m$, so that all cross-municipality differences in inequality come from the extensive margin. Since total land is normalized to 1 and the distribution is uniform, $l_i \sim U(0,\, 2/L_m)$. We can therefore define total tax revenue as:

$$T_m = \tau\left(1 - \frac{(L_m\alpha^*)^2}{4}\right).$$

See Appendix A.1 for the complete derivation.

Proposition 1 *Higher land concentration or a lower L_m unambiguously increases tax revenues in municipality m. In particular, the marginal change in T_m is given by:*

$$\frac{\partial T_m}{\partial L_m} = -\frac{\tau L_m \alpha^*}{2}\left(\alpha^* + \frac{L_m(\tau + p)}{\tau}\right) < 0$$

See Proof in Appendix A.2.

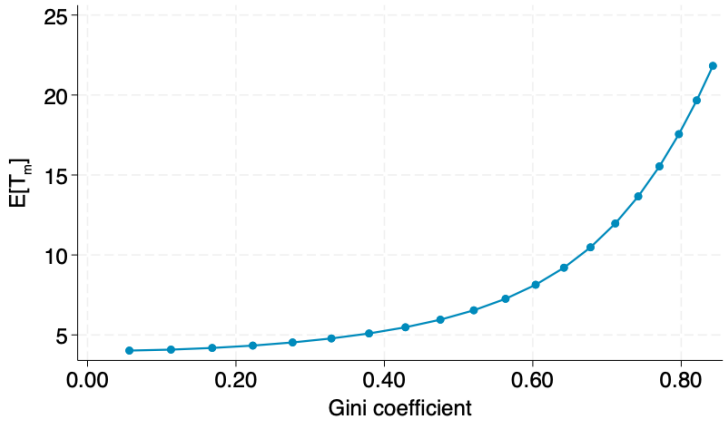
A smaller number of landowners makes tax enforcement easier, increasing the probability that an evader is detected. The higher detection risk raises the expected cost of evasion, lowers the compliance threshold α^*, and expands the set of landowners who choose to comply. As a result, tax revenues increase.

Furthermore, to study how tax revenues are affected by changes in the distribution of l_i, we rely on simulations. We fix model parameters to Colombian tax policy during the second half of the 20th century ($\tau = 0.4\%, p = 0.2\%$)[6] and hold L_m fixed at 0.5, so that all variation in predicted revenues reflects the intensive margin of inequality rather than differences in enforcement probability. We draw land shares from a LogNormal$(0, \sigma)$ distribution with $\sigma \in [0.1, 2.0]$, and measure inequality using the land Gini coefficient, computed analytically as $2\Phi(\sigma/\sqrt{2}) - 1$, where $\Phi(\cdot)$ is the standard normal CDF. For each value of σ, we numerically solve for the compliance threshold α^* and compute total tax revenue T_m. Figure 1 presents the results of these simulations and summarizes the second prediction of our model.

Proposition 2 *More unequal land distributions, represented by LogNormal$(0, \sigma)$ distributions with higher σ and thus higher land Gini, lead to higher tax revenues. Intuitively, when land is more concentrated, large owners have more incentive to comply because they benefit more from the public good (their absolute gains from improved infrastructure or security are larger).*

[6] See Section 3.

Figure 1: Gini Coefficient and $E[T_m]$



Note: This figure plots predicted tax revenues $E[T_m]$ against the land Gini from a model simulation. Land shares are drawn from a LogNormal$(0, \sigma)$ distribution with $\sigma \in [0.1, 2.0]$; the Gini is computed analytically as $2\Phi(\sigma/\sqrt{2}) - 1$, where $\Phi(\cdot)$ is the standard normal CDF. The compliance threshold $\alpha^* = 1 - (1 - L_m)(1 + p/\tau)$ is calibrated to Colombian property tax and penalty rates in 1960 ($\tau = 0.4\%$, $p = 0.2\%$) with $L_m = 0.5$ held fixed, so all variation in predicted revenues reflects the intensive margin of inequality rather than differences in enforcement probability.

3 Testing the model with municipality-level data from Colombia (1923-1960)

We use our model to analyze empirically how differences in land inequality are related to differences in tax revenues. We use data from Colombian municipalities during the mid 20th century. This context has several features that fit our model well.

First, property tax rates and penalties were set at the national level and applied uniformly across municipalities. Law 20 of 1908 established a tax rate of two pesos per thousand pesos of assessed value (0.2%), which was raised to four pesos per thousand (0.4%) through subsequent reforms, and stayed at that level for most of the century. Law 73 of 1935 introduced additional penalty of 50% for non-compliance, effectively setting the penalty rate at 0.2%. These fixed parameters meant that municipal authorities and local landowners had no formal role in setting tax rates, making it reasonable to assume landowners were tax policy takers.

Still, elites may have influenced implementation through informal bargaining. Even if they could not change nominal rates, landowners could delay cadastral updates, lobby for under-assessment, or pressure tax authorities to weaken enforcement (Sánchez-Talanquer, 2020). Such practices were more feasible in areas where political power was concentrated. Acemoglu et al. (2007) show that in 19th-century Colombia, political offices were held by a narrow subset of the wealthy. This suggests that while some powerful landowners may have influenced implementation, most were effectively subject to the tax policy as set. Our model's assumption of exogenous policy fits this majority's experience.

Second, throughout the 20th century, Colombian municipalities were in charge of collecting local taxes and funding basic public services. The property tax, established by Law 20 of 1908, and the industry and commerce tax (main sales tax), created by Law 97 of 1913, became the primary sources of municipal revenue. Municipal governments used these funds to build and maintain infrastructure such as roads, aqueducts, and sanitation systems, as well as to finance local health, education, and justice services. In 1948, Decree 2473 mandated that one-quarter of all property tax revenue be transferred to the Institute of Municipal Development (*Instituto de Fomento Municipal*), a national entity responsible for executing large-scale public infrastructure projects (Ricciulli-Marín et al., 2024). These institutional arrangements tied municipal revenue to the provision of public goods likely to increase land values rather than redistribute income.

Third, while we have limited cadastral data across time, our analysis focuses on cross-municipality variation in land inequality, which is highly persistent in Colombia. The colonial origins of large haciendas and weak redistributive land policy throughout the 19th and 20th centuries contributed to a remarkably stable landholding structure (Bértola and Ocampo, 2021; Safford and Palacios, 2002; LeGrand, 1986). Most land redistribution came from distributing public lands rather than dividing large estates, leaving the top of the land distribution largely intact. Thus, even with incomplete data, measuring land inequality at a given point in time is informative about long-run variation across municipalities.

These institutional features make Colombia in the early to mid-20th century a compelling setting to apply our model. Tax parameters were fixed nationally, limiting local discretion; municipal governments used revenues primarily to provide public goods; and the distribution of land ownership remained persistently unequal across municipalities. This combination enables us to bring the model to the data and test its predictions using historically grounded cross-sectional variation in inequality and fiscal outcomes.

3.1 Data

We build a panel of fiscal statistics for 702 municipalities from 1923 to 1960. As the set of municipalities changes over time, we create a fixed set by connecting units across four Population Censuses (1938, 1951, 1964, 1973) to the 1912 Census units. Fiscal statistics include municipal government revenues and expenditures. For the later years in our sample, income and expenditure sources, as well as the total cadastral value of land and the number of registered real estate properties, are available. More detailed cadastral data is scarce in Colombia, making it difficult to estimate land inequality. We only observe detailed cadastral records that allow us to compute a land Gini coefficient for municipalities in Cundinamarca in 1890. For the rest of the municipalities in our sample, we measure land inequality only at the extensive margin, using occupation data from the 1938 Population Census. For each municipality, we define inequality as the number of landowners as a share of agricultural workers. Inequality rises as this measure decreases. We are

interested in the relationship between the level of taxation (taxes per capita) and land concentration.

Municipal Fiscal Statistics We digitized local public finance data throughout the 20th century from three primary sources. The Office of the Comptroller General published fiscal statistics in the Yearly Statistical Reports (*Anuarios de Estadística*) from 1923 to 1936 and in the Financial Reports of the Comptroller (*Informes Financieros del Contralor*) from 1940 to 1956. For the period 1957-1967, the National Statistics Office (DANE) produced the Fiscal Statistics Report (*Estadísticas Fiscales*).

We collected information on total revenue and expenditure for all years. When available, we also digitized the main sources of tax revenue (direct and indirect taxes) and the main categories of local expenditure (governance, justice, health, education, public services, public investment, tax collection). For the year 1960, the reports include cadastral data: total value of land, and number of real estate properties registered in the cadastre.

Throughout the paper, we focus on three measures. First, tax revenue per capita. We also use tax revenue per hectare in robustness exercises. These tax revenues may come from direct or indirect taxes. Second, the total expenditure on tax collection as a proportion of total revenue. The average cost of raising one peso serves as a measure of tax collection efficiency. Third, the share of revenues devoted to infrastructure expenditures. Specifically, expenditures under the public services and public investment categories as a share of total expenditure.

Land Values For our simulations, we calibrate total land values for Colombian municipalities in 1960. We use the summary data from the fiscal statistics to obtain total land values at the municipality level. However, this figure applies only to properties registered in the cadastre. We scale up the value of land in the cadastre by dividing by the share of agricultural land with formal title in the 1960 Census of Agriculture.

Land Distribution Beyond summary statistics from the fiscal reports, detailed cadastral data for Colombia is scarce. Detailed cadastral data would allow us to measure inequality using the Gini coefficient within landowners for each municipality. We need to measure inequality for two purposes: first, to simulate landowners' decisions based on their land values and aggregate those decisions at the municipality level. Second, to measure land inequality for use in cross-municipality exercises.

For municipalities in the department of Cundinamarca, we document the detailed land distribution in 1890. Specifically, we build individual landowner shares of total land value, a vector of $l_i = [1, \ldots, L_m]$ for each m. We do this using detailed cadastral data from Acemoglu et al. (2007). For our cross-sectional results, we also calculate the 1890 land Gini coefficient based on land values for each of the 94 municipalities.

We leverage detailed occupation data in the 1938 Census of Population to proxy for land inequality at the extensive margin. For each municipality, we calculate the number of landowners as a share of agricultural workers. A high landowner share implies lower land concentration. For municipalities in Cundinamarca, the 1938 landowner share is negatively related to the 1890 land Gini coefficients (Figure B.1). Even though it is a coarse measure, the 1938 landowner share captures some of the cross-municipality variation in persistent inequality. While we aim to be precise when needed, we sometimes refer to "land inequality" to mean either the 1890 land Gini or the 1938 landowner share.

Benefits from Markets An important part of our argument relates to landowners' ability to benefit from markets. To explore this mechanism, we collect indirect information on market participation from the first coffee farm census (*Colombia Cafetera*, 1928). Municipalities engaged in coffee cultivation relied heavily on several markets. They exported most of their output, hired seasonal labor, relied on credit due to a five-year lag before production, and imported some machinery. We explore how inequality relates to taxation in places that stood to benefit more from markets, defined as those above the 10th percentile of coffee trees per capita in the 1920s. These are the early adopters of this crop.

Population, Economic Activity, Historical Fiscal Capacity, Land Conflict We interpolate total population figures from five Population Censuses (1912, 1928, 1938, 1951, and 1973) to measure tax revenues per capita more precisely each year. To measure differences in economic activity, we use the manufacturing and services share of total employment in 1912. From the 1948 Yearly Statistical Reports, we collect information on meat production, measured in number of cows per capita per year. We use this, together with our measure of coffee cultivation, to control for the agrarian economic structure. From the Panel CEDE (Acevedo and Bornacelly Olivella, 2014), we obtain identifiers for municipalities with high density of indigenous groups around 1550 and for the presence of *encomiendas* established by the Spanish Crown after 1530. Finally, to control for land conflict, we construct a dichotomous indicator for the presence of peasant-landowner legal disputes between 1900 and 1931, using data from LeGrand (1986), who compiled historical records of land conflicts during this period.

4 Empirical Results

We use data from Colombian municipalities to test our simple model of public good provision using two approaches. First, we calibrate the model using Colombia's tax policy parameters and simulate tax revenues based on our derived compliance rule and landowner-level data from municipalities in Cundinamarca with cadastral records from 1890. The correlation between simulated and observed tax revenues at the municipality level is 0.94. Moreover, the observed correlation between land Gini coefficients and tax

revenues per capita is positive. When using the 1938 landowner share instead of the Gini coefficient, the correlation is negative, but the interpretation remains the same: municipalities with higher land concentration (lower landowner share) tend to raise more tax revenue. These last two results are directly consistent with the two propositions derived from our theoretical framework in Section 2.3.

Second, we extend the analysis to the full sample to test the model's main implication: inequality has a direct positive effect on tax revenues when revenues are used to fund public goods that promote markets. Land concentration and tax revenues are positively correlated across our full sample of Colombian municipalities. This relationship remains stable over time and holds when using revenue per hectare instead of revenue per capita.

There may be other explanations for this correlation. Since our model suggests a causal role for inequality in determining equilibrium tax levels, we aim to go beyond OLS results consistent with the model's predictions. Ideally, we would exploit variation in land inequality that does not affect tax revenue through other channels. In the absence of such exogenous variation, we directly investigate four potential sources of omitted variable bias: economic activity, agrarian structure, historical fiscal capacity, and land conflict. We argue that none of these alternatives is sufficient to explain away the direct role of inequality in shaping tax outcomes.

4.1 Simulations

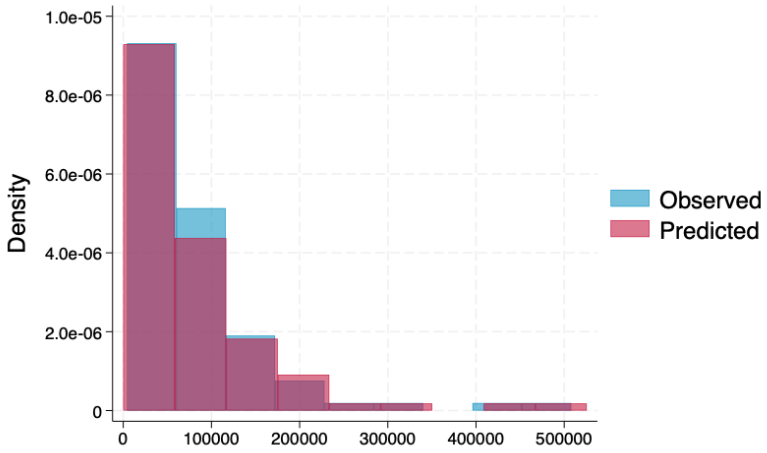
We calibrate the model to predict levels of taxation for municipalities m with different land distributions. We assume tax policy is constant across municipalities and follows the structure described in Section 2.3, with $\tau = 0.4\%$ and $p = 0.2\%$. Cadastral data from 1890 provides a sample of 94 municipalities, each with a vector of land distribution (l_i) and a count of landowners (L_m). We focus on the simplest version of the model, with the probability of enforcement linearly related to land concentration, linear costs of compliance, and a unitary fiscal multiplier.

For each municipality, we calculate α_m, the threshold of land value share above which landowners choose to pay taxes, using Equation 2. Given each landowner's land value and the threshold rule, we predict individual compliance decisions. We then aggregate compliant landowners' value shares and multiply them by the corrected 1960 land values (accounting for underreporting in the cadastre) and the property tax rate to generate a prediction for per capita tax revenue. The correlation between predicted and observed tax revenue is 0.94.

Figure 2 compares the distribution of observed and predicted tax revenues for the Cundinamarca sample, confirming the model's strong fit. Figure 3 illustrates simulated and observed relationships between land concentration and taxation. Panel 3a shows the correlation between the land Gini coefficient and taxes per capita, while Panel 3b presents the relationship between the share of landowners, L_m, and taxes per capita. The simulated patterns closely match those observed in the data and support the model's

main prediction that municipalities with greater land concentration (higher land Gini or lower share of landowners) collect higher tax revenues per capita.

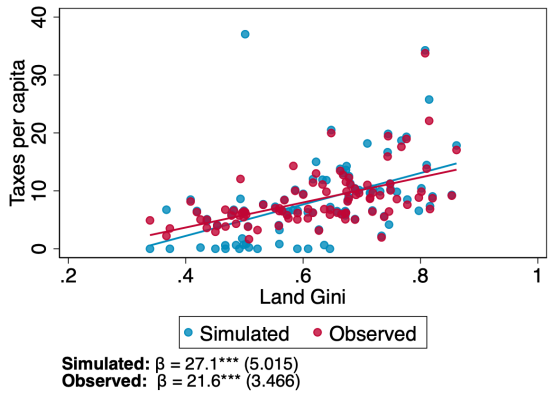
Figure 2: Predicted and Observed Tax Revenue, Cundinamarca Sample 1960



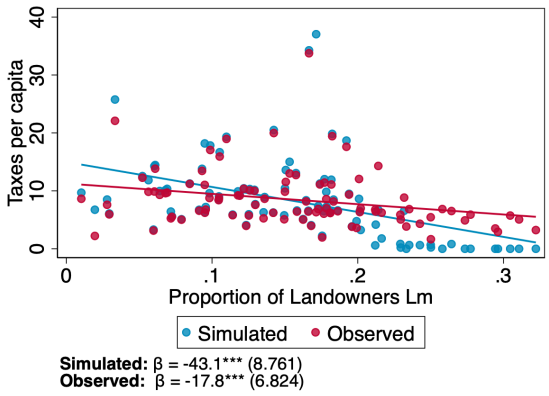
Note: This figure plots the distribution of observed versus predicted tax revenues using landowner-level data and the model structure. Landowner-level data correspond to the Cundinamarca land distribution in 1890, and model parameters are calibrated using historical data for 1960. Observed tax revenues are from Cundinamarca in 1960.

Figure 3: Taxation and Land Concentration (Cundinamarca Sample 1960)

(a) Land Gini



(b) Proportion of Landowners Lm



Note: This figure compares the simulated and observed correlations in the 1960 Cundinamarca sample. Panel (a) plots the correlation between land gini and taxes per capita; Panel (b) plots the correlation between the proportion of landowners L_m and taxes per capita. Both panels are scatter plots, where each dot represents one municipality in Cundinamarca. Correlation coefficients are reported with standard errors in parentheses.

4.2 Cross-municipality evidence

Results from the full sample, using the 1938 landowner share as a measure of inequality, tell a similar story to the Cundinamarca sample. Places where land is distributed among

more owners tend to collect lower revenues. Our (inverse) measure of land concentration, landowner share, is negatively correlated with tax revenues. Table 1, Columns (1) and (2) show the raw correlation and the within-department correlation between landowner share and tax revenues per capita in 1960, respectively. The landowner share is standardized, so coefficients reflect the change in per capita taxes associated with a one standard deviation increase in landowner share. The raw correlation indicates a decline of 24% relative to the mean, while the within-department correlation indicates a decline of about 14%.

Table 1: Share of Landowners and Taxation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: Tax revenue per capita							
Source:		All		Direct	Indirect		All	
Landowner share L_m (SD)	-1.789***	-1.088***	-0.864***	-0.739***	-0.125*	-0.678***	-0.669***	-0.670***
	(0.202)	(0.210)	(0.196)	(0.153)	(0.072)	(0.210)	(0.205)	(0.205)
Coffee trees per capita						-1.183	0.136	0.221
						(3.045)	(3.070)	(3.058)
Cattle production per capita						14.132***	14.274***	14.195***
						(5.008)	(5.365)	(5.315)
Indigenous population, 1535 (=1)							0.445	0.458
							(0.554)	(0.549)
Encomienda, 1550 (=1)							1.110**	1.148**
							(0.469)	(0.473)
Land conflict, 1900-1931 (=1)								0.335
								(0.744)
Observations	702	702	702	702	702	702	702	702
R-squared	0.063	0.261	0.351	0.354	0.284	0.361	0.367	0.367
State FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment Share Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Capital Dummy	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	7.519	7.519	7.519	5.574	1.945	7.519	7.519	7.519
Oster δ								1.077

Note: This table presents the coefficients of a regression of taxes per capita on the standardized share of landowners for 702 municipalities. Coefficients can be interpreted as the change in taxes per capita associated with a one standard deviation increase in the landowner share. All columns, except (4) and (5), use total tax revenues as the dependent variable. Columns (4) and (5) use direct and indirect taxes per capita, respectively. Employment controls include manufacturing and services employment shares. Oster δ is reported for column (8) using $R_{max} = 1.3 \times \tilde{R}^2$. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The negative correlation is relatively stable over time, despite major changes in tax policy, external shocks, political violence, and economic development between 1923 and 1960 that may have had differential effects for different levels of inequality (Figure B.2). For instance, Colombia implemented the first land reform, established universal male suffrage, and reformed the tax code between 1935 and 1936. However, the correlation between landowner share and tax revenues is not statistically different before 1935 than after. We do not see a difference in the correlation coefficients for years before the start of civil war (1943), in the middle of the war (1953), or 5 years after the National Front accord ended partisan warfare (1962).

While we are testing a causal claim that comes out of our model, the correlations in Table 1 cannot be interpreted causally. There might be countless factors that are correlated with both land inequality and the local government's fiscal capacity, which would inval-

idate the causal interpretation of the correlations. Ideally, we would identify variation in inequality that affects tax revenue only through the distribution of land. However, this is challenging for two reasons.

First, tax revenues come out of equilibrium decisions by taxpayers and governments. Any factor correlated with the willingness to pay or the ability to raise taxes would not be a valid instrument. Second, inequality in the distribution of land has been extremely persistent in Colombia. The correlation between our two measures of inequality, spaced out by 50 years, is around 0.6. Potential sources of variation could be either shocks and policies that change the distribution of land or deep-rooted factors that explain variation in persistent inequality. However, in the first case, we would also require two equivalent measures of inequality over time (at least). But even with a good measure of changes in land inequality over time, most shocks (e.g., trade or natural disasters) and most policies that deal with land distribution will also affect landowners' willingness to pay taxes directly. For instance, land reforms may generate quasi-exogenous variation in inequality but by reconsidering property rights they may also affect landowners' willingness to pay taxes.

Although we cannot fully resolve endogeneity, we assess whether four potential sources of omitted variable bias drive our results: economic activity, agrarian structure, historical fiscal capacity, and land conflict. While this may not be a comprehensive analysis, we cover the main potential issues that could explain away our results. We think the data are consistent with a direct positive relationship between land inequality and the government's ability to raise taxes.

Addressing Potential Sources of Omitted Variable Bias

While it may be impossible to find the true causal estimate of land inequality on tax revenues, we do argue that inequality has a direct effect on tax collection. In our model the mechanism is the higher willingness to pay taxes of landowners in unequal municipalities when they stand to gain from government revenues via increases in land value. In this section, we examine alternative explanations for our findings.

Economic Activity, Economies of Scale Places with high land inequality may also be richer and more productive. Differences in the level of economic activity or the output per hectare may be driving the correlation between land inequality and tax revenues. For instance, if there are economies of scale in agriculture, the average land value of large farms may be higher in more unequal municipalities, leading to higher tax revenues.

We explore whether these explanations are consistent with the data in two ways. First, the correlation between inequality and tax revenue is robust to controlling for the 1912 manufacturing and services employment share and an indicator variable for department capitals (Column (3), Table 1), which are a direct measure of cross-sectional differences in economic development. We include these controls in all other specifications. The co-

efficient drops in magnitude, to about -11% of the average revenue per capita. Second, we divide tax revenues by source. Indirect taxes (mostly sales taxes) might respond more to differences in economic activity across municipalities. Columns (4) and (5) show that the correlation between direct tax revenue and inequality is roughly twice as large as the correlation between indirect taxes and land inequality (13.2% and 6.4% relative to the average, respectively). The overall correlation between land inequality and tax revenues is mostly driven by direct and not indirect taxes.

Agrarian Structure The structure of the agricultural sector in each municipality may be an important omitted factor. If the "basket" of products a place produces is correlated with both the land distribution and the ability to raise taxes, there would be identification concerns. For instance, cattle ranching in Colombia is a land-intensive activity with economies of scale (Meisel, 2023), which leads to municipalities specializing in cattle to have very unequal land distributions. If producing cattle made it easier for the government to assess land values, for instance, there would be an omitted variable bias concern.

Results in Column (6) show that the correlation between inequality and tax revenues is robust to controlling for two relevant dimensions of the agrarian structure: coffee cultivation (measured as coffee trees per capita) and cattle ranching (measured as the number of heads of cattle slaughtered per person in 1948). Cattle production is in fact positively correlated with tax revenues. Controlling for these two measures of the agrarian structure does lead to a reduction in the correlation coefficient's magnitude to about 9% of the average revenue per capita for a change of one standard deviation in a municipality's landowner share.

Historical Fiscal Capacity It might be the case that historical fiscal capacity is driving the results. The Spanish Crown distributed rents from Conquest in the form of fiscal institutions, called *encomiendas*. Encomiendas, beyond distributing indigenous labor and tribute, laid the foundation for the Colonial and modern state. Faguet et al. (2024) document that places with encomiendas in 1550 are places with higher state capacity throughout time, with higher ability to raise taxes and provide public goods. As encomiendas were grants that depended on a relationship with the Crown, places with high prevalence of encomienda are also unequal places, where the initial distribution of resources to a limited number of conquistadores persisted over time.

The coefficient for landowner share is robust to controlling for the presence of indigenous groups in 1535 and the presence of encomiendas in 1550 (Column (7)). As emphasized by Faguet et al. (2024), places where the Spanish Crown established encomiendas raise more tax revenues per capita in 1960. The presence of indigenous groups, a coarse measure of population density around 1500, is not correlated with taxes per capita. Only the historical presence of fiscal institution is correlated with tax collection in the mid 20th century, not the initial population density. But controlling for these differences does not

lead to a change in the coefficient of interest.

Redistribution Under Threat An additional concern is that the positive association between land concentration and tax revenues may reflect landowners' strategic response to the threat of peasant rebellion, as formalized by Acemoglu and Robinson (2006), rather than voluntary compliance driven by the benefit from public good provision and the threat of detection. To study this potential mechanism, we control for the extensive margin of peasant-landowner legal disputes between 1900 and 1931, using data from LeGrand (1986), who compiled historical records of land conflicts during the early 20th century. Colombian historiography assigns considerable explanatory power to the idea that peasant revolts respond to historic land conflicts, and we interpret this dichotomous measure as a coarse proxy for latent revolutionary threat. Controlling for this variable does not meaningfully change the coefficient on landowner share (Column (8)). The conflict measure itself is positively related to tax revenues but the coefficient is small and not statistically significant, offering no support for the redistribution under threat mechanism.

Sensitivity to Unobserved Confounders To bound the overall threat from unobserved confounding, we compute the Oster (2019) δ statistic. Under the assumption that selection on unobservables is proportional to selection on observables, δ measures how much more important unobservables would need to be, relative to the full set of included controls, to drive the estimated coefficient to zero. Following Oster (2019), we set $R_{\max} = 1.3 \times \tilde{R}^2$, where $\tilde{R}^2$ is the R^2 from the fully controlled regression, as a benchmark for the maximum variance that a regression including both observed and unobserved controls could explain. A value of $\delta \geq 1$ is typically taken as evidence of robustness to unobserved confounding, since it requires unobservables to be at least as jointly explanatory as all observed controls. For our preferred specification (Column (8)), $\delta = 1.077$, indicating that unobservables would need to account for slightly more variation than our entire set of observed controls to fully explain away the estimated relationship. Given the extensive set of controls we include, this threshold is unlikely to be met, lending further credibility to the estimated effect.

Land extension Finally, it could be that places with high land concentration have also more extensive plots of land and more hectares of land in the aggregate, resulting in higher property tax revenues. To rule out this explanation, we reproduce Table 1 using taxes per hectare of land as main dependent variable. Table B.1 shows that our main results remain unchanged using this alternative definition of tax collection.

Coordination and Markets

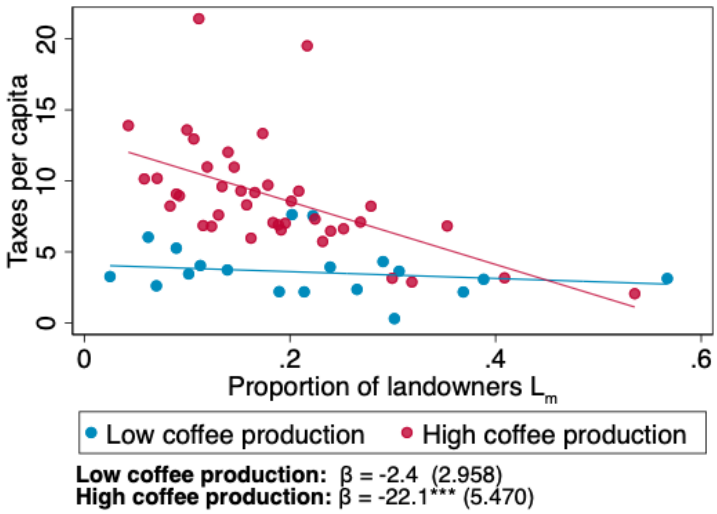
So far, we have presented evidence that aligns with the model's main prediction: land inequality is positively associated with tax revenues, even after accounting for poten-

tial sources of omitted variable bias. This section turns to the model's two underlying mechanisms and presents empirical evidence consistent with each. First, land inequality facilitates coordination among landowners to pay taxes when the government uses tax revenue to finance public goods that raise land values. Second, higher land concentration increases tax collection efficiency, since landowners have fewer peers to hide behind.

Gains from Markets We compare the correlation between land inequality and tax revenues for places specializing in coffee cultivation against other municipalities. The first part of the model's argument relies on the idea that public goods may enhance markets, increasing land values. Those gains will be concentrated in places that stand to gain a lot from the development of markets. In the Colombian setting, those are distinctively the places engaged in coffee cultivation. Between 1920 and 1970, coffee was the country's most important economic activity. Coffee exports were 2/3 of total exports and the main source of foreign exchange. The bulk of these exports was produced by small to medium farms in the Andean region, and most of the production was exported (Palacios, 2002).

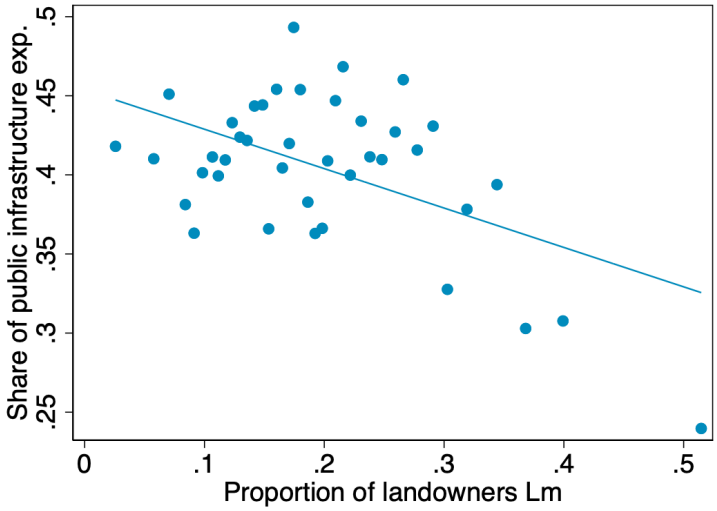
Landowners in places specializing in coffee cultivation frequently interacted in several markets. Starting a coffee farm required considerable capital, since trees start producing coffee cherries after 5 years of being planted. A single, labor-intensive harvest season generated labor shortages where labor markets did not function properly. Most importantly, long distance trade was built around financial contracts for the exchange of goods, letters of exchange, and foreign currency. In the Colombian context, coffee growers stand to gain a lot from public goods that enhance how markets operate (Deas, 1976).

Figure 4: Taxation and Land Concentration by Coffee Production



Note: This figure plots a binscatter that illustrates the correlation between the proportion of landowners L_m and 1960 tax revenues per capita, separately for municipalities with high coffee production (above the 10th percentile of coffee trees per person in 1925) and low coffee production (rest).

Figure 5: Public Infrastructure Expenditure and Land Concentration



Note: This figure shows a binscatter illustrating the relationship between the share of landowners (L_m) and the share of public infrastructure expenditures. Infrastructure expenditure includes public services spending, which at the time consisted primarily of investments in water and sewage infrastructure.

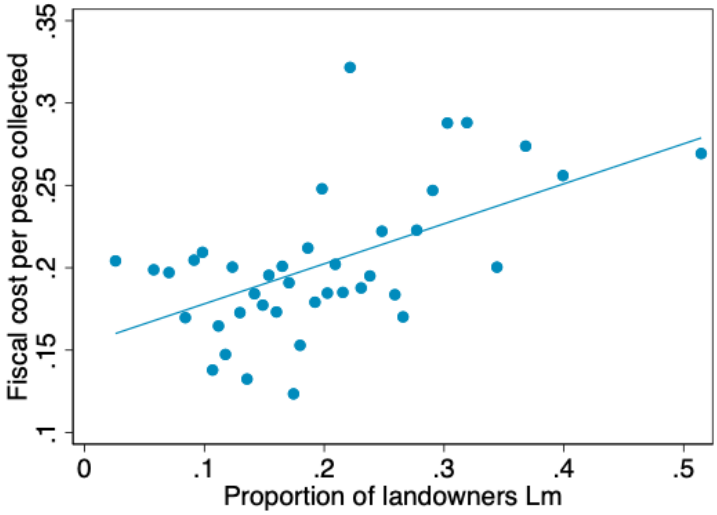
Figure 4 compares the relationship between land inequality and tax revenues for two groups of municipalities. Diamonds represent municipalities beyond the 10th percentile in the 1925 coffee trees per capita distribution. Circles represent municipalities not engaged in coffee cultivation. The relationship is steeper for the first group, coffee growing municipalities. Figure 5 further shows that in municipalities with higher land concentration, public spending is disproportionately directed toward investments that raise land values, such as infrastructure and basic public services (e.g., water and sanitation). Taken together, these findings support our proposed mechanism: when public revenues increase land values, particularly in places that stand more to gain from markets, large landowners have stronger incentives to comply with taxation. While there may be other explanations for this result, we think it provides compelling evidence about the role of markets to understand when landowners can coordinate into paying taxes.

Tax collection efficiency Our model also argues that land concentration increases tax collection efficiency, leading to higher tax revenues. To explore this idea, Figure 6 presents the correlation between land inequality and tax collection efficiency, measured as the total expenditure on tax collection over total revenue. On average, places with higher inequality (to the left on the x-axis) spend considerably less for each peso of revenue.

This pattern is consistent with the model's enforcement mechanism: when land ownership is concentrated among a few large landowners, the government faces a lower fiscal cost of identifying and monitoring taxpayers. Land is an immobile, visible asset, so the location, extent, and value of large landholdings are comparatively easy for tax authori-

ties to observe and assess when there are fewer landowners, unlike settings where wealth is dispersed across many small holdings. As enforcement becomes more credible, large landowners face a higher probability of detection if they evade, which raises the expected cost of evasion, increasing compliance and total tax revenues.

Figure 6: Fiscal Cost of Tax Collection and Land Concentration



Note: the figure shows a bin-scatter that illustrates the correlation between 1938 landowner share (landowners as proportion of agricultural workers) and 1960 tax efficiency for the full sample. Tax efficiency is defined as total expenditure in tax collection over total revenue.

5 Discussion of Results and Conclusion

"One of the basic arguments linking political centralization with economic reward rests upon the desire of people to benefit from the gains in welfare which can be reaped from markets. (...) In the context of this argument, the contribution of the state is to provide order and peace and thereby to render production and exchange possible for members of society. The origins of the state, then, lie in the welfare gains that can be reaped through the promotion of markets."

Robert Bates (1987). Essays on the Political Economy of Rural Africa.

This paper is an attempt to examine Robert Bates' insight about the process of state formation and its relationship with development. Based on Adam Smith's ideas in Book V of the Wealth of Nations, Bates summarizes an argument that had a long-standing tradition in explaining Sub-Saharan Africa's pattern of development (Fenske, 2014; Herbst, 2014; Osafo-Kwaako and Robinson, 2013; Hopkins, 2019). Gains from centralization are proportional to the welfare gains from engaging in markets.

We translate this idea into a tax compliance model, where landowners decide whether to pay taxes or evade them, taking tax policy as given. Two mechanisms drive our results.

First, public investment raises the returns to land, giving landowners a direct private incentive to fund the state through taxes. Second, tax enforcement is more effective when wealth is concentrated, because fewer, wealthier taxpayers are easier to monitor.

We interpret "gains in welfare which can be reaped from markets" as improvements in the marginal returns to production inputs that get priced in land values. Importantly, if we treated the gains from government action as pure public goods, we would get a classic free-rider result, with a different prediction for the relationship between inequality and tax revenues.

Historical data from twentieth-century Colombia support the model's predictions: municipalities with more concentrated land ownership exhibit higher tax compliance and greater local revenues. The evidence also supports the model's two underlying mechanisms: the positive relationship is stronger in places with greater potential gains from market development, and land concentration reduces the average fiscal cost of collecting one peso of revenue.

The positive relationship between inequality and taxation presented in this paper is context-dependent and reflects conditions characteristic of early stages of development, when wealth is concentrated in visible and immobile assets such as land, and public spending primarily facilitates market development rather than redistribution. As economies develop, both mechanisms may weaken. The increasing importance of mobile financial assets makes wealth easier to conceal and reduces the enforcement advantages associated with concentration, while a larger redistributive role of the state lowers the private benefits that wealthy taxpayers derive from financing public expenditure. Under these conditions, higher inequality may no longer increase tax revenues and can instead reduce tax compliance, consistent with evidence from contemporary economies (Hollenbach and Silva, 2019).

References

Acemoglu, D., Bautista, M. A., Querubín, P., and Robinson, J. A. (2007). Economic and Political Inequality in Development: The Case of Cundinamarca, Colombia.

Acemoglu, D. and Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Acevedo, K. and Bornacelly Olivella, I. D. (2014). Panel municipal del CEDE.

Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1(3):323–338.

Baland, J.-M. and Platteau, J.-P. (1997). Wealth inequality and efficiency in the commons, Part I: The unregulated case. *Oxford Economic Papers*, 49(4):451–482.

Baland, J.-M. and Platteau, J.-P. (1999). The ambiguous impact of inequality on local resource management. *World Development*, 27(5):773–788.

Banerjee, A., Iyer, L., and Somanathan, R. (2007). Public action for public goods. In Schultz, T. P. and Strauss, J. A., editors, *Handbook of Development Economics*, volume 4, pages 3117–3154. Elsevier.

Bates, R. H. (1987). *Essays on the Political Economy of Rural Africa*. University of California Press.

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, 76(2):169–217.

Bergstrom, T., Blume, L., and Varian, H. (1986). On the private provision of public goods. *Journal of Public Economics*, 29(1):25–49.

Bértola, L. and Ocampo, J. A. (2021). *El desarrollo económico de América Latina desde la independencia*. Fondo de cultura económica.

Besley, T. (1998). Investment incentives and property rights. In Newman, P., editor, *New Palgrave Dictionary of Economics and the Law*. Palgrave Macmillan, New York.

Besley, T. (2020). State Capacity, Reciprocity, and the Social Contract. *Econometrica*, 88(4):1307–1335.

Besley, T. and Persson, T. (2009). The Origins of State Capacity: Property Rights, Taxation, and Politics. *American Economic Review*, 99(4):1218–1244.

Besley, T. and Persson, T. (2013). Taxation and Development. In Auerbach, A. J., Chetty, R., Feldstein, M., and Saez, E., editors, *Handbook of Public Economics*, volume 5, pages 51–110. Elsevier.

Boehm, J. (2022). The Impact of Contract Enforcement Costs on Value Chains and Aggregate Productivity. *The Review of Economics and Statistics*, 104(1):34–50.

Coury, M., Kitagawa, T., Shertzer, A., and Turner, M. (2022). The Value of Piped Water and Sewers: Evidence from 19th Century Chicago.

Deas, M. (1976). Una hacienda cafetera de cundinamarca: Santa bárbara (1870-1912). *Anuario colombiano de historia social y de la cultura*, (8):74–100.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, 78(6):1863–1903.

Dincecco, M. and Katz, G. (2016). State Capacity and Long-Run Economic Performance. *The Economic Journal*, 126(590):189–218.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American Economic Growth: A "Market Access" Approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Engerman, S. L. and Sokoloff, K. L. (2002). Factor Endowments, Inequality, and Paths of Development among New World Economies. *Economía*, 3(1):41–109.

Faguet, J.-P., Matajira Gaitán, C. A., and Sánchez Torres, F. J. (2024). Encomienda, the Colonial State, and Long-Run Development in Colombia.

Fenske, J. (2014). Ecology, Trade, and States in Pre-Colonial Africa. *Journal of the European Economic Association*, 12(3):612–640.

Galiani, S. and Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, 94(9):700–729.

Garfias, F. and Sellars, E. A. (2026). Fiscal legibility and state development: Theory and evidence from colonial mexico. *American Journal of Political Science*, 70(1):5–20.

Herbst, J. (2014). *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press, Princeton, New Jersey.

Hoffman, P. T. and Norberg, K. (2002). *Fiscal Crises, Liberty, and Representative Government 1450-1789*. Stanford University Press.

Hollenbach, F. M. and Silva, T. N. (2019). Fiscal capacity and inequality: Evidence from brazilian municipalities. *The Journal of Politics*, 81(4):1434–1445.

Hopkins, A. G. (2019). *An Economic History of West Africa*. Routledge, London, second edition.

Hornbeck, R. (2010). Barbed Wire: Property Rights and Agricultural Development*. *The Quarterly Journal of Economics*, 125(2):767–810.

Kitchens, C. and Fishback, P. (2013). Flip the Switch: The Spatial Impact of the Rural Electrification Administration 1935-1940.

Klepper, S. and Nagin, D. (1989). Tax Compliance and Perceptions of the Risks of Detection and Criminal Prosecution. *Law & Society Review*, 23(2):209–240.

LeGrand, C. (1986). *Frontier expansion and peasant protest in Colombia, 1850-1936*. University of New Mexico Press.

Lewis, J. and Severnini, E. (2020). Short- and long-run impacts of rural electrification: Evidence from the historical rollout of the U.S. power grid. *Journal of Development Economics*, 143:102412.

Meisel, A. (2023). *Santa Bárbara de Las Cabezas: La gran hacienda del caribe colombiano 1742-194*. Universidad del Norte.

Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of political Economy*, 89(5):914–927.

North, D. C. and Weingast, B. R. (1989). Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England. *The Journal of Economic History*, 49(4):803–832.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups, With a New Preface and Appendix*. Harvard University Press.

Osafo-Kwaako, P. and Robinson, J. A. (2013). Political centralization in pre-colonial Africa. *Journal of Comparative Economics*, 41(1):6–21.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Palacios, M. (2002). *Coffee in Colombia, 1850-1970*. Cambridge University Press.

Ricciulli-Marín, D., Bonet-Morón, J., and Pérez-Valbuena, G. J. (2024). Cien años de finanzas públicas territoriales en colombia. *Lecturas de Economía*, (100):85–133.

Safford, F. and Palacios, M. (2002). *Colombia: Fragmented land, divided society*. Oxford University Press New York.

Sokoloff, K. L. and Zolt, E. M. (2007). Inequality and the Evolution of Institutions of Taxation: Evidence from the Economic History of the Americas. In *The Decline of Latin American Economies: Growth, Institutions, and Crises*, pages 83–138. University of Chicago Press.

Sánchez-Talanquer, M. (2020). One-Eyed State: The Politics of Legibility and Property Taxation. *Latin American Politics and Society*, 62(3):65–93.

Wallis, J. J. (2003). The property tax as a coordinating device: Financing Indiana's Mammoth Internal Improvement System, 1835–1842. *Explorations in Economic History*, 40(3):223–250.

Appendix A Derivations and Proof of Proposition 1

A.1 Derivation of Total Tax Revenue T_m

With total land normalized to 1 and L_m landowners, each landowner holds on average $1/L_m$, so $l_i \sim U(0, 2/L_m)$. The fraction of landowners with $l_i \geq \alpha^*$ is $1 - \frac{L_m \alpha^*}{2}$, giving a number of compliers:

$$N_C = L_m \left(1 - \frac{L_m \alpha^*}{2}\right).$$

Since $l_i \sim U(\alpha^*, 2/L_m)$, conditional on complying, the average land holding of a complier is $E[l_i \mid l_i \geq \alpha^*] = \frac{\alpha^* + 2/L_m}{2}$. Total tax revenue is the number of compliers times the tax rate times this average:

$$T_m = \tau N_C \cdot \frac{\alpha^* + 2/L_m}{2} = \frac{\tau L_m}{2}\left(1 - \frac{L_m \alpha^*}{2}\right)\left(\alpha^* + \frac{2}{L_m}\right).$$

Expanding and simplifying:

$$T_m = \tau \left(1 - \frac{(L_m \alpha^*)^2}{4}\right). \tag{3}$$

A.2 Proof of Proposition 1

We show that $\partial T_m / \partial L_m < 0$ unambiguously.

From equation (3), $T_m = \tau\left(1 - \frac{(L_m \alpha^*)^2}{4}\right)$. Differentiating with respect to L_m:

$$\frac{\partial T_m}{\partial L_m} = -\frac{\tau}{2} L_m \alpha^* \cdot \frac{\partial (L_m \alpha^*)}{\partial L_m}.$$

From equation (2), $\frac{\partial \alpha^*}{\partial L_m} = 1 + \frac{p}{\tau} = \frac{\tau + p}{\tau} > 0$, so:

$$\frac{\partial (L_m \alpha^*)}{\partial L_m} = \alpha^* + L_m \frac{\partial \alpha^*}{\partial L_m} = \alpha^* + \frac{L_m(\tau + p)}{\tau} > 0.$$

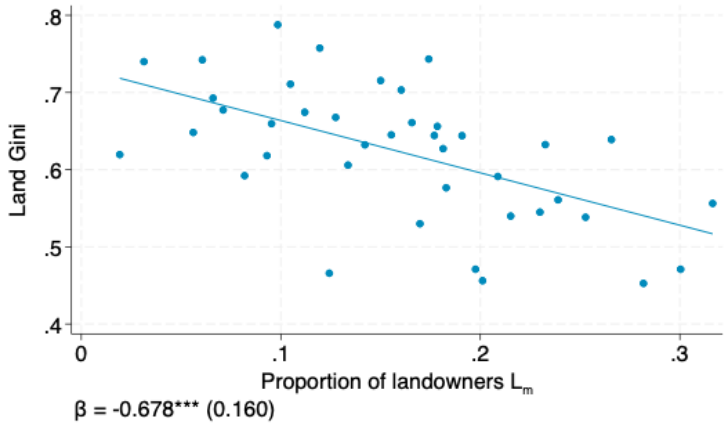
Substituting:

$$\frac{\partial T_m}{\partial L_m} = -\frac{\tau L_m \alpha^*}{2}\left(\alpha^* + \frac{L_m(\tau + p)}{\tau}\right) < 0.$$

Since $L_m > 0$, $\alpha^* > 0$, $\tau > 0$, and $p > 0$, the derivative is unambiguously negative. A decrease in L_m (higher concentration) lowers α^* and increases the share of land owned by compliers, unambiguously raising total tax revenues. □

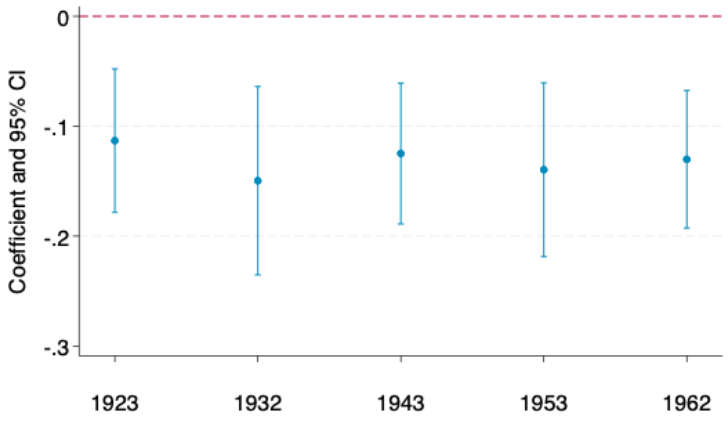
Appendix B Supporting Exhibits

Figure B.1: Share of Landowners Lm and Land Gini Cundinamarca 1890



Note: This figure plots a binscatter that illustrates the relationship between the share of landowners in 1938 and the 1890 land Gini for each municipalites in the Cundinamarca sample. Correlation coefficient with standard error in parentheses is reported.

Figure B.2: Taxation and Land Concentration Over Time



This figure plots the correlation between the share of landowners and tax revenues per capita for selected years, estimated from a regression that includes all controls from column (8) of Table 1. Given the changes in levels of taxation, the dependent variable in this case is the natural logarithm of taxes per capita.

Table B.1: Share of Landowners L_m and Taxation per Hectare

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable: Tax revenue per Hectare							
Source:		All		Direct	Indirect		All	
Landowner share L_m (SD)	-1.872***	-1.657**	-1.234**	-0.876**	-0.357*	-1.221*	-1.229*	-1.223*
	(0.565)	(0.649)	(0.612)	(0.417)	(0.205)	(0.676)	(0.687)	(0.684)
Coffee trees per capita						-8.231*	-5.863	-6.058
						(4.311)	(4.462)	(4.532)
Cattle production per capita						4.996	8.494	8.697
						(7.028)	(6.922)	(6.863)
Indigenous population, 1535 (=1)							-0.674	-0.703
							(1.139)	(1.148)
Encomienda, 1550 (=1)							3.223***	3.112***
							(0.856)	(0.849)
Land conflict, 1900-1931 (=1)								-0.968
								(1.059)
Observations	676	676	676	676	676	676	676	676
R-squared	0.019	0.090	0.200	0.201	0.188	0.201	0.210	0.210
State FE	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Employment Share Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Capital Dummy	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dep. Var.	7.442	7.442	7.442	5.479	1.963	7.442	7.442	7.442
Oster δ								2.729

Note: This table presents the coefficients of a regression of taxes per hectare (1960) on the share of landowners L_m (1938) for 676 municipalities that have information on hectares of land from the agricultural census. Proportion of landowners is standardized, so coefficients can be interpreted as the change in taxes moving one standard deviation. Dependent variable is the level of pesos per hectare of land raised in taxes in 1960. All columns, except for (4) and (5), use total tax revenues. Columns (4) and (5) use direct and indirect taxes per hectare of land as dependent variable. Employment controls are the 1912 manufacturing and services employment shares. Oster δ is reported for column (8) using $R_{max} = 1.3 \times \tilde{R}^2$. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$