DISRUPTING STATE LEGITIMACY: Evidence from 43,000 Colonial Documents*

Augustin Bergeron and Abdoulaye Ndiaye

September 9, 2026

Abstract

Does state legitimacy shape citizens' resistance to state demands and, ultimately, state capacity? We study colonial Nigeria, where British rule imposed similar fiscal institutions across neighboring areas but disrupted an established religious fiscal order in the former Sokoto Caliphate. Under Sokoto, Islamic law defined both citizens' tax obligations and rulers' authority to tax. Colonial rule replaced Sokoto's fiscal rules with a common colonial regime that was also imposed on neighboring areas, where it did not displace a comparable pre-existing order. We assemble 43,648 tax-related colonial documents and use large language models to measure tax sentiment and formal grievances. Exploiting the former Sokoto boundary in a spatial regression discontinuity design, we find substantially greater tax resistance where colonial rule displaced Sokoto's fiscal order. The additional resistance is disproportionately associated with Muslims, partly concentrated on a tax rejected under Sokoto rule, and more likely to challenge the colonial state's right to tax. Within former Sokoto, resistance is also greater where exposure to Sokoto's state-building and administrative institutions was stronger. Differences in colonial taxation, administrative capacity, enforcement, and public-goods provision do not explain the results. Our findings show that state capacity depends not only on states' ability to administer and enforce policy, but also on whether citizens regard their demands as legitimate.

*Bergeron (abergeron@fas.harvard.edu) - Harvard University and NBER; Ndiaye (an1490@stern.nyu.edu) - New York University. For outstanding research assistance, we thank Fatima Abba, Pearl Anson-Tsang, Henry David Jackson, Adrien Jalouneix, Gabriele Mole', Diana Nichvolodova, and Kelvin Torto. We thank Vittorio Bassi, Sam Bazzi, Anne Brockmeyer, Katherine Casey, Melissa Dell, Mike Droste, Edward Glaeser, Anders Jensen, Yasir Khan, Gabriel Kreindler, David Lagakos, Paul Lovejoy, Ben Marx, Eduardo Montero, Suresh Naidu, Paulina Oliva, Elias Papaioannou, Nicola Persico, Nancy Qian, James Robinson, Simone Schanner, David Schonholzer, Ambra Seck, Marco Tabellini, Leonard Wantchekon, Jeffrey Weaver, Jonathan Weigel, and Noam Yuchtman for excellent comments; and seminar participants at Kaduna State University and USC.

1 Introduction

State capacity is central to economic development: states enforce rules, provide public goods, and raise the revenue needed to finance them (Besley and Persson 2009, 2013). A growing empirical literature shows how investments in administration, territorial organization, and fiscal institutions expand states' ability to raise revenue, enforce policy, and provide public goods (Dincecco and Katz 2016; Chambru et al. 2024; Cantoni et al. 2024; Chiovelli et al. 2025). Yet the effectiveness of state-building may depend not only on the administrative institutions states create, but also on whether citizens view the authority behind state demands as legitimate. Indeed, theoretical, historical, and observational studies argue that the perceived legitimacy of the state shapes compliance and resistance (Levi 1988; Tyler 2006; Besley 2020; Robinson 2023; Greif and Rubin 2024; Besley and Dray 2024). Causal evidence, however, remains scarce. This paper provides such evidence by studying a setting in which an established source of state legitimacy was abruptly displaced.

Our setting is a disruption of fiscal authority in colonial Nigeria. Before British conquest, much of northern Nigeria belonged to the Sokoto Caliphate (1804–1903), a centralized Islamic polity in which taxation was embedded in a religious fiscal order. Islamic law defined subjects' fiscal obligations, understood in part as religious duties, and constrained rulers' authority to tax; attempts to impose taxes outside that framework provoked opposition (Last 1967; Lovejoy 2016). British occupation transformed this order. Official collection of *zakat* (a form of Islamic wealth tax) as a distinct religious obligation was curtailed, religious distinctions governing liability for poll and land taxes were removed, and *jangali*, a cattle tax rejected by Sokoto reformers as lacking a basis in Islamic law, was reintroduced. Colonial rule therefore changed both *who* claimed the authority to tax and *what* taxes members of different religious groups were required to pay. Elsewhere in Nigeria, British rule likewise expanded and reorganized taxation but generally did not displace a comparable centralized, religiously legitimated fiscal order.

This contrast provides unusual variation in a historically rooted source of legitimacy. While much of the economics literature emphasizes legitimacy derived from state performance—effectiveness, fairness, enforcement, or public-goods provision (Levi 1988; Tyler 2006; Besley 2020; Besley and Dray 2024; Besley et al. 2026), other work emphasizes that legitimacy can also rest on historically formed beliefs about rightful authority and the scope of legitimate state action (Greif and Rubin 2022, 2024). Colonial Nigeria allows us to distinguish this source from differences in the contemporary fiscal regime or state performance: nearby communities faced similar colonial tax institutions, but only in former Sokoto did those institutions displace an established religious fiscal order that defined both taxpayers'

obligations and rulers' authority to tax.This yields a simple prediction: if inherited sources of authority and obligation shape responses to taxation, resistance should be greater where colonial rule displaced that order.

Testing this prediction requires systematically observing tax resistance. We assemble a new corpus of 43,648 tax-related colonial documents from the three zonal offices of the National Archives of Nigeria in Kaduna, Ibadan, and Enugu, spanning the entire colonial period. The records span routine fiscal administration, including correspondence, financial and tax records, reports, petitions, and formal grievances, allowing us to trace how taxation was discussed and contested across space and over much of the colonial period. We use large language models (LLMs) to recover locations and dates and to construct two complementary measures of resistance. The first captures tax sentiment, distinguishing negative discussions of taxation from neutral or positive administrative records. The second identifies formal tax complaints, defined as written grievances about taxation submitted to colonial authorities. Because the corpus captures resistance within routine administration rather than only around known episodes of unrest, it allows us to recover less visible forms of opposition that are often missing from historical studies focused on revolts, riots, or rebellions. This is particularly valuable in coercive colonial settings, where repression can limit overt collective action, making it an incomplete measure of underlying resistance to state demands.

Because LLMs may misread complex historical text or hallucinate (Korinek 2023; Amirizaniani et al. 2024; Dell 2025), we validate the two main outcomes against independent human judgments. Five research assistants independently coded a random sample of 100 documents for tax sentiment and formal tax complaints. Following a Turing-style approach (Turing 1950; Lagakos et al. 2025), we benchmark both the human coders and the LLM against the modal human classification. On both outcomes, the LLM performs within the range of human coders.

Our main empirical strategy exploits the former Sokoto Caliphate boundary in a spatial regression discontinuity design. We compare nearby locations on opposite sides of the historical frontier that differed in their exposure to Sokoto's fiscal order but were subsequently incorporated into the same colonial state and subjected to similar tax institutions. We estimate local linear specifications in distance to the boundary, allowing slopes to differ across sides, with border-segment and archive fixed effects and a rich set of geographic and pre-Sokoto controls. We report estimates across fixed bandwidths from 50 to 200 kilometers and at the MSE-optimal bandwidth (Cattaneo et al. 2020).

The credibility of the design rests on comparability across the historical boundary. Historical accounts emphasize military reach and ecological constraints (Kirk-Greene 1965; Last 1967; Lovejoy 2016). However, we find no discontinuities in the geographic and ecological

characteristics most directly implicated by these accounts, or in a much broader set of pre-Sokoto characteristics. The boundary also cuts through several ethnic homelands (Murdock 1967), and the estimates remain similar when comparisons are restricted within ethnic territories. A second concern is that the Sokoto frontier may later have become aligned with differences in colonial fiscal policy. In practice, direct-tax institutions converged substantially across Nigeria, and the historical Sokoto boundary cut across the colonial protectorates, which were the main administrative units through which tax policy was organized. The results also remain similar when we restrict the sample to the Northern Protectorate, where the colonial fiscal framework is held fixed.

We find substantially greater tax resistance inside the former Sokoto Caliphate. At the MSE-optimal bandwidth, documents just inside the boundary are 9.3 percentage points more likely to express negative tax sentiment, about 40 percent above the mean just outside. The discontinuity in formal complaints is larger still: documents inside former Sokoto are 18.2 percentage points more likely to contain a written tax grievance, relative to a control mean of 7 percent. The effect therefore appears both in the broader language surrounding taxation and in an explicit form of resistance directed at the colonial state. It remains positive and statistically significant across the bandwidths we consider.

We conduct a series of checks addressing three threats to the interpretation of the discontinuity. First, the results could reflect differences in the composition of the surviving archive rather than differences in tax resistance. The three archives contribute different types of documents, and their composition may also vary over time. The estimates are robust to document-type fixed effects and to archive-by-decade and archive-by-year fixed effects, limiting the comparison to similar records drawn from the same archive and period. A related concern is uneven spatial coverage: some locations generate many more surviving documents than others, which could give heavily represented places disproportionate weight. We find no discontinuity in document counts at the Sokoto boundary, and the results remain significant when observations are reweighted to give locations equal weight. They are also robust to reweighting by local population from the 1952 Census, addressing the possibility that more populous locations mechanically generate more surviving records.

Second, the historical boundary is measured with some uncertainty and includes areas whose relationship to Sokoto is less clear-cut. The estimates remain similar in donut specifications that exclude observations closest to the frontier, where geocoding error is most plausible. They are also robust to alternative handling of areas adjacent to Kanem–Bornu, where some Islamic fiscal institutions were present, and to excluding the Jos Plateau, whose incorporation into Sokoto is debated in the historical literature. Third, the results are not sensitive to the econometric implementation of the spatial RD: the discontinuity in our

main outcomes remains similar under alternative spatial functional forms, while province-level clustering and spatially correlated standard errors leave the substantive conclusions unchanged.

A boundary design cannot rule out every unobserved characteristic that may change at the historical frontier. We therefore complement the spatial regression discontinuity with a distinct source of variation entirely within the former Caliphate. Sokoto institutions were not equally rooted throughout its territory: some locations were more closely connected to the process of state formation and to the administrative structures through which the Caliphate was subsequently governed. If the boundary discontinuity reflects the legacy of those institutions rather than an unrelated feature of the frontier, colonial tax resistance should also be greater in more intensively exposed locations. We construct two measures from Last (1967): proximity to Usman dan Fodio's 1804–1817 consolidation corridor, capturing exposure to the formative process of Sokoto state-building, and proximity to the Caliphate's major administrative centers, capturing connection to its subsequent administrative structure. Both predict greater colonial-era tax resistance within former Sokoto. To distinguish institutional exposure from Sokoto presence more generally, we use proximity to the royal ribats, fortified settlements associated primarily with the Caliphate's military and defensive presence, as a placebo. Proximity to the ribats does not predict greater resistance. The within-Caliphate evidence therefore points to a more specific pattern: resistance is greatest where exposure to Sokoto was most closely tied to state-building and administration, not simply where Sokoto's historical presence was stronger.

We next ask whether the form of resistance bears the imprint of the fiscal order colonial rule displaced. Sokoto's history yields two predictions specific to the fiscal-legitimacy interpretation. First, the additional resistance should be disproportionately associated with Muslims. Sokoto's fiscal order was explicitly Islamic, and colonial rule changed Muslims' fiscal obligations especially sharply: official *zakat* collection was curtailed, while poll and land taxes from which Muslims in core Sokoto areas had generally been exempt were imposed without the same religious distinction. We identify the named complainant in each formal grievance and classify extracted names as Muslim, Christian, or unclear. Formal complaints by Muslim complainants are substantially more common inside former Sokoto. Moreover, among documents containing a formal complaint, the share involving a Muslim complainant is also higher across bandwidths. This pattern is not mechanically explained by religious composition: Muslim population share is smooth at the boundary in the 1952 Census.[1] Consistent with the same interpretation, formal complaints increase with Muslim

[1]The Census was conducted nearly half a century after the fall of the Caliphate. While it cannot establish that religious composition was balanced before colonial rule, it does, however, show that the discontinuity

population share within former Sokoto, while no comparable positive relationship appears outside it.

The second prediction concerns which taxes are contested. Sokoto's fiscal history provides a particularly sharp case: the cattle levy *jangali*. Before the jihad, Usman dan Fodio's reform movement rejected *jangali* as lacking a legitimate basis in Islamic law and replaced it with the legally prescribed *zakat*; colonial authorities later reintroduced *jangali* and officially banned *zakat*. If the resistance we document reflects the displacement of Sokoto's fiscal order, complaints inside former Sokoto should therefore be disproportionately directed at *jangali*. We use a separate LLM classification to identify formal complaints centered specifically on the tax, excluding documents in which it appears only incidentally or as an administrative line item. *Jangali*-related grievances are rare outside the former Caliphate but account for part of the increase in formal complaints at the Sokoto boundary. Thus, the excess resistance is partly concentrated on a tax that the displaced fiscal order had explicitly rejected as illegitimate.

As complementary evidence, we construct a direct text-based measure of contested fiscal legitimacy, identifying documents that challenge the authority or normative basis of taxation rather than merely expressing dissatisfaction with its administration or enforcement. This measure also rises discontinuously at the former Sokoto boundary. Taken together, the three exercises point to distinct implications of the same historical mechanism: the additional resistance is disproportionately associated with Muslims, partly concentrated on a tax the displaced order had explicitly rejected as illegitimate, and more likely to challenge the rightfulness of taxation itself.

The patterns above are difficult to reconcile with several alternative explanations. Differences in colonial policy or state performance could generate greater dissatisfaction on one side of the boundary. Still, they do not naturally predict that the additional resistance should come disproportionately from Muslims, concentrate partly on a tax explicitly rejected under Sokoto rule, and more often challenge the rightfulness of taxation itself. We nevertheless examine these alternatives directly. Observable population composition appears unlikely to account for the result. If former Sokoto contained more literate, wealthier, or more Muslim populations, these differences could independently generate more written tax resistance. Yet Muslim population share, Latin-script literacy, and occupational structure are smooth at the boundary in the 1952 Census, and controlling for these characteristics leaves the results intact. Nor does the gap appear to reflect differences in colonial taxes or their burden. The results persist after 1927, when direct-tax systems had largely converged across the former

in Muslim complainants is not mirrored by a corresponding difference in Muslim population shares later in the colonial period.

protectorates, and within the Northern Protectorate, where locations on either side of the Sokoto boundary faced the same colonial tax regime. Aggregate fiscal accounts likewise do not indicate that former Sokoto was taxed more heavily.

We also find little evidence that the colonial state was weaker or more coercive inside former Sokoto. Documents from former Sokoto are, if anything, less likely to report poor tax administration; civil servants per capita vary smoothly at the boundary; and references to coercion, punishment, and tax evasion exhibit no consistent discontinuity. Finally, we assemble measures of colonial public-goods provision, including transport, medical, administrative, and communications infrastructure, and find no discontinuity at the historical boundary. Controlling for these measures leaves the main estimates essentially unchanged. Taken together, these exercises provide little evidence that differences in population composition, colonial taxation, administrative capacity, enforcement, or public-goods provision account for the resistance gap.

We finally ask whether the historical legitimacy gap remains visible today. Persistence could arise because post-independence taxation largely inherited the colonial fiscal system rather than restoring the pre-colonial Islamic order, because beliefs about the state's right to tax persist across generations, or both. Using Afrobarometer rounds 5–9 for Nigeria, we find that respondents on the former Sokoto side of the boundary are significantly less likely to agree that tax authorities have the right to make people pay taxes, with similar estimates across bandwidths. Trust in the tax department and downstream measures of tax morale point in the same direction but are estimated less precisely. The contemporary evidence is therefore consistent with a persistent difference in fiscal legitimacy, with the strongest result for whether tax authorities are seen as having the right to compel payment.

The paper makes four contributions. First, it contributes to the literature on state capacity by providing causal evidence on legitimacy as a distinct determinant of state capacity. A growing empirical literature shows how administrative, territorial, and fiscal institutions expand states' ability to raise revenue and implement policy (Dincecco and Katz 2016; Chambru et al. 2024; Cantoni et al. 2024; Chiovelli et al. 2025). Theoretical, historical, and observational work argues that state effectiveness may also depend on citizens' acceptance of state authority, which shapes compliance and resistance (Levi 1988; Tyler 2006; Besley 2020; Robinson 2023; Greif and Rubin 2024; Besley and Dray 2024), but causal evidence on this channel remains scarce. We provide such evidence by exploiting variation in whether colonial state-building disrupted an established source of legitimate fiscal authority. The results show that effective state capacity depends not only on what states can administer or enforce, but also on whether citizens accept state demands as legitimate.

Second, the paper contributes to understanding the sources of state legitimacy. Much of

the economics literature emphasizes legitimacy rooted in state performance (Besley 2020; Besley and Dray 2024; Besley et al. 2026), but legitimate authority also depends on historically formed beliefs about who has the right to rule and which actions authorities are entitled to take (Greif and Rubin 2022, 2024). We provide causal evidence on this historically rooted source of legitimacy. In colonial Nigeria, a similar fiscal regime generated substantially greater resistance where it displaced an established religious fiscal order that had defined both taxpayers' obligations and rulers' authority to tax. The results therefore show that responses to state demands depend not only on how the state performs or how its policies are administered, but also on whether those demands accord with inherited understandings of legitimate authority and obligation.

Third, the paper contributes to the literature on historical institutions. A large literature studies the persistent effects of colonial and precolonial institutions (Acemoglu et al. 2001, 2002; Banerjee and Iyer 2005; Dell 2010; Michalopoulos and Papaioannou 2013; Dell and Olken 2020; Lowes and Montero 2021; Dupas et al. 2024). Much of the work on colonialism treats colonial institutions as the relevant institutional shock, with less attention to how their consequences depend on the political, legal, and religious orders they encounter. We study this interaction directly. In former Sokoto, colonial taxation displaced a religious fiscal order that had defined both taxpayers' obligations and rulers' authority to tax; nearby areas faced colonial taxation without the displacement of a comparable source of fiscal legitimacy. More broadly, the results show that the consequences of state-building depend not only on the institutions states introduce, but also on their relationship to the institutional order already in place.

Fourth, the paper contributes to the growing use of computational methods to recover economic information from unstructured text (Gentzkow et al. 2019; Korinek 2023; Ludwig et al. 2025; Lagakos et al. 2025; Asirvatham et al. 2026). Historical archives contain rich evidence on state–citizen interactions, but much of it remains difficult to study systematically because it is dispersed across large collections of heterogeneous records. We use LLMs to recover tax attitudes, formal grievances, locations, dates, complainant characteristics, and the content of disputes from 43,648 colonial documents. This approach also expands the measurement of historical tax resistance beyond salient but rare episodes of revolt or riot, allowing us to recover less visible forms of opposition to taxation systematically across space and time (Burg 2004). More broadly, the paper illustrates how computational methods can turn large historical archives into systematic measures of political and economic behavior that can be combined with modern empirical designs.

The remainder of the paper proceeds as follows. Section 2 presents the historical setting; Section 3 presents the archival corpus, the outcomes, and their validation; and Section 4

presents the spatial RD design. Section 5 reports the main results, robustness checks, and the within-Caliphate evidence; Section 6 interprets the Sokoto discontinuity using evidence on fiscal legitimacy and alternative explanations; Section 7 documents persistence in contemporary survey data; and Section 8 concludes.

2 Historical Context

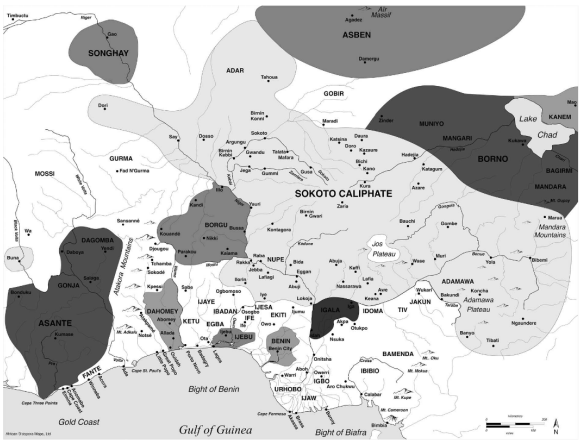
British colonization brought territories with very different pre-colonial fiscal institutions under an increasingly common system of taxation. The contrast at the center of our analysis is between the former Sokoto Caliphate, where taxation had been embedded for roughly a century in a centralized Islamic fiscal order, and neighboring areas where colonial rule did not displace a comparable religiously grounded system of fiscal authority.

2.1 The Sokoto Caliphate and Its Fiscal Order

The Sokoto Caliphate (1804–1903) was a centralized Islamic polity founded by Usman dan Fodio's reform movement (Gwandu 1977). Its emirates were governed under Islamic law through a system of courts, administration, and taxation (Last 1967; Hiskett 1973; Chafe 1999; Lovejoy 2016). Panel A of Figure 1 shows the territorial extent of the Caliphate, from Lovejoy (2016) and Figure E1 shows the movements of Usman dan Fodio during its early consolidation, mapped by Last (1967).

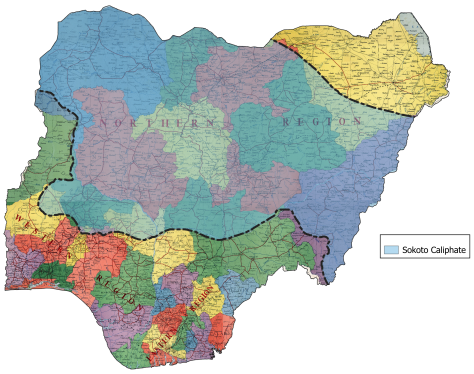
Taxation formed part of this broader Islamic legal order. *Zakat* was a religious levy on Muslims who met specified wealth thresholds, including owners of livestock and other taxable assets, and was to be redistributed among categories of beneficiaries prescribed by Islamic law, including the poor and needy. *Kharaj* was a land tax associated with non-Muslim landholders or conquered lands that remained in the possession of their occupants, while *jizya* was a poll tax imposed on non-Muslims in exchange for protection and exemption from military service (Last 1967; Hiskett 1973; Lovejoy 2016). These distinctions tied fiscal obligations to religious status and to rules governing who could be taxed and on what basis. Moreover, collection was organized through the Caliphate's administrative hierarchy: district and village authorities collected taxes locally under the supervision of emirate officials, while emirates remitted revenue to Sokoto.[2] Fiscal obligations were therefore not simply demands for revenue; they were embedded in a recognized Islamic legal framework and implemented through emirate institutions.

[2]In Muri, district heads (*hakimai*) collected major taxes, including *zakat* and *jizya*, with the assistance of village heads (*dagatai*) (Tukur 1979).



Map 40.1 The Sokoto Caliphate.
Source: Lovejoy, 2016, *Jihad in West Africa during the Age of Revolutions*, map 3.2.
Copyright: Henry Lovejoy.

(a) Historical Extent of the Sokoto Caliphate



(b) Sokoto Boundary on 1924 Map of Nigeria

Figure 1: The Sokoto Caliphate and Its Historical Boundary

Notes: Panel (a) shows the territorial extent of the Sokoto Caliphate from Lovejoy (2016). Panel (b) shows the Caliphate boundary digitized from Lovejoy (2016), georeferenced using stable historical landmarks, including named settlements and rivers, and overlaid on a 1924 map of Nigeria.

The Islamic fiscal framework did more than define taxpayers' obligations; it also distinguished legitimate from illegitimate forms of taxation and, in some cases, constrained rulers' fiscal authority. Last (1967, pp. 103–104) provides two revealing examples. The cattle levy known as *jangali*, which predated the jihad that established the Caliphate, was rejected by Usman dan Fodio and his followers as an arbitrary tax without a recognized basis in Islamic law. In its place, the reformers recognized *zakat*, the religiously prescribed levy on livestock and other qualifying wealth. Separately, when Amir al-mu'minin 'Umar attempted to introduce a form of head tax, public opposition prevented him from doing so on grounds of Islamic law:

> *"Sokoto was exceptional in that no tax other than alms (zakāt) was formally collected from Muslims.*[49] *A tax on the cattle Fulani, called jangali, used to be levied before the jihād. Since jangali was unpopular and illegal according to the Shaikh, the term was expressly dropped in favour of the legal zakāt.* [50] *Zakāt in theory should be disbursed locally and immediately for the legally prescribed purposes, and none forwarded. Kharāj (land-tax), a term which became synonymous with tax generally, was not, it seems, paid in Sokoto.*[51] *All land had been declared waqf—land held for the benefit of the Community, of which only the usufruct can be transferred and, then, theoretically, only for a limited period. As the pres-*

ence of the Shaikh attracted a great number of immigrants, all of whom would be Muslims, the land round Sifawa and Sokoto was settled by people who were unwilling to pay any tax like kharāj. Thus when Amir al-mu'minin 'Umar tried to introduce a form of head tax, he was prevented on grounds of Islamic Law by public opinion.[52] "

Together, these examples show that Islamic law constrained both taxpayers and rulers: it defined legitimate fiscal obligations and limited the taxes rulers could impose. Outside the Caliphate, fiscal arrangements were considerably more heterogenous. Depending on the polity, revenue could take the form of tribute, tolls, market and trade levies, and labor obligations (Hopkins 1973). These arrangements could themselves rest on recognized political or customary authority. What distinguished Sokoto was the scale and territorial reach of a common Islamic legal framework that linked fiscal obligations to religious status.[3]

2.2 Colonial Taxation and Fiscal Authority

British conquest incorporated the former Caliphate into a common colonial state. Sokoto fell in 1903, and the territories that became Nigeria were ultimately amalgamated into the Colony and Protectorate of Nigeria in 1914.[4] Importantly for our analysis, the former Sokoto frontier did not become a colonial administrative boundary. The boundary between the Northern and Southern Protectorates instead emerged from the reorganization of pre-existing British jurisdictions rather than from the territorial limits of the Caliphate (Figure D1, Panel b).[5]

Colonial tax policy followed these new administrative boundaries. Tax systems initially differed across the Northern and Southern Protectorates and continued to evolve, but the major instruments of direct taxation became increasingly similar over the colonial period. Figure 2, constructed from the *Blue Books* and the *Reports from the Committee of Taxation*, documents this convergence. Some regional differences persisted, notably *jangali* in the Northern Protectorate and export taxation in the Southern Protectorate. Crucially, these differences followed colonial administrative boundaries rather than the former Sokoto frontier. Because the two sets of boundaries did not coincide, nearby locations on opposite

[3]The Caliphate's southern frontier illustrates this contrast. In areas such as Zuru and parts of the Jos Plateau, many communities remained outside effective Sokoto incorporation and retained political and fiscal institutions that were not organized around the Caliphate's Islamic legal order (Last 1967; Lovejoy 2016).

[4]Appendix A.4 provides a detailed chronology of British expansion and administrative reorganization in Nigeria.

[5]When the Royal Niger Company surrendered its charter in 1900, its southern territories were combined with the Niger Coast Protectorate to form Southern Nigeria, while the remainder became Northern Nigeria. The resulting division therefore followed the geography of earlier British administration rather than the former Sokoto frontier.

sides of the former Caliphate could face the same colonial tax regime despite their different pre-colonial fiscal histories.

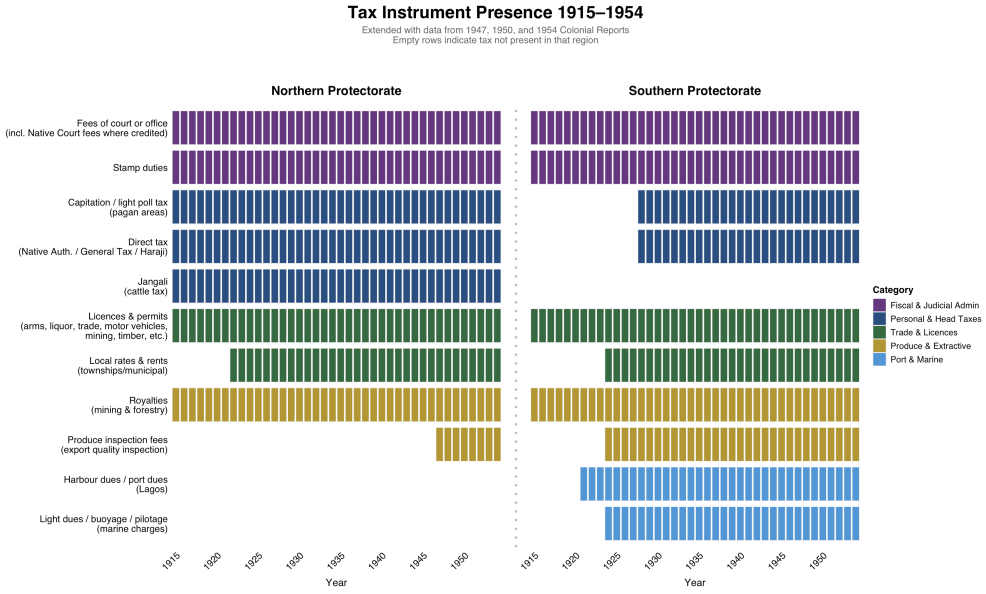


Figure 2: Tax Instruments in the Northern and Southern Protectorates

Notes: The figure compares tax instruments in the Northern and Southern Protectorates over the colonial period. Data: *Blue Books* and *Reports from the Committee of Taxation*.

In former Sokoto, colonial authorities continued to rely on emirs and subordinate local officials to administer and collect taxes (Kirk-Greene 1965; Mamdani 1996), but substantially altered the rules and authority underlying taxation. Official collection of *zakat* as a distinct religious obligation was curtailed as colonial authorities sought to bring all revenue collection under a standardized colonial fiscal system. Colonial taxation also moved away from religious status as a basis for determining tax liability. Under Sokoto rule, Muslims in the Caliphate's core territories had been exempt from *jizya* and *kharaj*, taxes analogous to poll and land taxes; colonial authorities imposed comparable taxes without preserving the same confessional distinctions. Colonial authorities also reintroduced *jangali* in the Northern Protectorate as part of a broader effort to simplify direct taxation and extend it to pastoral wealth.[6] The cattle tax provided a practical way to tax mobile livestock-owning populations. In former Sokoto, however, this restored a levy that Sokoto's founders had rejected as lacking a legitimate basis in Islamic law. Historical accounts indicate that these changes themselves generated resistance. In core Sokoto and Gwandu, the introduction of poll taxation and *jangali* provoked grievances among Muslims who had previously been exempt from these

[6]The reform formed part of Lugard's reorganization of direct taxation in Northern Nigeria, which sought to consolidate numerous existing levies into a more standardized fiscal system and to bring pastoral wealth within the tax base (Lugard 1907; Adebayo 1995).

levies (Lovejoy 1990; Jumare 1998). Colonial rule therefore retained much of the machinery of tax collection while changing which taxes were owed, the authority under which they were imposed, and whether they had a recognized basis in Islamic law.

3 Data and Measurement

3.1 The Sokoto Boundary and Predetermined Characteristics

We digitize the boundary of the Sokoto Caliphate from Lovejoy (2016) and georeference it using stable historical landmarks, including named settlements and rivers. Panel B of Figure 1 overlays the resulting boundary on a 1924 map of Nigeria.

The location of the historical boundary may itself reflect geographic or pre-existing economic conditions. Historical accounts emphasize military reach and ecological constraints, particularly along Sokoto's southern frontier (Kirk-Greene 1965; Last 1967; Lovejoy 2016). We therefore assemble a broad set of predetermined characteristics both to assess balance at the boundary and to use as controls in the empirical analysis. Geographic characteristics include elevation and ruggedness (GDEM v2.0); temperature, precipitation, length of the growing period, and suitability for cassava, maize, yam, cocoa, millet, sorghum, and palm oil (GAEZ v4.0); malaria ecology (Malaria Atlas Project); tsetse suitability (Alsan 2015); and distance to the coast and to the nearest river. We complement these measures with two pre-Sokoto historical characteristics: population density in 1800 (HYDE 3.1) and slave-trade intensity (Nunn 2008). Section 4 shows that these characteristics vary smoothly across the Sokoto boundary.

3.2 Colonial Tax Records

We assemble a new corpus of 43,648 tax-related colonial documents from the three zonal offices of the National Archives of Nigeria. The collection spans the country and much of the colonial period, providing unusually broad coverage of day-to-day fiscal administration and state–citizen interactions. The Kaduna archive contributes 23,032 documents from northern Nigeria; the Enugu and Ibadan archives contribute 20,616 documents from the South-East and South-West. Appendix Figure B1 shows the physical archive, while Figure B2 shows examples of documents.

The documents record the day-to-day operation of colonial fiscal administration, including correspondence, financial and tax records, reports, petitions, and complaints. Importantly, the corpus is not selected around known episodes of tax resistance, but instead captures

resistance as it appears within a much broader body of routine administrative material. An LLM classification of document type confirms this breadth: correspondence between authorities is the most common category (71%), followed by financial and tax records (32%) and taxation reports (20%), while complaints, requests, or petitions account for about 7% of documents.[7] This breadth allows us to measure tax contestation beyond the salient revolts, riots, and rebellions that dominate much of the historical literature (Burg 2004), recovering individual grievances and other less visible forms of resistance. This is particularly valuable in colonial Nigeria, where overt collective resistance could be met with violent repression, so the absence of open revolt may substantially understate opposition to state demands.[8]

3.2.1 Extracting Location and Date

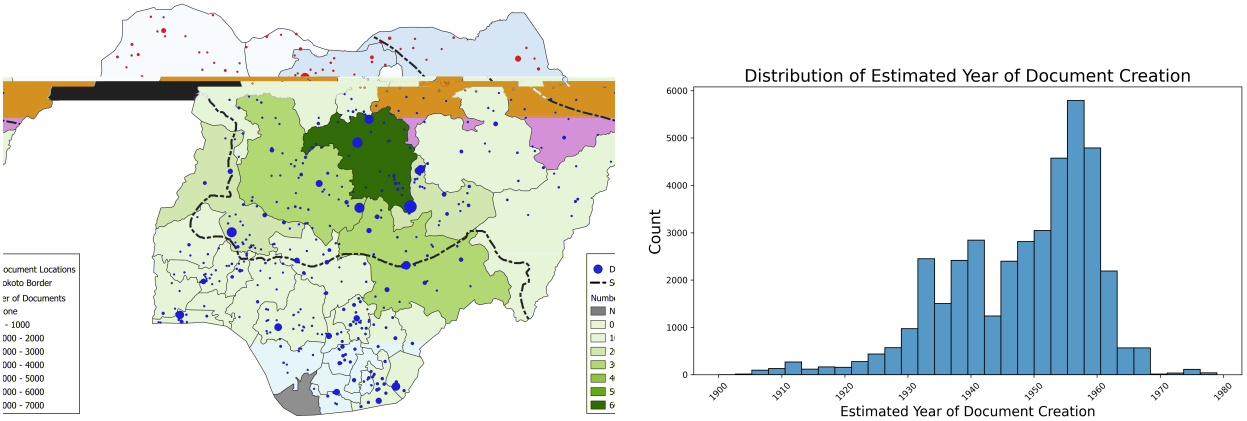
Each document is assigned a geographic location and year. We first use an LLM to identify the document's location of origin. We then standardize the resulting place names to account for spelling variants, OCR errors, abbreviations, and changes in naming conventions, and match them to modern localities using Google Maps.[9] Locations that cannot be matched remain unmatched. The resulting corpus contains documents from locations on both sides of the former Sokoto boundary. Panel A of Figure 3 shows their spatial coverage, while Figure B3 plots the document locations on a 1924 map of Nigeria.

We similarly recover each document's year from its contents rather than relying exclusively on archival folder labels. We first construct a master year estimate using a hierarchical rule-based procedure that reconciles dates extracted from the title and first page with the distribution of years mentioned throughout the document. We then use an LLM to identify the most precise creation date supported by the document. When the LLM-implied year is sufficiently close to the master estimate, we use it; otherwise, we retain the master estimate. Panel B of Figure 3 shows the temporal coverage of the resulting corpus.

[7] Documents may belong to multiple categories. The remaining categories are correspondence between authorities and civilians (10%), investigative or special reports (9%), historical records (8%), legal and policy documents (4%), cover pages (3%), minutes (2%), periodic administrative reports (1.7%), unidentifiable material (0.5%), and correspondence between civilians (0.04%).

[8] For example, British forces violently suppressed the 1906 Satiru uprising in the former Sokoto Caliphate; armed force was also used against the Ekumeku resistance in southern Nigeria and against protesters during the 1929 Women's War (Mohammed 2024; Igbafe 1971; Matera et al. 2012).

[9] Historical place names in Nigeria exhibit substantial continuity, although their standardized spellings have sometimes changed over time. To assess the feasibility of matching historical localities to contemporary geographic sources, we manually checked 100 legible settlements appearing across the 1924 map of Nigeria mentioned above, including both major towns and smaller settlements. All 100 could be identified in contemporary geographic sources after allowing for spelling variants. This exercise suggests that, conditional on recovering the historical place name, contemporary geographic sources provide a reliable basis for matching historical localities.



(a) Documents by Province (b) Estimated Dates of Creation

Figure 3: Document Location and Year

Notes: Panel (a) shows the number of documents by province, with the Sokoto Caliphate boundary overlaid. Panel (b) shows the distribution of estimated document creation dates.

3.2.2 Outcomes: Tax Sentiment and Complaints

We construct two complementary measures of tax resistance. Tax sentiment captures the broader climate surrounding taxation, including opposition expressed by taxpayers and problems recorded by colonial administrators. This approach builds on the broader economics literature using text as data (Gentzkow et al. 2019) and on evidence that sentiment extracted from economic text can recover meaningful variation closely related to human assessments (Shapiro et al. 2022). Formal tax complaints provide a narrower but more direct measure: written grievances submitted to an authority. Together, the two measures capture both diffuse opposition and explicit resistance directed toward the colonial state.

Tax Sentiment. We ask an LLM to classify the sentiment toward taxation expressed or reported in each document on an ordinal scale from 0 to 4, ranging from very negative to very positive; documents that cannot be reliably interpreted receive a separate score of -1. Very negative documents describe acute conflict or distress surrounding taxation, including riots, violent clashes, and punitive actions. Negative documents capture nonviolent opposition or problems, including petitions, tax evasion or refusal, allegations of corruption, and administrative disputes. Neutral documents consist largely of routine fiscal administration, while positive classifications capture successful collection, cooperation, praise, or other favorable assessments. Appendix B.4 reproduces the complete classification prompt.

Because individual LLM responses can vary, we classify each document multiple times and assign its modal classification, using the median to break ties. Overall, 64% of documents are classified as neutral and 34% as negative or very negative; only 2% are classified as positive

or very positive. Figure 4, Panels (a) and (b), shows the distribution of tax sentiment and the evolution of negative sentiment over the colonial period.

Formal Tax Complaints. We define a formal tax complaint as a written grievance about taxation submitted to an authority, including petitions, letters of protest, and appeals against tax assessments. We identify these complaints using a separate LLM classification of tax-related issues. A document qualifies only when it contains an explicit grievance submitted to or reported to an authority; discussion of taxation or an administrative problem alone is insufficient. Appendix B.5 reproduces the classification instructions.[10] Formal tax complaints appear in 15% of the corpus. Figure 4, Panels (c) and (d), shows their overall frequency and evolution over the colonial period.

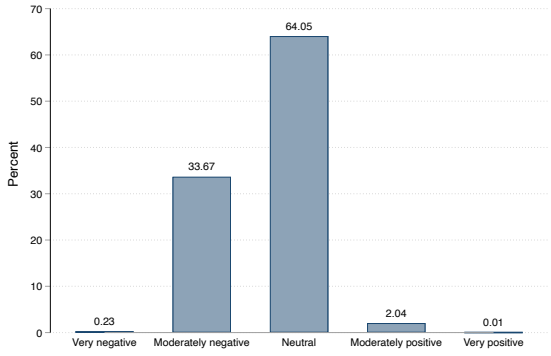
Both measures become more prevalent over the colonial period, particularly in its later decades. This pattern is consistent with historical accounts of expanding anticolonial and nationalist mobilization in Nigeria from the 1930s and especially after World War II (Falola and Heaton 2008). We do not interpret these aggregate trends causally, since the composition of surviving archival records may itself change over time. Their correspondence with the historical record nevertheless provides a useful check that the measures capture meaningful variation in the documents.

3.3 Validating the LLM Measures

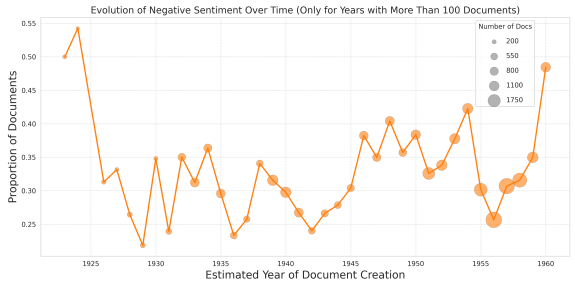
We validate our two primary outcomes against independent human readings of the same documents. The exercise serves two purposes. First, we ask whether tax sentiment and formal complaints can be measured consistently from the historical text by examining agreement across independent human readers. Second, we ask whether the LLM recovers these classifications as reliably as the human readers themselves. Both outcomes require substantive interpretation rather than simple keyword matching: measuring tax sentiment requires distinguishing routine fiscal administration from descriptions of problems or resistance, while identifying formal complaints requires determining whether a document contains an actual grievance submitted to an authority. Human validation is therefore particularly important because the classifications may themselves be ambiguous to human readers, while LLMs may also misinterpret complex historical text or assign classifications unsupported by the underlying document (Korinek 2023; Dell 2025).

Neither outcome has an independently observable ground truth. Following the Turing-test approach of Lagakos et al. (2025), five research assistants independently classify a random audit sample of 100 documents using the same substantive definitions as the LLM. We use

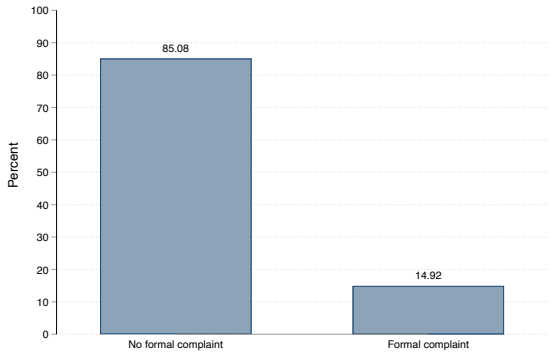
[10] The other categories described in Appendix B.5 become outcomes later in the paper when we distinguish the legitimacy mechanism from alternative explanations.



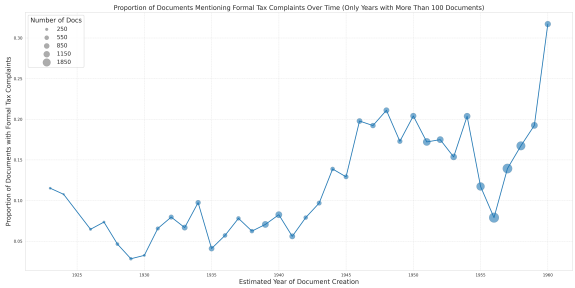
(a) Distribution of Tax Sentiment



(b) Negative Sentiment over Time



(c) Share of Formal Tax Complaints



(d) Formal Tax Complaints over Time

Figure 4: Tax Resistance in the Colonial Records

Notes: Panels (a) and (b) show the distribution of tax sentiment and the share of documents expressing negative tax sentiment by year. Panels (c) and (d) show the overall share of documents classified as formal tax complaints and the share containing a formal tax complaint by year.

the modal human classification for each document as a common benchmark in two ways. First, agreement between individual research assistants and the modal human classification measures how consistently each outcome can be recovered from the text. Second, agreement between the LLM and the same benchmark measures how reliably the LLM reproduces the human classifications. For tax sentiment, which is ordinal, performance is measured by the average absolute distance between each rater's classification and the modal human score. For formal complaints, which are binary, performance is measured by the share of classifications that agree with the modal human classification.

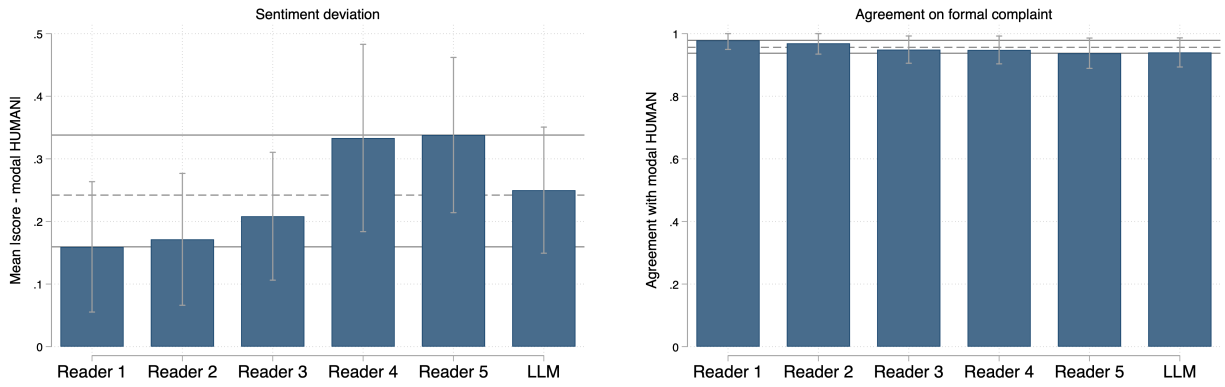
Figure 5 reports the results. We first establish the range of performance among human readers. For formal complaints, individual research assistants agree with the modal human classification on approximately 94–98% of documents. For tax sentiment, their average absolute deviation from the modal human score ranges from approximately 0.16 to 0.34 points on the 0–4 scale. The high degree of agreement indicates that independent human readers classify both outcomes consistently.

We then compare the LLM with the human readers using the same benchmark. Its average absolute deviation from the modal sentiment score is approximately 0.25, squarely within the range of the human readers. For formal complaints, the LLM agrees with the modal human classification on approximately 94% of documents, again within the human range. For neither outcome can we reject equality between the LLM's performance and that of the median human reader. The LLM thus passes our Turing-style validation test: its performance is statistically indistinguishable from that of a typical human reader.

This validation is consequential for the scale of our empirical analysis. Based on the time required to classify the audit sample, coding all 43,648 documents would require nearly three years of full-time work by a single research assistant.[11] The LLM allows us to apply the same classification procedure consistently across the full corpus while achieving performance comparable to independent human readers. This makes it possible to construct systematic measures of tax resistance from an archival record whose scale would otherwise make comprehensive human coding difficult.[12]

[11] At approximately 60 documents per working day, coding the full corpus would require about 728 working days. Many documents are only a few pages long, but each must be read and classified along multiple dimensions used in the analysis.

[12] We conduct analogous validation exercises for the secondary outcomes used later in the paper, including state coercion or punishment, tax evasion, corruption, and administrative failure, and reach the same two conclusions: human readers classify these outcomes consistently, and LLM performance falls within the human range (Appendix Figures C1–C5). We apply the same validation procedure to the direct measure of fiscal legitimacy discussed in Section 6.



(a) Tax Sentiment (b) Formal Tax Complaints

Figure 5: Human Agreement and LLM Performance

Notes: The figure benchmarks five independent human research assistants and the LLM against the modal human classification in a random audit sample of 100 documents. Panel (a) reports each rater's average absolute distance from the modal tax-sentiment score; lower values indicate greater agreement. Panel (b) reports agreement with the modal formal-complaint classification; higher values indicate greater agreement. Error bars report 95% confidence intervals.

4 Empirical Strategy

Our empirical strategy exploits the former boundary of the Sokoto Caliphate. British conquest eliminated this boundary as a fiscal jurisdiction, but it continued to separate locations with different pre-colonial fiscal histories. Inside the boundary, colonial taxation displaced the Islamic fiscal order described in Section 2; immediately outside it, colonial rule did not displace a comparable Sokoto fiscal order. We therefore compare tax resistance in nearby locations on opposite sides of the historical frontier.

Specifically, we estimate

$$y_{d,l} = \alpha + \beta \cdot SokotoCaliphate_{d,l} + f(location_l) + X_l'\gamma + \delta_{s(l)} + \zeta_{a(d)} + \epsilon_{d,l}, \tag{1}$$

where $y_{d,l}$ is the outcome for document d from location l, and $SokotoCaliphate_l$ indicates that location l lies within the former Caliphate. The function $f(distance_l)$ is linear in distance to the historical boundary, with separate slopes on either side. X_l contains the predetermined geographic and pre-Sokoto characteristics described in Section 3.1. Boundary-segment fixed effects $\delta_{s(l)}$ restrict comparisons to locations facing the same stretch of the frontier, while archive fixed effects $\zeta_{a(d)}$ absorb level differences across the Kaduna, Enugu, and Ibadan collections. The coefficient β measures the discontinuity in tax resistance at the former Sokoto boundary. Our baseline specification is local linear. We report estimates

at bandwidths of 50, 100, 150, and 200 kilometers and at the MSE-optimal bandwidth (Cattaneo et al. 2020). Standard errors are clustered by document location, the level at which treatment varies. Section 5 examines alternative spatial functional forms and approaches to inference.

Identification requires that, absent exposure to Sokoto's fiscal order, other determinants of colonial tax resistance would vary smoothly across the historical boundary. Two concerns are particularly important. First, Sokoto may have expanded into locations that already differed systematically from neighboring areas. Second, the historical frontier may subsequently have become aligned with differences in colonial policy or state capacity. We examine these threats in turn.

Predetermined Characteristics. The Sokoto boundary was not randomly assigned. Historical accounts emphasize military reach and ecological constraints on cavalry in shaping the limits of Caliphate expansion, particularly along its southern frontier (Kirk-Greene 1965; Last 1967; Lovejoy 2016).[13] We therefore begin by examining the characteristics most directly implicated by these accounts, including terrain, vegetation, agricultural ecology, and tsetse suitability. We then broaden the balance exercise to a larger set of predetermined characteristics that could have shaped settlement, economic activity, or Sokoto expansion: elevation and ruggedness; temperature, precipitation, growing conditions, and crop suitability; malaria ecology; distance to rivers and the coast; population density in 1800; and slave-trade exposure.

Appendix Table D1 applies the RD specification to each characteristic over bandwidths from 50 to 200 kilometers. The estimated discontinuities are small, show no systematic pattern, and none is statistically significant. We therefore find no evidence of discontinuities either in the characteristics most directly related to the historical determinants of Sokoto expansion or in a broader set of observable pre-Sokoto characteristics.

A related concern is that Sokoto expansion followed pre-existing ethnic or livelihood patterns. For example, the Caliphate may have disproportionately incorporated Fulani or other pastoral populations that were independently more resistant to taxation.[14] The *Ethnographic Atlas* of Murdock (1967) shows that the Sokoto boundary cuts through several ethnic homelands rather than coinciding with their borders (Appendix Figure D1, Panel a). The main estimates also remain similar with ethnic-territory fixed effects, so that identification comes from comparisons within the same ethnic homeland on opposite sides of the Sokoto boundary

[13] Sokoto's armies relied heavily on horses, making expansion more difficult where dense vegetation and difficult terrain impeded cavalry and where tsetse-borne disease limited horse survival.

[14] Usman dan Fodio and many leaders of the jihad were Fulani, and members of this leadership subsequently became prominent in the emirate structures established after the conquest of the Hausa states. The Caliphate itself, however, incorporated a much broader population and was not ethnically homogeneous (Last 1967).

(Appendix Table D4). This specification also absorbs fixed differences across ethnic groups in pre-colonial characteristics, including reliance on pastoral livelihoods.[15]

Colonial Fiscal Environment. A second concern is that the historical boundary may later have become aligned with differences in the colonial environment. If locations inside former Sokoto faced systematically different taxes, heavier fiscal demands, weaker administration, harsher enforcement, or lower public-goods provision, the discontinuity could reflect colonial policy or state capacity rather than the pre-colonial fiscal order it displaced.

The institutional setting helps separate these possibilities. As documented in Section 2.2, the former Sokoto frontier did not become a colonial administrative boundary, and direct-tax institutions became increasingly similar across Nigeria. Because colonial administrative boundaries cut across the historical frontier, substantial portions of the RD compare nearby locations governed within the same protectorate (Appendix Figure D1, Panel b). We exploit this directly by re-estimating the RD within the Northern Protectorate, where locations on either side of the Sokoto boundary were subject to the same colonial fiscal regime; the estimates remain similar to the baseline results (Appendix Table D5). The discontinuity also persists after 1927, when direct-tax instruments had largely converged across the former protectorates (Appendix Table G3).

Section 6.3 examines the remaining dimensions of the colonial environment directly, including fiscal burden, administrative capacity, enforcement, and public-goods provision. We find little evidence that any of these changes discontinuously at the former Sokoto boundary or accounts for the estimated resistance gap.

5 Results

5.1 Tax Resistance at the Sokoto Boundary

Table 1 reports the main estimates of equation (1). Colonial tax resistance rises sharply at the former Sokoto boundary. At the MSE-optimal bandwidth, documents just inside the former Caliphate are 9.3 percentage points more likely to express negative tax sentiment, relative to a mean of 24 percent just outside the boundary. This represents an increase of roughly 40 percent. The discontinuity in formal complaints is larger still: documents inside former Sokoto are 18.2 percentage points more likely to contain a written tax grievance submitted to an authority, relative to an outside-Sokoto mean of 7 percent. Crossing the historical boundary therefore more than triples the probability that a document contains a

[15] Ethnic groups differed substantially in their reliance on pastoralism, and Fulani populations that became prominent in the emirate structures were historically more pastoral.

formal tax complaint.

The result is not specific to the MSE-optimal bandwidth. Across fixed bandwidths from 50 to 200 kilometers, the estimated discontinuity ranges from 14.3 to 19.4 percentage points for negative sentiment and from 8.1 to 19.9 percentage points for formal complaints; every estimate is positive and statistically significant. Figure 6 shows the corresponding raw spatial pattern. Both outcomes rise on the Sokoto side along substantial portions of the historical frontier rather than around a single cluster of locations. The discontinuity is also present throughout the colonial period and, if anything, grows in later decades.[16]

Table 1: More Negative Tax Sentiment and More Tax Complaints in Sokoto

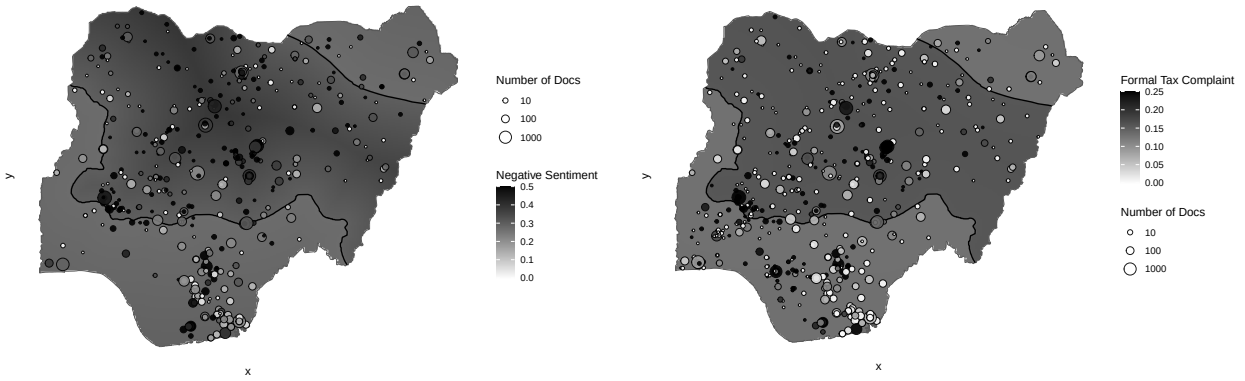
	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.093**	0.165***	0.194***	0.157***	0.143***
	(0.043)	(0.043)	(0.040)	(0.034)	(0.034)
Observations	7739	6016	9278	16810	24609
Control Mean	0.240	0.246	0.256	0.313	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.182***	0.190***	0.199***	0.097***	0.081***
	(0.018)	(0.013)	(0.025)	(0.023)	(0.028)
Observations	6121	6016	9278	16810	24609
Control Mean	0.070	0.071	0.081	0.105	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1). The outcome is an indicator for negative tax sentiment (Panel A) and for a formal tax complaint (Panel B). The unit of observation is a document. Column (1) uses the MSE-optimal bandwidth (Cattaneo et al. 2020); columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km around the Sokoto Caliphate boundary. All specifications include a local linear function of distance to the boundary estimated separately on each side, geographic controls, archive fixed effects, and border-segment fixed effects. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[16] We do not place strong interpretation on this temporal pattern because both the colonial environment and the composition of the surviving archive changed over time.



(a) Negative Tax Sentiment (b) Formal Tax Complaints

Figure 6: Discontinuities at the Sokoto Boundary

Notes: Each panel plots the raw outcome by location on either side of the Sokoto Caliphate boundary: negative tax sentiment in panel (a) and formal tax complaints in panel (b).

We also compare the archival measures with more visible forms of collective resistance. Appendix H assembles a separate dataset of 100 tax protests, riots, and revolts from the colonial documents and historical sources. These events are strongly positively associated with both negative tax sentiment and formal complaints (Appendix Table H1), providing an independent check that the archival measures capture meaningful opposition to taxation. Yet protests and riots themselves do not exhibit a precisely estimated discontinuity at the Sokoto boundary (Appendix Table H2). This contrast highlights an advantage of the archival measures: protests, riots, and revolts are rare, high-threshold manifestations of resistance, whereas negative sentiment and formal complaints capture grievances that need not escalate into collective action. In colonial Nigeria, where overt resistance could carry severe costs and was sometimes violently repressed, the absence of revolt need not imply the absence of resistance.[17]

5.2 Robustness

We assess the robustness of the spatial discontinuity to three broad concerns: the composition and spatial coverage of the archival sample, uncertainty over the historical boundary, and the econometric implementation of the spatial RD. Appendix D reports the corresponding results.

[17]The Satiru revolt of 1906 provides a stark example: British forces violently suppressed the uprising, killing or imprisoning many participants and leaders and sharply curtailing subsequent revolutionary resistance (Mohammed 2024).

Archival composition and coverage. A first concern is that the discontinuity reflects differences in the composition of the surviving archival record rather than differences in tax resistance. The three archives might preserve different types of documents, and the composition of their holdings may also vary over time. Our baseline archive fixed effects absorb average differences across archives. We further add document-type fixed effects, restricting comparisons to similar types of records, and archive-by-decade and archive-by-year fixed effects, restricting comparisons to documents from the same archive and period. The discontinuities remain positive and statistically significant across these specifications, although the estimates are attenuated in the most restrictive comparisons (Appendix Tables D6, D7, and D8). The results therefore do not appear to be driven by systematic differences in the types or timing of records represented in the archive.

A second concern is uneven archival coverage across locations. Because the unit of observation is a document, locations with more surviving records receive greater weight in the baseline estimates. If archival intensity is related to underlying tax resistance, this unequal representation could affect the estimated discontinuity. We first find no evidence that the number of surviving documents changes at the Sokoto boundary (Appendix Table D9). We then reweight observations by the inverse number of documents from each location, giving each location equal weight; the resulting estimates remain similar to the baseline (Appendix Table D10). Archival coverage may also reflect underlying population size. Yet division-level population from the 1952 Census varies smoothly at the Sokoto boundary (Table D2),[18] and reweighting the main specifications using these population figures again produces similar estimates (Appendix Table D11). The results therefore do not appear to be driven by differential archival coverage or by disproportionate weight on locations that contribute more surviving records.

Boundary measurement and historical ambiguity. A second concern is uncertainty in the location of the historical frontier. The Sokoto boundary is digitized from a nineteenth-century historical map and, although numerous named settlements and stable geographic features permit close georeferencing, its location cannot be measured perfectly. Following Michalopoulos and Papaioannou (2013), we address this uncertainty by excluding observations whose treatment assignment is most sensitive to the historical reconstruction. We therefore estimate a donut RD that drops locations within 10 kilometers of the boundary. The resulting estimates remain similar to the baseline (Appendix Table D12).

We also examine whether the results depend on the treatment of historically ambiguous

[18] We use the 1952–53 Census because it provides population counts at the division level throughout our study area and was the first comprehensive enumeration of Nigeria's population. Earlier censuses relied substantially on administrative estimates rather than complete enumeration.

regions. Kanem–Bornu was a neighboring Islamic polity with its own Islamic legal and fiscal institutions, making its classification less clear-cut than that of most areas outside Sokoto.[19] We therefore test robustness to excluding observations in Kanem–Bornu and to classifying Kanem–Bornu as treated; neither exercise materially changes the results (Appendix Table D13). We conduct analogous exercises for the Jos Plateau, whose incorporation into the Caliphate is disputed in the historical literature.[20] Excluding the Plateau leaves the main conclusions unchanged, while redrawing the boundary to place it outside Sokoto yields estimates of similar sign and magnitude, though with less precision in some specifications (Appendix Table D14). The main discontinuity is therefore robust to alternative classifications of the historically ambiguous portions of the frontier.

Specification and inference. Finally, we examine whether the results depend on the spatial specification or the assumptions used for inference. Replacing the baseline local linear function of distance to the boundary with a quadratic yields similar estimates. We also dispense with distance to the boundary altogether and instead model spatial trends linearly in latitude and longitude; the estimated discontinuities remain similar (Appendix Tables D15 and D16), and Appendix Figure D2 shows the corresponding spatial fit. We then allow for substantially broader dependence in the regression residuals. Clustering standard errors at the province rather than document-location level leaves both discontinuities precisely estimated (Appendix Table D17). Conley standard errors allowing for spatial correlation within 25 kilometers widen the confidence intervals, as expected, but leave the qualitative conclusion unchanged: the sentiment and complaint discontinuity remains significant (Appendix Table D18). Thus, neither the estimated discontinuities nor the statistical conclusions depend materially on the baseline spatial functional form or error structure.

5.3 Within-Caliphate Variation

We complement the boundary design with variation entirely within former Sokoto. This exercise has two advantages. First, it does not rely on the location of the historical frontier or on comparisons with areas whose pre-colonial fiscal institutions may have differed in ways that are difficult to observe fully. Second, it allows us to ask whether resistance is greater specifically where Sokoto's political and administrative institutions were more deeply rooted,

[19] Kanem–Bornu was a long-standing Islamic polity with established Islamic legal and fiscal institutions (Lavers 1971). Unlike Sokoto, however, its fiscal system combined Islamic taxes with other levies inherited from the pre-existing political order (Cohen 1967). Bornu therefore represents an ambiguous case for our treatment definition: it shared some of Sokoto's Islamic fiscal institutions without adopting the same reformist fiscal order.

[20] Communities on the Jos Plateau repeatedly resisted nineteenth-century incursions associated with the Caliphate and retained substantial political independence until British conquest (Tambo 1978).

rather than where Sokoto's historical presence was stronger more generally.

We construct two measures of institutional exposure from Last (1967). The first captures proximity to the formative process of Sokoto state-building. Between 1804 and 1817, Usman dan Fodio moved through a sequence of locations during the jihad and early consolidation of the Caliphate. We digitize these locations from the map and accompanying historical account in Last (1967) (Figure E1) and connect them using an elevation-based least-cost path, producing the consolidation corridor shown in Appendix Figure E2. Proximity to this corridor captures exposure to locations closely associated with the formation of the Sokoto political order.[21]

The second measure captures proximity to the Caliphate's subsequent administrative structure. Using another map in Last (1967), we digitize major administrative locations, including emirate centers and territories administered under the Vizier (Appendix Figure E3). These locations were principal nodes through which the Caliphate's political, judicial, and fiscal institutions were exercised.[22] The two measures thus capture exposure to the Caliphate's formation and to the institutions through which it was subsequently governed.

Table 2 reports the results. All specifications are restricted to locations within the former Caliphate and include the geographic and pre-Sokoto controls used in the main analysis together with archive fixed effects. In Panel A, greater proximity to the reconstructed consolidation corridor predicts significantly more negative tax sentiment, while the estimate for formal complaints is positive but less precise. Panel B provides a discrete alternative to the reconstructed corridor: locations that Last (1967) records Usman dan Fodio as having visited are matched to geographically and historically comparable locations elsewhere within the Caliphate. Visited locations exhibit significantly more formal complaints and more negative tax sentiment at the 10 percent level. Panel C turns to the Caliphate's subsequent administrative structure, where greater proximity to major administrative centers predicts significantly more negative tax sentiment and more formal complaints. Across the three specifications, resistance is therefore greater in locations more closely connected to both Sokoto's formation and its subsequent administration.

One concern is that these gradients could simply capture proximity to places where Sokoto power or historical activity was more concentrated. We examine this possibility using the royal ribats mapped by Last (1967). Established between 1839 and 1845, these fortified set-

[21] The locations include sites from which Usman dan Fodio and his followers organized military campaigns, established political authority, and consolidated the new Islamic order before the Caliphate's emirate structure was fully institutionalized (Last 1967).

[22] After consolidation, the Caliphate was governed through a decentralized emirate system. Emirs exercised political, judicial, and fiscal authority within their territories while recognizing the authority of Sokoto, and the Vizier played a central role in coordinating administration and governing territories under more direct Sokoto authority (Last 1967).

Table 2: Within-Caliphate Exposure to Sokoto State-Building and Tax Resistance

	Neg. Sentiment	Formal Tax Complaint
	(1)	(2)
Panel A: Log Proximity to Shaikh's Path (OLS, within 100km)		
Log proximity	0.024**	0.033
	(0.024)	(0.022)
Obs.	6,233	6,233
Control mean	0.358	0.131
Geo controls	Yes	Yes
Pre-Sokoto controls	Yes	Yes
Archive FE	Yes	Yes
Panel B: Matching on Shaikh Visit Locations (5km buffer, 200km range)		
Shaikh visit	0.053*	0.076***
	(0.036)	(0.018)
Obs.	1,913	1,913
Panel C: Log Proximity to Sokoto Administrative Centres (within 100km)		
Log proximity	0.054**	0.037**
	(0.023)	(0.018)
Obs.	13,530	13,530
Control mean	0.333	0.151
Geo controls	Yes	Yes
Pre-Sokoto controls	Yes	Yes
Archive FE	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. SE clustered by city. Panel B uses location-level nearest-neighbour matching with geo controls and Pre-Sokoto controls as matching covariates; document-level OLS includes matched-pair FEs and geo + Pre-Sokoto controls.

Notes: The table relates colonial-era tax resistance to exposure to early Sokoto state-building within the former Caliphate. The unit of observation is a document; outcomes are indicators for negative tax sentiment, column (1), and a formal tax complaint, column (2). Panel A is a document-level OLS of the outcome on log proximity to Usman dan Fodio's 1804–1817 path, among documents within 100 km of the path. Panel B uses location-level nearest-neighbor matching of locations the Shaikh visited (5 km buffer) to comparable locations within a 200 km range, with geographic and pre-Sokoto controls as matching covariates; the document-level OLS includes matched-pair fixed effects. Panel C is a document-level OLS on log proximity to the Caliphate's administrative centers, from Last (1967), within 100 km. Standard errors, clustered by city, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

tlements primarily capture Sokoto's military and defensive presence (Appendix Figure E4). If resistance reflected Sokoto presence more generally, proximity to the ribats should also predict greater resistance. It does not: proximity to the royal ribats is not positively associated with either negative tax sentiment or formal complaints and, if anything, the estimated relationships are negative (Appendix Table E1).

The within-Caliphate evidence therefore points to institutional rather than generic historical exposure. Resistance is greater near locations associated with Sokoto's state formation and administrative structure, but not near sites associated primarily with its military presence.

6 Interpreting the Sokoto Discontinuity

The boundary and within-Caliphate results show that colonial tax resistance was greatest where Sokoto's institutions had been most deeply rooted. We now ask what accounts for this pattern. Our interpretation is that colonial rule disrupted an established source of fiscal legitimacy. Although the colonial regime preserved important elements of Sokoto's machinery of tax collection, it changed both the basis of fiscal authority and the obligations it imposed. Official collection of *zakat* as a religious duty was curtailed, religious distinctions governing liability for other taxes were weakened or removed, and levies such as *jangali*, which Sokoto reformers had rejected on Islamic grounds, were reintroduced.

This interpretation predicts not merely more resistance, but a distinctive pattern of resistance. The increase should be disproportionately associated with Muslims, whose fiscal obligations changed most sharply; it should disproportionately concern taxes whose legitimacy had been contested under the Sokoto order; and resistance should more often challenge the authority or normative basis of taxation itself. We test these predictions in turn before assessing alternative explanations based on population composition, colonial fiscal policy, and state performance.

6.1 Who Resists?

Our first test asks who is responsible for the formal complaints in the archival record. The colonial transformation of taxation was not religiously neutral. Under Sokoto rule, Muslims in the Caliphate's core territories were principally subject to religiously sanctioned obligations such as *zakat*, while some taxes subsequently imposed by the colonial state had previously been restricted to other populations or rejected altogether. If disruption of this fiscal order contributed to resistance, the increase in formal complaints inside former Sokoto

should therefore be disproportionately associated with Muslims.

For each formal complaint, we identify the named complainant and classify the name as Muslim, Christian, or unclear. The procedure has two stages. An LLM first identifies the individual actually making the complaint, excluding officials and individuals merely mentioned in the document. We then match the extracted names to curated repositories of Muslim and Christian names; unmatched names are classified by the LLM using linguistic and cultural naming conventions. Appendix B.7 describes the procedure in detail.

Formal complaints involving Muslim complainants rise sharply at the Sokoto boundary. Figure 7 shows the composition of complainants on either side, while Table 3 reports the corresponding spatial discontinuities. Panel A considers the unconditional probability that a document contains a formal complaint made by a Muslim. Such complaints account for only about 0.5 percent of documents just outside the former Caliphate. At the MSE-optimal bandwidth, crossing into former Sokoto increases their incidence by 1.9 percentage points, with positive estimates at every reported bandwidth. Panel B asks whether conditional on observing a formal complaint, the complainant become more likely to be Muslim? The estimated discontinuity is positive at every reported bandwidth, although it is imprecisely estimated at the MSE-optimal bandwidth. At 50 kilometers, for example, the Muslim share among complainants rises by 23 percentage points relative to a control mean of 11 percent. The increase in formal resistance inside former Sokoto is therefore disproportionately associated with Muslim complainants.

This pattern is unlikely to be mechanically explained by a larger Muslim population on the Sokoto side since Muslim population share varies smoothly at the historical boundary in the 1952 Census (Appendix Table D2).[23] We then use the same Census data to test a complementary implication of the fiscal-legitimacy interpretation. If colonial taxation generated greater resistance by displacing an Islamic fiscal order, Muslim population share should be more strongly associated with resistance within former Sokoto than outside it. Consistent with this prediction, formal complaints increase with division-level Muslim population share within former Sokoto, whereas the corresponding relationship outside the Caliphate is negative (Appendix Table F1). This pattern again points to greater resistance where the population whose fiscal position was most affected by the change in fiscal order was more prevalent.

[23]Because the Census postdates the fall of the Caliphate by nearly half a century, we do not interpret this result as evidence that religious composition was balanced before colonization. Rather, it shows that the discontinuity in Muslim complainants is not mirrored by a corresponding difference in Muslim population shares later in the colonial period.

Table 3: Who Complains? Muslim Complainants at the Sokoto Boundary

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Muslim Formal Complaint					
In Sokoto	0.019**	0.012	0.020**	0.040**	0.053*
	(0.008)	(0.008)	(0.008)	(0.020)	(0.027)
Observations	7141	5831	8263	15709	23392
Control Mean	0.005	0.006	0.006	0.013	0.015
Panel B: Muslim, Conditional on Complaint					
In Sokoto	0.111	0.234***	0.117*	0.135*	0.232**
	(0.069)	(0.046)	(0.068)	(0.080)	(0.098)
Observations	1042	976	1068	1921	3552
Control Mean	0.090	0.111	0.103	0.138	0.096
Geographic Controls	No	No	No	No	No
Archive FEs	No	No	No	No	No
Border Segment FEs	No	No	No	No	No
Pre-Sokoto Controls	No	No	No	No	No

Standard errors in parentheses. Sample excludes 865 Bornu-area observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates at the Sokoto Caliphate boundary. In Panel A, the outcome is an indicator for a formal tax complaint with a Muslim complainant; in Panel B, the outcome is an indicator for a Muslim complainant conditional on the document being a formal tax complaint. The unit of observation is a document. Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. As indicated in the table, these specifications include no controls or fixed effects. The sample excludes 865 Bornu-area observations. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

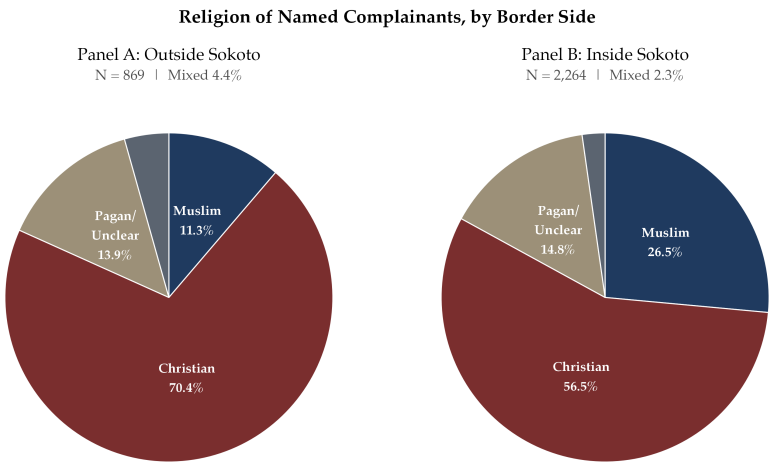


Figure 7: Complainant Religion on Either Side of the Sokoto Boundary

Notes: The figure shows the composition of complainant religion among formal tax complaints, inside and outside the former Sokoto Caliphate. Complainant names are extracted by an LLM and classified as Muslim, Christian, or unclear using curated name repositories, with unmatched names classified by the LLM.

6.2 Which Taxes Are Contested?

Our second test examines whether resistance disproportionately concerns taxes whose legitimacy had been contested under Sokoto rule. *Jangali*, a tax on cattle, provides the sharpest case. As discussed in Section 2.1, *jangali* had been rejected by Usman dan Fodio's reform movement as lacking a legitimate basis in Islamic law and replaced by the legally prescribed *zakat*. Colonial authorities later reintroduced *jangali* in the Northern Protectorate as part of an effort to extend direct taxation to pastoral wealth, making it possible to tax mobile livestock-owning populations. In former Sokoto, colonial rule therefore reinstated a levy that the Sokoto fiscal order had deemed illegitimate. Complaints against *jangali* provide a particularly direct test of whether colonial resistance concentrated on a tax rejected under the displaced fiscal order.

We use an LLM to identify formal complaints centered specifically on *jangali*. Appendix B.8 reproduces the classification prompt.[24] Table 4 shows a clear discontinuity in *jangali*-related grievances. Such complaints are essentially absent immediately outside the former Caliphate but become more common on the Sokoto side. In Panel A, the probability that a document contains a formal complaint centered on *jangali* rises at the boundary, with

[24] A similar historical logic applies to colonial poll and land taxes. Under Sokoto rule, Muslims in the Caliphate's core territories had generally not been subject to poll and land taxes analogous to *jizya* and *kharaj*; under colonial rule, comparable taxes were imposed without preserving the same religious distinctions. In practice, however, colonial records use several overlapping terms for poll and land taxation, making these levies substantially harder to identify consistently. We therefore focus on *jangali*, for which both the historical prediction and the textual classification are unusually sharp.

positive and statistically significant estimates at most reported bandwidths. Panel B shows a similar pattern when we condition on observing a formal complaint: the share of complaints centered on *jangali* generally increases on the Sokoto side.[25] Taken together, these results show that part of the excess formal resistance inside former Sokoto was directed specifically at a levy that its displaced fiscal order had rejected as illegitimate.

Table 4: Which Taxes Are Contested? Jangali Complaints at the Sokoto Boundary

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Jangali-Related Formal Complaint					
In Sokoto	0.002***	0.002***	0.001*	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Observations	6262	5748	8174	14879	21933
Control Mean	0.000	0.000	0.001	0.002	0.001
Panel B: Jangali-Related, Conditional on Complaint					
In Sokoto	0.084**	0.029***	-0.073***	0.114***	0.035
	(0.037)	(0.017)	(0.014)	(0.022)	(0.027)
Observations	889	889	975	1088	2087
Control Mean	0.000	0.000	0.015	0.057	0.056
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Pre-Sokoto Controls	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. Sample excludes 865 Bornu-area observations. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates at the Sokoto Caliphate boundary. In Panel A, the outcome is an indicator for a formal tax complaint specifically centered on *jangali*, the cattle tax; in Panel B, the outcome is an indicator for a *jangali*-related complaint conditional on the document being a formal tax complaint. The unit of observation is a document. Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. All specifications include geographic controls, pre-Sokoto controls, archive fixed effects, and border-segment fixed effects. The sample excludes 865 Bornu-area observations. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The preceding tests infer the legitimacy mechanism from two features of resistance—who complains and which taxes are contested—without requiring us to identify legitimacy claims directly. As complementary evidence, we construct a text-based measure that identifies whether a document challenges the authority on which a tax rests.[26] Appendix Figure C5

[25] Because *jangali* complaints are extremely rare outside former Sokoto, these conditional estimates are less stable across bandwidths and are negative at 100 kilometers.

[26] The prompt, reproduced in Appendix B.9, defines tax legitimacy in terms of rightful authority, continuity, and moral justification, and explicitly distinguishes it from the level of taxation, coercion or enforcement, administrative capacity, and general positive or negative sentiment.

validates this measure using the same comparison with independent human readers as in Section 3.3. The contested-legitimacy measure exhibits a sharp discontinuity at the Sokoto boundary: documents from former Sokoto are 8.3 percentage points more likely to contest the legitimacy of taxation at the MSE-optimal bandwidth, and the estimates remain positive and statistically significant across the reported bandwidths. They are also larger when we restrict attention to classifications in which the LLM expresses greater confidence. Taken together, the three exercises show that both the composition and content of resistance align with the fiscal changes produced by the displacement of Sokoto's order.

6.3 Assessing Alternative Explanations

The composition and content of resistance already narrow the set of plausible alternatives. Differences in colonial policy or state performance could generate greater dissatisfaction on one side of the boundary, but they do not naturally predict that the additional complaints should come disproportionately from Muslims, concern a tax rejected under Sokoto's fiscal order, and more often challenge the authority underlying taxation. We nevertheless examine the main competing explanations directly: differences in population composition, colonial tax instruments and burdens, the administration and enforcement of taxation, and public-goods provision. Appendix G reports the corresponding data, maps, and estimates.

Population Composition. A first possibility is that the Sokoto boundary separates populations that differed in ways that independently generated more tax resistance. As a large and institutionally developed pre-colonial polity, Sokoto may have contained populations that were more educated, wealthier, or more commercially engaged, and therefore better able to submit written grievances or more exposed to taxation. Differences in religious composition could likewise mechanically generate more complaints by Muslims on the Sokoto side. We examine these possibilities using the 1952 Census, which reports Muslim population share, Latin-script literacy, and occupational structure. None exhibits a discontinuity at the historical boundary (Appendix Table D2).[27] The main estimates are also robust to conditioning directly on these characteristics: controlling for literacy leaves both resistance discontinuities precisely estimated, while controlling for occupational shares produces the same qualitative pattern (Appendix Tables G1 and G2). Observable differences in population composition therefore do not appear to account for the resistance gap.

[27]The Census was conducted nearly half a century after the fall of the Caliphate, and no comparably disaggregated population data are available for the pre-colonial period. We therefore do not interpret these balance tests as evidence that population composition was identical before conquest; rather, they show that the resistance discontinuity is not accompanied by corresponding differences in these observable characteristics later in the colonial period.

Tax Instruments and Burden. A second possibility is that former Sokoto faced different or heavier colonial taxation. We reconstruct colonial tax instruments from the *Blue Books* and the *Reports from the Committee of Taxation*. Although some regional differences remained—most notably *jangali* in the North and export taxation in the South—the broad structure of direct taxation became increasingly similar across Nigeria (Figure 2). The Sokoto discontinuity persists after 1927, when direct-tax instruments had largely converged across the former protectorates (Appendix Table G3), and when we restrict the sample to the Northern Protectorate, comparing locations subject to the same colonial tax regime on either side of the historical boundary (Appendix Table D5). Differences in statutory tax instruments therefore cannot readily explain the resistance gap.

Greater resistance might instead reflect a heavier tax burden. Available aggregate fiscal accounts provide no evidence in this direction. Revenue data from the same colonial sources show that the bulk of colonial revenue was raised in southern Nigeria rather than in areas formerly governed by Sokoto (Appendix Figure G5). These accounts do not measure the effective burden faced by each locality and therefore cannot rule out local differences in fiscal incidence. They nevertheless provide no indication that the Sokoto side was systematically more heavily taxed.

Tax Administration and Enforcement. Greater resistance could also arise from how colonial taxation was administered or enforced. A weaker or more abusive bureaucracy could generate more grievances even under identical statutory taxes, while harsher enforcement could induce greater opposition to collection.

Documents from former Sokoto are, if anything, less likely to describe poor tax administration. The evidence on corruption is mixed, with no significant discontinuity at the MSE-optimal bandwidth and positive estimates only at wider bandwidths (Appendix Table G4). An independent measure of administrative presence points in the same direction: civil servants per 1,000 population in the 1931 Census vary smoothly at the Sokoto boundary, and controlling for this measure leaves the main estimates intact (Appendix Tables D2 and G5). References to state coercion and punishment likewise show no consistent discontinuity across bandwidths, while references to tax enforcement show no significant discontinuity at any reported bandwidth (Appendix Table G6). The resistance gap therefore does not appear to reflect a discontinuously weaker or more coercive colonial tax administration.

Public Goods Provision. Finally, greater resistance could reflect receiving fewer services in return for taxation. We assemble independent measures of colonial public investment spanning transportation, health and education, administrative infrastructure, utilities, and communications. Appendix Figures G7–G9 describe and map these measures. We find no

systematic discontinuity in public-goods provision at the historical boundary, and conditioning on these measures leaves the estimated resistance gap intact (Appendix Tables D2 and G7).

Taken together, these exercises provide little evidence that the Sokoto discontinuity reflects differences in observable population composition, colonial tax instruments or burdens, the administration or enforcement of taxation, or public-goods provision. Nor do these alternatives naturally account for the distinctive composition and content of resistance documented above. The evidence therefore favors an interpretation in which colonial resistance was shaped by the displacement of Sokoto's fiscal order rather than simply by differential exposure to colonial taxation or state performance.

7 Persistence

Did the colonial disruption of fiscal legitimacy leave a persistent legacy? Two channels could plausibly transmit the historical difference into the present. First, post-independence taxation largely inherited colonial fiscal institutions rather than restoring the pre-colonial Islamic fiscal order, so the state's claim to tax in former Sokoto was not re-embedded in its earlier religious foundation. Second, beliefs about the rightful scope of state authority may persist across generations even as institutions change. We do not attempt to distinguish between these channels. Instead, we ask whether the historical boundary continues to predict differences in contemporary attitudes toward taxation.

We use Afrobarometer data for Nigeria, rounds 5–9, and estimate the spatial boundary specification on individual survey responses. We distinguish two measures of fiscal legitimacy from two downstream measures of tax morale. Fiscal legitimacy is captured by agreement that tax authorities have the right to make people pay taxes and by trust in the tax department. Tax morale is captured by whether respondents believe it is wrong not to pay taxes owed on their income and whether they report having refused, or being willing to refuse, to pay a tax in the past year.

Table 5 shows the strongest evidence for persistence in the outcome most directly tied to the paper's historical mechanism. Respondents on the former Sokoto side of the boundary are substantially less likely to agree that tax authorities have the right to make people pay taxes. The estimate is negative and statistically significant at the MSE-optimal bandwidth and at every fixed bandwidth, ranging from 0.23 to 0.27 standard units across specifications. The second legitimacy measure, trust in the tax department, is also lower at every bandwidth, but the estimates are not statistically significant. The two tax-morale outcomes are likewise directionally consistent with lower acceptance of taxation in former Sokoto but are less

precisely estimated: agreement that it is wrong not to pay taxes is lower throughout and marginally significant at the MSE-optimal and 100-kilometer bandwidths, while self-reported tax refusal is higher in the direction implied by lower tax morale but statistically insignificant.

Taken together, the Afrobarometer evidence is consistent with persistence in fiscal legitimacy, but the result is concentrated in the measure that most directly captures the state's perceived right to tax. The broader measures of trust and tax morale move in the same direction but are estimated less precisely.

Table 5: Persistent Gap in Fiscal Legitimacy: Afrobarometer Evidence

	(1) MSE-Optimal	(2) 100km	(3) 200km	(4) 300km
Panel A: Tax Authorities Always Have the Right to Make People Pay				
In Sokoto	-0.264**	-0.271**	-0.240***	-0.232***
	(0.133)	(0.129)	(0.084)	(0.071)
Observations	3426	3788	9001	11965
Panel B: Trust in the Tax Department				
In Sokoto	-0.246	-0.143	-0.116	-0.100
	(0.174)	(0.144)	(0.097)	(0.082)
Observations	1039	1584	3745	4920
Panel C: Wrong for People Not to Pay Taxes Owed on their Income				
In Sokoto	-0.101*	-0.107*	-0.101	-0.069
	(0.058)	(0.059)	(0.065)	(0.061)
Observations	1055	1152	2789	3703
Panel D: Has/Would Have Refused to Pay a Tax in the Past Year				
In Sokoto	-0.104	-0.122	-0.075	-0.028
	(0.172)	(0.154)	(0.116)	(0.097)
Observations	888	1570	3820	5057
Afrobarometer Round FE	Yes	Yes	Yes	Yes
Border Segment FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes
Historical Controls	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates at the Sokoto Caliphate boundary using Afrobarometer data for Nigeria, rounds 5–9. The unit of observation is a survey respondent. Outcomes are agreement that tax authorities always have the right to make people pay taxes (Panel A), trust in the tax department (Panel B), agreement that it is wrong not to pay taxes owed on one's income (Panel C), and having refused, or saying one would have refused, to pay a tax in the past year (Panel D). Column (1) uses the MSE-optimal bandwidth; columns (2)–(4) use fixed bandwidths of 100, 200, and 300 km. All specifications include Afrobarometer round fixed effects, border-segment fixed effects, individual controls, geographic controls, and historical controls. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

8 Conclusion

Using nearly 50,000 colonial administrative documents, this paper shows that resistance to taxation was substantially greater where colonial rule displaced an established source of fiscal legitimacy. In the Sokoto Caliphate, taxation had been embedded in an Islamic fiscal order that defined both taxpayers' obligations and rulers' authority to tax. British rule preserved important elements of the machinery of collection but changed the institutional basis of taxation: official *zakat* collection was curtailed, religious distinctions governing tax liability were weakened or removed, and levies such as *jangali* that Sokoto reformers had rejected as illegitimate were reintroduced. At the former Caliphate boundary, documents from the Sokoto side express more negative tax sentiment and contain substantially more formal tax complaints than documents from nearby locations outside it.

Several pieces of evidence tie this discontinuity to the fiscal order that colonial rule displaced. Within former Sokoto, resistance is greater in locations more closely connected to the Caliphate's state-building and administrative institutions, but not near sites associated primarily with its military presence. The form of resistance also follows historically specific predictions: the additional complaints are disproportionately associated with Muslims, whose fiscal obligations changed especially sharply under colonial rule, and part of the increase is concentrated on *jangali*, a tax the Sokoto order had explicitly deemed illegitimate. A direct text-based measure likewise shows greater contestation of the rightfulness of taxation itself. These patterns are difficult to explain with a generic increase in dissatisfaction. Consistent with that interpretation, we find little evidence that the discontinuity is driven by differences in observable population composition, colonial tax instruments or fiscal burden, administrative capacity, enforcement, or public-goods provision.

The historical difference also remains visible in contemporary attitudes toward taxation. In Afrobarometer data, respondents in former Sokoto are less likely to agree that tax authorities have the right to compel payment. Other measures of trust and tax morale move in the same direction but are estimated less precisely. The persistence evidence is therefore strongest for the outcome most directly connected to the historical mechanism: whether the state is perceived to have the right to tax.

The broader implication is that state capacity depends not only on the institutions states build or the coercive and administrative resources they command, but also on how new institutions relate to pre-existing sources of authority and obligation. A reform that expands administrative reach can nevertheless weaken effective policy implementation if it displaces institutions that previously made state demands legitimate. In this sense, the consequences of state-building depend not only on what states create, but also on what they replace.

The results raise broader questions about how states can build new fiscal institutions when they inherit competing sources of authority. When does institutional replacement erode legitimacy, and when can states incorporate pre-existing norms into new systems of taxation? How quickly can legitimacy be rebuilt once disrupted, and through which institutions—political representation, public-goods provision, religious authority, or administrative performance? Understanding how new state institutions interact with the sources of authority citizens already recognize may help explain why similar reforms generate durable compliance in some settings and persistent resistance in others.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001, *91* (5), 1369–1401. 7

_ , _ , and _ , "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 2002, *117* (4), 1231–1294. 7

Adebayo, Akanmu G., "Jangali: Fulani Pastoralists and Colonial Taxation in Northern Nigeria," *The International Journal of African Historical Studies*, 1995, *28* (1), 113–150. 11

Afigbo, A. E., *The Warrant Chiefs: Indirect Rule in Southeastern Nigeria, 1891–1929*, Longman, 1972. 96

Afolabi, Abiodun, "Tax Revolts in Yorubaland, 1925–1955," *Working Paper*, 2023. 96

Alsan, Marcella, "The Effect of the TseTse Fly on African Development," *American Economic Review*, 2015, *105* (1), 382–410. 12

Amirizaniani, Maryam, Elias Martin, Maryna Sivachenko, Afra Mashhadi, and Chirag Shah, "Can llms reason like humans? assessing theory of mind reasoning in llms for open-ended questions," in "Proceedings of the 33rd ACM International Conference on Information and Knowledge Management" 2024, pp. 34–44. 2

Asirvatham, Hemanth, Elliott Mokski, and Andrei Shleifer, "GPT as a Measurement Tool," Technical Report, National Bureau of Economic Research 2026. 7

Banerjee, Abhijit and Lakshmi Iyer, "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*, 2005, *95* (4), 1190–1213. 7

Besley, Timothy, "State capacity, reciprocity, and the social contract," *Econometrica*, 2020, *88* (4), 1307–1335. 1, 6, 7

_ **and Sacha Dray**, "Trust and state effectiveness: The political economy of compliance," *The Economic Journal*, 2024, *134* (662), 2225–2251. 1, 6, 7

_ **and Torsten Persson**, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 2009, *99* (4), 1218–1244. 1

_ **and** _ , "Taxation and Development," *Handbook of Public Economics*, 2013, *5*, 51–110. 1

_ **, Christopher Dann, and Sacha Dray**, "Growth experiences and trust in government," *The Quarterly Journal of Economics*, 2026, *141* (2), 1761–1822. 1, 7

Burg, David F., *A World History of Tax Rebellions: An Encyclopedia of Tax Rebels, Revolts, and Riots from Antiquity to the Present*, New York: Routledge, 2004. 7, 13

Byfield, Judith, "Taxation, Women, and the Colonial State: Egba Women's Tax Revolt," *Meridians: feminism, race, transnationalism*, 2003, *3* (2), 250–277. 96

Cantoni, Davide, Cathrin Mohr, and Matthias Weigand, "The Rise of Fiscal Capacity: Administration and State Consolidation in the Holy Roman Empire," *Econometrica*, 2024, *92* (5), 1439–1472. 1, 6

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik, "A Practical Introduction to Regression Discontinuity Designs: Foundations," *Cambridge Elements: Quantitative and Computational Methods for Social Science*, 2020. 2, 19, 21

Chafe, Kabiru Sulaiman, *State and economy in the Sokoto Caliphate: policies and practicies in the metropolitan districts, 1804-1903*, Ahmadu Bello University Press, 1999. 8

Chambru, Cédric, Emeric Henry, and Benjamin Marx, "The Dynamic Consequences of State Building: Evidence from the French Revolution," *American Economic Review*, 2024, *114* (11), 3578–3622. 1, 6

Chiovelli, Giorgio, Leopoldo Fergusson, Luis R. Martínez, Juan David Torres, and Felipe Valencia Caicedo, "Bourbon Reforms, State Capacity, and Revolution in the Spanish Empire," Working Paper 2504, Universidad de Montevideo 2025. 1, 6

Cohen, Ronald, *The Kanuri of Bornu* Case Studies in Cultural Anthropology, New York: Holt, Rinehart and Winston, 1967. 24

Conley, Timothy G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45. 76, 77

Dell, Melissa, "The Persistent Effects of Peru's Mining *Mita*," *Econometrica*, 2010, *78* (6), 1863–1903. 7

_ , "Deep learning for economists," *Journal of Economic Literature*, 2025, *63* (1), 5–58. 2, 15

_ **and Benjamin A. Olken**, "The Development Effects of the Extractive Colonial Economy: The Dutch Cultivation System in Java," *Review of Economic Studies*, 2020, *87* (1), 164–203. 7

Dincecco, Mark and Gabriel Katz, "State Capacity and Long-Run Economic Performance," *The Economic Journal*, 2016, *126* (590), 189–218. 1, 6

Dupas, Pascaline, Ola Kanaan, and Flora Ponsot, "Colonial Legacies: Forced Labor and the Persistence of Informal Taxation," *Working Paper*, 2024. 7

Falola, Toyin, *Colonialism and Violence in Nigeria*, Indiana University Press, 2009. 96

_ **and Matthew M. Heaton**, *A History of Nigeria*, Cambridge: Cambridge University Press, 2008. 15

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, "Text as Data," *Journal of Economic Literature*, 2019, *57* (3), 535–574. 7, 14

Greif, Avner and Jared Rubin, "Political Legitimacy in Historical Political Economy," in "The Oxford Handbook of Historical Political Economy," Oxford University Press, 2022. 1, 7

_ **and** _ , "Endogenous Political Legitimacy: The Tudor Roots of England's Constitutional Governance," *The Journal of Economic History*, 2024, *84* (3), 655–689. 1, 6, 7

Gwandu, Abubaker Aliu, "Abdullahi b. fodio as a Muslim jurist." PhD dissertation, Durham University 1977. 8

Hiskett, Mervyn, *The Sword of Truth: The Life and Times of the Shehu Usuman dan Fodio*, Oxford University Press, 1973. 8

Hopkins, Antony G., *An Economic History of West Africa*, London: Longman, 1973. 10

Igbafe, Philip A., "Western Ibo Society and its Resistance to British Rule: The Ekumeku Movement 1898–1911," *The Journal of African History*, 1971, *12* (3), 441–459. 13

Jr, Adell Patton, *The Ningi Chiefdom and the African Frontier: Mountaineers and Resistance to the Sokoto Caliphate, ca. 1800–1908*, The University of Wisconsin-Madison, 1975. 46

Jumare, Ibrahim M., "Colonial Taxation in the Capital Emirates of Northern Nigeria," *African Economic History*, 1998, (26), 83–108. 12

Kirk-Greene, Anthony H. M., *The Principles of Native Administration in Nigeria: Selected Documents, 1900–1947*, Oxford University Press, 1965. 2, 11, 12, 19, 47

Korinek, Anton, "Language Models and Cognitive Automation for Economic Research," NBER Working Paper 30957, National Bureau of Economic Research 2023. 2, 7, 15

Lagakos, David, Stelios Michalopoulos, and Hans-Joachim Voth, "American life histories," Technical Report, National Bureau of Economic Research 2025. 2, 7, 15

Last, Murray, *The Sokoto Caliphate*, London: Longmans, 1967. 1, 2, 4, 8, 9, 10, 12, 19, 25, 26, 45, 47, 76, 79, 80, 81

Lavers, John E., "Islam in the Bornu Caliphate: A Survey," *Odù: Journal of Yoruba and Related Studies*, 1971, *5*, 27–53. 24

Levi, Margaret, *Of Rule and Revenue*, University of California Press, 1988. 1, 6

Lovejoy, Paul E., "Revolutionary Mahdism and Resistance to Colonial Rule in the Sokoto Caliphate, 1905–1906," *The Journal of African History*, 1990, *31* (2), 217–244. 12

_ , *Jihad in West Africa during the Age of Revolutions*, Ohio University Press, 2016. 1, 2, 8, 9, 10, 12, 19, 47

Lowes, Sara and Eduardo Montero, "The Legacy of Colonial Medicine in Central Africa," *American Economic Review*, 2021, *111* (4), 1284–1314. 7

Ludwig, Jens, Sendhil Mullainathan, and Ashesh Rambachan, "Large language models: An applied econometric framework," Technical Report, National Bureau of Economic Research 2025. 7

Lugard, Frederick D., "Memorandum on the Taxation of Natives in Northern Nigeria," Technical Report Colonial Reports–Miscellaneous No. 40, Cd. 3309, Colonial Office, London 1907. 11

Mamdani, Mahmood, *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*, Princeton University Press, 1996. 11

Matera, Marc, Misty L. Bastian, and Susan Kingsley Kent, *The Women's War of 1929: Gender and Violence in Colonial Nigeria*, London: Palgrave Macmillan, 2012. 13

Michalopoulos, Stelios and Elias Papaioannou, "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 2013, *81* (1), 113–152. 7, 23

Mohammed, Abubakar Sokoto, *The Satiru Revolt of Peasants and Slaves in Sokoto Caliphate*, Lagos: Malthouse Press, 2024. 13, 22

Murdock, George Peter, *Ethnographic Atlas*, University of Pittsburgh Press, 1967. 3, 19, 58, 60, 62

Nunn, Nathan, "The Long-Term Effects of Africa's Slave Trades," *Quarterly Journal of Economics*, 2008, *123* (1), 139–176. 12

Okoye, Dozie, Roland Pongou, and Tite Yokossi, "New Technology, Better Economy? The Heterogeneous Impact of Colonial Railroads in Nigeria," *Journal of Development Economics*, 2019, *140*, 320–346. 90

Robinson, James A., "Tax Aversion and the Social Contract in Africa," *Journal of African Economies*, 2023, *32* (Supplement 1), i33–i56. 1, 6

Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson, "Measuring News Sentiment," *Journal of Econometrics*, 2022, *228* (2), 221–243. 14

Tambo, David C., "The "Hill Refuges" of the Jos Plateau: A Historiographical Examination," *History in Africa*, 1978, *5*, 201–223. 24

Thomas, Martin, *Violence and Colonial Order: Police, Workers, and Protests in the European Colonial Empires, 1918–1940*, Cambridge University Press, 2012. 96

Tukur, Mahmud Modibbo, "The Imposition of British Colonial Domination on the Sokoto Caliphate, Borno and Neighbouring States, 1897–1914: A Reinterpretation of Colonial Sources." PhD dissertation, Ahmadu Bello University 1979. 8

Turing, Alan M., "Computing Machinery and Intelligence," *Mind*, 1950, *59* (236), 433–460. 2

Tyler, Tom R., *Why People Obey the Law*, Princeton University Press, 2006. 1, 6

Usman, Yusuf Bala, *The Manipulation of Religion in Nigeria, 1977–1987*, Vanguard Publishers, 1987. 96

Part I

Appendix

A Additional Historical Background **45**
A.1 Islamic Taxes Before and During British Rule . . . 45
A.2 Tax Legitimacy in Pre-Colonial Sokoto . . . 45
A.3 Frontier Evidence: Ningi and Zuru . . . 46
A.4 Colonial Nigeria: Key Dates . . . 46
A.5 The Sokoto Caliphate Cavalry . . . 47
A.6 The Satiru Revolt, 1904–06 . . . 47

B Data and Measurement Details **47**
B.1 The National Archives . . . 47
B.2 Historical Map of Locations . . . 47
B.3 Document Topics . . . 51
B.4 Tax Sentiment Prompt . . . 51
B.5 Tax Issue Prompt . . . 52
B.6 Distinctive Vocabulary of the Outcomes . . . 52
B.7 Measuring Complainant Religion . . . 52
B.8 Jangali Complaint Prompt . . . 54
B.9 Tax Legitimacy Prompt . . . 54

C Additional Validation **55**

D RD Validity and Robustness **55**
D.1 Balance at the Boundary . . . 58
D.2 Effects over Time . . . 58
D.3 Ethnic Homelands and Protectorates . . . 58
D.4 Results Within the Northern Protectorate . . . 63
D.5 RD Figures in Latitude and Longitude . . . 63
D.6 Archival Coverage . . . 64
D.7 Spatial and Population Coverage . . . 64
D.8 Geocoding Uncertainty . . . 64
D.9 Border Specificity . . . 64
D.10 Functional Form . . . 72
D.11 Inference . . . 76

E Within-Caliphate Exposure **76**

F Additional Mechanism Evidence **79**

G Alternative Explanations, Supporting Evidence **82**
G.1 Demography . . . 83

G.2 Tax Instruments 87
G.3 Tax Burden 87
G.4 Tax Administration 87
G.5 Tax Enforcement 88
G.6 Public Goods Provision 88

H Tax Revolts and Riots **88**
H.1 Data Construction 96
H.2 Historical Sources 96
H.3 Results 96

A Additional Historical Background

A.1 Islamic Taxes Before and During British Rule

This appendix describes the principal taxes of the Sokoto Caliphate's Islamic fiscal order and their fate under British rule.

Zakat. A religious and charitable levy required of Muslims who meet wealth thresholds, traditionally redistributed to specified beneficiary categories, including the poor and needy. The colonial administration banned its official collection as a religious obligation.

Kharaj. An Islamic land tax imposed on non-Muslim landowners or on conquered lands that remained in the possession of their original owners, assessed based on the annual agricultural yield of the land. Under colonial rule it was banned or reclassified and replaced by a land tax based on land and produce, imposed irrespective of religion.

Jizya. A poll tax on non-Muslims, in exchange for protection and exemption from military service. Under colonial rule, *jizya* was abolished and replaced by secular head and poll taxes imposed irrespective of religion.

Jangali. The cattle tax occupies a distinct position. It predated the Caliphate, which rejected it on Islamic grounds: as documented in Appendix A.2, the term was expressly dropped in Sokoto as unpopular and illegal according to the Shaikh. The colonial state reintroduced it. Section 6.2 shows that *jangali* accounts for part of the increase in formal complaints at the Caliphate boundary.

A.2 Tax Legitimacy in Pre-Colonial Sokoto

Section 2.1 quotes the passage on revenue in pre-colonial Sokoto from Last (1967, pp. 103–104). The passage's footnotes, reproduced from Last (1967), are:

> "[49] The legal taxes and regulations for using them are given, for example, in 'Uthmān b. Fodiye, Kitāb al-farq, and 'Abdullāh b. Fodiye, Dīyā' al-hukkām. As books by 'Abdullāh and the Shaikh' were the textbooks for the Sokoto administration, their authority is quoted for the law. It is clear however that the law was not always fulfilled in detail, and that legal terminology became used as rough descriptions of traditional practice, modified after the jihād.
>
> [50] Kitāb al-farq (ed. Hiskett, B.S.O.A.S., 1960), p. 561. Letter of 'Ali b. Bello to Khalilu b. 'Abdullāh in 'Abd al-Qādir b. Gidado, Majmū'.
>
> [51] The Hausa term was kurdin kasa ('land tax'). Cf. J. A. Burdon, Notes on Tribute in Sokoto Province (7 March 1904), and E. J. Arnett, Gazetteer of Sokoto Province (p. 69). As

administrators responsible for collecting tax in Sokoto, it is quite possible they were misled; but though revenue reached Sokoto from the surrounding areas, it probably did not come as kurdin kasa. Moreover, no tolls or craft taxes were paid in Sokoto, according to al-haji Junaidu.

[52] P. G. Harris, Sokoto Provincial Gazetteer, p. 105. Cf. Burdon, loc. cit."

A.3 Frontier Evidence: Ningi and Zuru

Two frontier polities illustrate where the Caliphate's fiscal legitimacy ended.

Ningi. In 1846, Mallam Hamza, an Islamic scholar and the founder of the Ningi state, fled Kano to escape the land tax (*kurdin kasa*) and allied with non-Muslim mountaineers who had long resisted Caliphate taxation and slave raiding. Sokoto's right to tax was never recognized; the state raided rather than taxed. "Warring was their farming" (Patton Jr 1975, pp. 235–236).

Zuru. Zuru was a comparable frontier enclave, along with Kari-Kari, Gwari, Kebbi, and Kwatarkwashi, that Sokoto never incorporated into its fiscal structure; tribute demands remained contested throughout the nineteenth century. Formal resistance ended in 1902: facing British forces, the leader Dan Yaya refused surrender, saying "Your lies are finished" before arranging his own death rather than submitting.

A.4 Colonial Nigeria: Key Dates

- **1861:** Lagos annexed as a British colony (coastal foothold).
- **1885–1893:** Oil Rivers Protectorate; renamed Niger Coast Protectorate (1893–1900).
- **1886–1900:** Royal Niger Company (chartered company) administers the "Niger Territories" under company rule.
- **1900:** Britain assumes direct control of chartered territories; creation of the Protectorate of Northern Nigeria and the Protectorate of Southern Nigeria (alongside Lagos).
- **1903:** Conquest of Kano and Sokoto; consolidation of British authority in the North.
- **1906:** Lagos Colony merged into Southern Nigeria.
- **1914:** Amalgamation into the Colony and Protectorate of Nigeria (administrative north–south division largely retained).

A.5 The Sokoto Caliphate Cavalry

Historians emphasize military reach and ecological constraints in shaping Sokoto's southern boundary (Kirk-Greene 1965; Last 1967; Lovejoy 2016). Appendix Figure A1 shows the Sokoto Caliphate cavalry.



Figure A1: The Sokoto Caliphate Cavalry

A.6 The Satiru Revolt, 1904–06

Appendix Figure A2 illustrates an example of tax revolt in the early colonial period: the Satiru revolt of 1904–06. Appendix H describes the systematic data on tax revolts and riots.

B Data and Measurement Details

B.1 The National Archives

The documents come from the three zonal offices of the National Archives of Nigeria. Appendix Figure B1 shows the Kaduna archive, which holds the northern collection, and the physical state of the records. Figure B2 reproduces two examples of documents, including a page from a tax-complaints file.

B.2 Historical Map of Locations

Appendix Figure B3 shows the geolocated document locations on the 1924 map of Nigeria, together with the Sokoto Caliphate boundary.

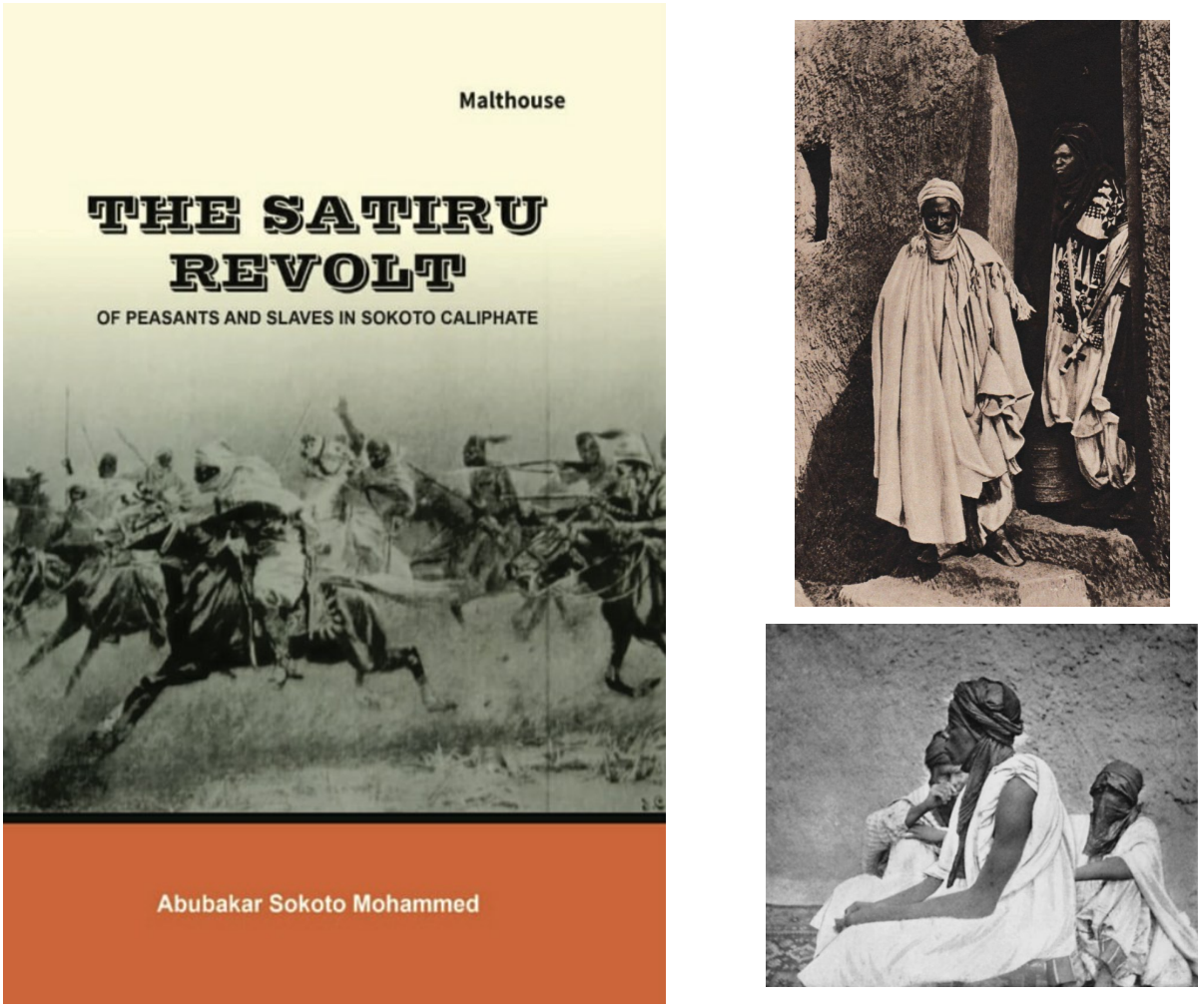
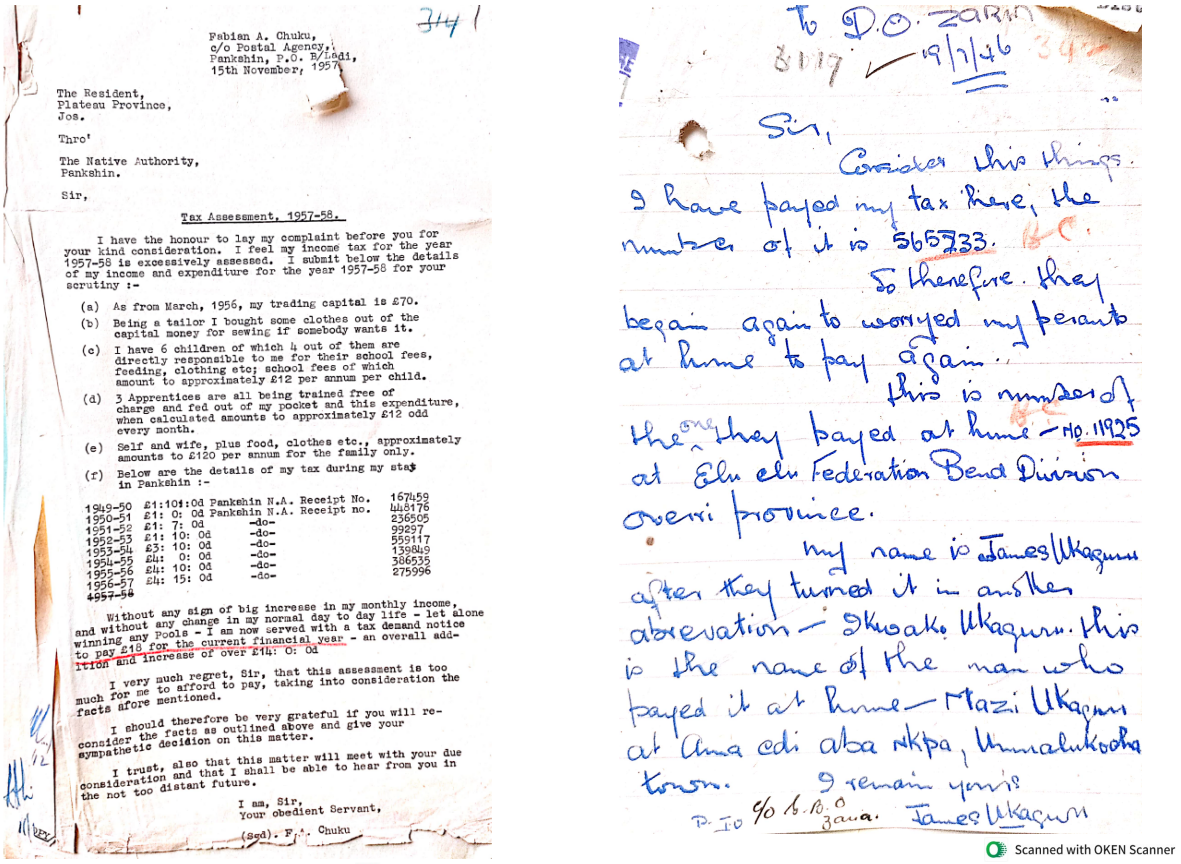


Figure A2: The Satiru Revolt, 1904–06



Figure B1: The National Archives, Kaduna



Fabian A. Chuku,
c/o Postal Agency,
Pankshin, P.O. B/Ladi,
15th November, 1957

The Resident,
Plateau Province,
Jos.

Thro'

The Native Authority,
Pankshin.

Sir,

Tax Assessment, 1957-58.

I have the honour to lay my complaint before you for your kind consideration. I feel my income tax for the year 1957-58 is excessively assessed. I submit below the details of my income and expenditure for the year 1957-58 for your scrutiny :-

(a) As from March, 1956, my trading capital is £70.

(b) Being a tailor I bought some clothes out of the capital money for sewing if somebody wants it.

(c) I have 6 children of which 4 out of them are directly responsible to me for their school fees, feeding, clothing etc; school fees of which amount to approximately £12 per annum per child.

(d) 3 Apprentices are all being trained free of charge and fed out of my pocket and this expenditure, when calculated amounts to approximately £12 odd every month.

(e) Self and wife, plus food, clothes etc., approximately amounts to £120 per annum for the family only.

(f) Below are the details of my tax during my stay in Pankshin :-

1949-50	£1:10:0d	Pankshin N.A. Receipt No.	167459
1950-51	£1: 0: 0d	Pankshin N.A. Receipt no.	448176
1951-52	£1: 7: 0d	-do-	236505
1952-53	£1: 10: 0d	-do-	99297
1953-54	£3: 10: 0d	-do-	559117
1954-55	£4: 0: 0d	-do-	139849
1955-56	£4: 10: 0d	-do-	386535
1956-57	£4: 15: 0d	-do-	275996
~~1957-58~~			

Without any sign of big increase in my monthly income, and without any change in my normal day to day life - let alone winning any Pools - I am now served with a tax demand notice to pay £18 for the current financial year - an overall addition and increase of over £14: 0: 0d

I very much regret, Sir, that this assessment is too much for me to afford to pay, taking into consideration the facts afore mentioned.

I should therefore be very grateful if you will reconsider the facts as outlined above and give your sympathetic decidion on this matter.

I trust, also that this matter will meet with your due consideration and that I shall be able to hear from you in the not too distant future.

I am, Sir,
Your obedient Servant,

(Sgd). F. A. Chuku

To D.O. Zaria
8117 — 19/1/46

Sir,

Consider this things.
I have payed my tax here, the number of it is 565733.
So therefore they begain again to worryed my peasants at home to pay again.
this is numbers of the one they payed at home — No. 11925 at Edu edu Federation Bend Division Ovenri province.
my name is James Ukagwu after they turned it in another abrevation — Okoako Ukagwu. this is the name of the man who payed it at home — Nazi Ukagwu at Ama edi aba nkpa, Umuahukoaha town. I remain yours
P.I.O c/o N.A.O Zaria. James Ukagwu

Scanned with OKEN Scanner

Figure B2: Examples of Documents

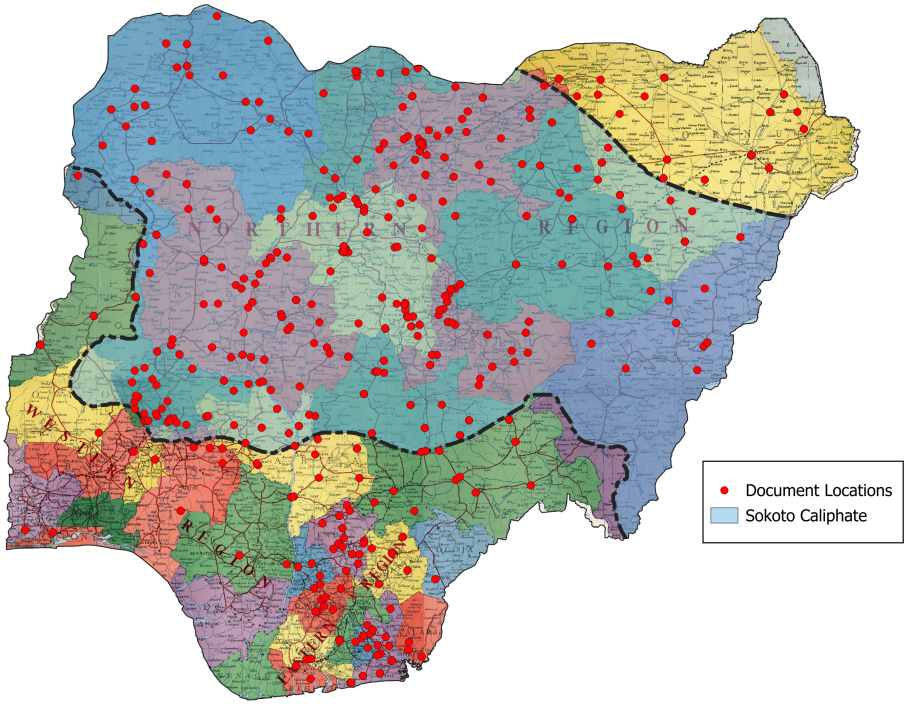


Figure B3: Document Locations on the 1924 Map of Nigeria

Notes: The figure shows the geolocated document locations, matched to historical localities on the 1924 map, with the Sokoto Caliphate boundary overlaid.

B.3 Document Topics

A similar LLM procedure classifies document topics; documents may cover multiple topics. The categories are:

- Colonial Taxation Policies & Administration
- Tax Enforcement & Collection Methods
- State Finance & Revenue
- Indigenous Authority & Control
- Demographics & Social Organization
- Tax Resistance and Evasion
- Health and Labor Conditions
- Violent Conflicts
- Unidentifiable Topic

B.4 Tax Sentiment Prompt

The tax sentiment score of Section 3.2.2 is produced by the following prompt, reproduced verbatim.

> You are an expert historical analyst assigned to a data coding task. Your task is to classify the document's sentiment by returning a single integer grade from -1 to 4.
>
> - **-1 (unanalyzable)**: The text is garbled, corrupted, or consists of indecipherable characters due to poor OCR of handwriting. It is impossible to determine its original meaning, even though it may be long
> - **0 (very negative)**: Documents detailing acute crisis, violence, or severe social breakdown. The tone is one of emergency, conflict, or extreme distress. Look for descriptions of: violent clashes, riots, military or police actions (punitive expeditions), reports of overt threats against officials, or official communications expressing severe condemnation or announcing punishments.
> - **1 (negative)**: Content describing problems, disagreements, or complaints from either officials or civilians. The tone is clearly negative because it focuses on an issue, but it does not describe violence or riots. Look for: formal petitions from civilians against taxes or chiefs, reports of tax evasion or refusal to cooperate, accusations of corruption against officials, discussions of legal disputes, or official complaints about administrative difficulties.
> - **2 (neutral)**: Impersonal, objective, or purely administrative content. The document's primary purpose is to record or transmit information without significant emotional or persuasive language. Look for: financial ledgers, tax rolls, population or livestock censuses, routine meeting minutes, gazetteers, and standard transmittal letters ('I forward herewith...').
> - **3 (moderately positive)**: Documents indicating successful administration, cooperation, and the smooth functioning of the colonial system. The tone is one of approval,

satisfaction, or positive progress. Look for: reports of tax collection being completed on time and without issue, notes of praise for a specific chief or community, records of successful dispute resolutions, or positive reports on public works projects (e.g., wells, roads).

- **4 (very positive)**: Content expressing exceptional approval, success, or enthusiasm. The tone is highly positive and celebratory, going beyond routine satisfaction. Look for: reports giving special praise to individuals or communities for their excellent work, declarations of loyalty, reports of exceptional economic prosperity (e.g., 'bumper crop'), or the celebration of a major administrative achievement.

B.5 Tax Issue Prompt

The tax-issue classification underlying the formal tax complaint outcome of Section **??** is produced by the following prompt, reproduced verbatim.

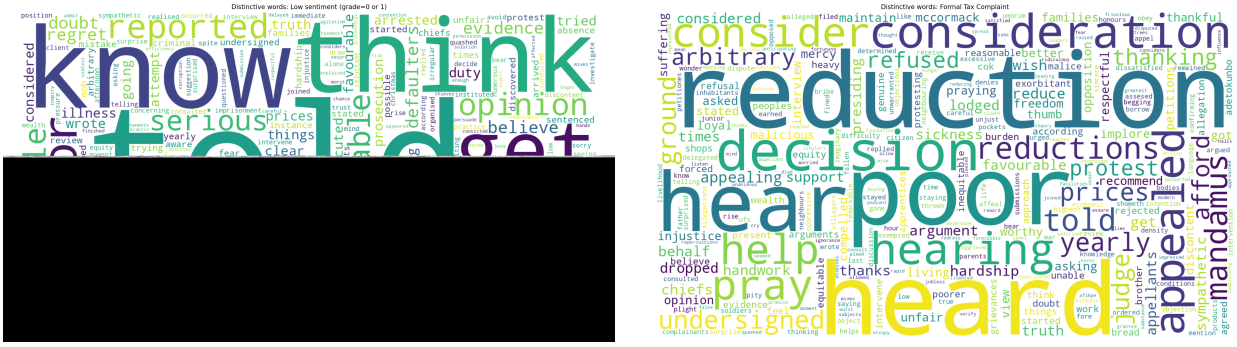
> You are a meticulous historical analyst. Your task is to read a document from colonial Nigeria and identify all described issues related to taxation, from administrative problems to open rebellion. A conflict over a chieftaincy dispute or a report on road building are NOT tax issues. You must link the event described to the subject of taxation.
>
> - Formal Tax Complaint: A written complaint about taxation submitted to an authority
> - Tax Evasion: Secretive actions to illegally avoid paying taxes
> - Corruption: The abuse of an official's position for personal financial gain
> - Bad Administration: Problems of incompetence, inefficiency, or systemic dysfunction
> - State Coercion or Punishment: Actions by authorities to enforce tax laws
> - Tax Riot: A localized and often spontaneous outbreak of violence
> - Tax Revolt: An organized and widespread challenge to the government's authority
> - No Evidence of Tax Issues: A routine document, or a document describing issues unrelated to taxation
>
> Output: Tax issue detected (yes/no), tax issue types, justification

B.6 Distinctive Vocabulary of the Outcomes

Appendix Figure B4 shows the words most distinctive of documents coded as expressing negative tax sentiment and of documents coded as formal tax complaints, the two outcomes described in Sections 3.2.2.

B.7 Measuring Complainant Religion

For each document already flagged as a formal tax complaint, the religion of the complainant is identified in two stages.



(a) Negative Tax Sentiment (b) Formal Tax Complaints

Figure B4: Distinctive Vocabulary of the Two Outcomes

Notes: Panel (a) shows the words most distinctive of documents coded as expressing negative tax sentiment; panel (b) the words most distinctive of documents coded as formal tax complaints.

Name Extraction. An LLM (Gemini 2.5 Flash) reads the OCR'd document text and extracts the actual complainant's name, excluding officials or titleholders writing *about* someone else's complaint. Multiple complainants in one document are distinguished as joint (filing together) versus independent (separately listed), and the fullest available form of a name is preferred over initials. The excerpt below is from the extraction prompt, verbatim:

> "Analyze the text below, a subdocument from Northern Nigeria during the colonial period.
>
> This subdocument has been automatically flagged as likely containing a formal tax complaint, but that tag can be wrong. Read the actual text before assuming a complaint exists.
>
> Otherwise, identify ONLY the person who made/filed the complaint (the complainant), not the person or authority the complaint is addressed to, and not any other people merely mentioned in the text.
>
> EXCLUDE ANYONE HOLDING AN OFFICIAL, ADMINISTRATIVE, OR TRADITIONAL TITLE, even if they are personally disputing their own tax. This includes British colonial officials (Resident, District Officer, etc.) AND indirect-rule/Native Authority titleholders (Emir, Chief, District Head, etc.).
>
> OFFICIALS WRITING ABOUT SOMEONE ELSE'S COMPLAINT: an official or titleholder frequently writes, forwards, or responds to a complaint that was actually made by a DIFFERENT, specific named individual. In these cases, extract that OTHER specific named person as the complainant, not the official/writer.
>
> MULTIPLE NAMED CO-COMPLAINANTS: if two or more SPECIFIC named individuals jointly file a single complaint, list all of them, separated by '&' in both First_Name and Last_Name, in the same order."

Religion Classification. Each extracted name is matched against curated repositories of Muslim and non-Muslim names and variants; if no match is found, the LLM classifies

the name as Muslim, Christian, or unclear using linguistic and cultural naming conventions. Repository match rates differ by category: Muslim names are matched roughly twice as often as Christian names, so Christian and unclear classifications rely more heavily on LLM judgment.

B.8 Jangali Complaint Prompt

The *jangali* classification of Section 6.2 is implemented as a single, two-step LLM prompt; the excerpt below is from that prompt, verbatim.

> "You are a meticulous historical analyst doing document coding for a research dataset.
>
> Read the document text below (OCR'd colonial/historical material). Determine (1) whether this document contains a FORMAL TAX COMPLAINT, an explicit grievance about taxation directed at or reported to an authority, such as a petition, letter of protest, or appeal against an assessment, and if so, (2) whether that complaint is specifically about JANGALI, the colonial-era per-head cattle tax, as distinct from general tax, income tax, or other levies, even when discussed alongside it.
>
> EXCLUDE from Step 2: Jangali appearing only as a line item on an administrative form with no grievance attached; complaints about tax in general where Jangali is named only in passing; complaints about a different tax where Jangali is merely listed alongside it.
>
> Provide a short quote or paraphrase grounding the determination."

B.9 Tax Legitimacy Prompt

The contested-legitimacy measure of Section 6.2 is produced by the following prompt, reproduced verbatim.

> You are a meticulous historical analyst doing document coding for a research dataset. Read the document text below. Your job is to measure TAX LEGITIMACY expressed or reported in the document. Definition (use this throughout). Tax legitimacy= perceived continuity, rightful authority, and moral justification behind taxation. This is DISTINCT from tax rates or the level of taxation, enforcement/coercion intensity (arrests, seizures, patrols), administrative capacity/competence, general "sentiment" (positive/negative tone) unless it is explicitly about rightfulness/justification.
>
> Important rules:
>
> - Do NOT infer legitimacy from violence/coercion alone. Only treat it as legitimacy-relevant if the text frames the taxation/authority as rightful or not rightful
> - If the document is purely administrative (tax roll, transmittal letter) with no legitimacy language, default to score = 2 ("unclear") with low-to-moderate confidence

- If there are multiple taxes/authorities discussed and they differ, score the legitimacy of the PRIMARY tax demand in the document; mention the heterogeneity briefly in the justification
- Be conservative: when evidence is thin, use score 2 and lower confidence rather than guessing

- **-1 (unanalyzable)**: The text is garbled/indeciphrable, cannot reliably infer meaning. If the text is readable but purely administrative, do NOT sure -1, use 2.
- **1 (contested)**:
  - Very low legitimacy: taxation is framed as unlawful/usurped/illicit, explicit claims it violated religion/custom/law, noncompliance framed as justified ("no right to tax" logic)
  - Low legitimacy: clear questioning of rightfulness/fairness/authority, frequent legitimacy-based grievances, "arbitrary", "foreign", "against customs", "against religion", "no consent", "abuse of authority," but not total repudiation throughout
- **2 (unclear)**: mostly administrative or factual, or mixed signal with no dominant legitimacy frame, legitimacy is not clearly asserted or challenged
- **3 (endorsed)**:
  - High legitimacy: taxation is framed as rightful and within recognized authority, compliance is framed as a duty, reference to lawful authority/customary basis/accountability disputes may exist but do not challenge the right to tax
  - Very high legitimacy: explicit moral/religious/consitutional justification, strong language of duty/obligation/loyalty, voluntary compliance, taxation linked to legitimate rule and accepted norms

C Additional Validation

This appendix reports the Turing test of Section 3.3 for the secondary LLM classifications: state coercion or punishment, tax evasion, corruption, and administrative failure, as well as for the contested-legitimacy measure of Section 6.2. In each case, every rater, the five human research assistants and the LLM, is benchmarked against the modal human score on the audit sample of 100 randomly selected documents. Appendix Figures C1, C2, C3, and C4 report the results for the four secondary classifications, and Appendix Figure C5 for contested legitimacy, in the full and balanced audit samples.

D RD Validity and Robustness

This appendix collects the evidence supporting the validity of the spatial RD design of Section 4 and the robustness checks summarized in Section 5.2.

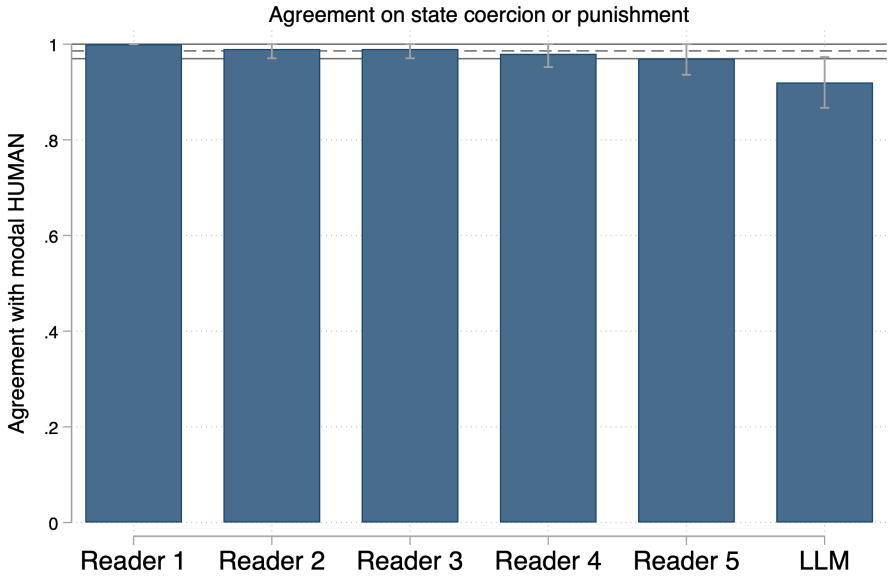


Figure C1: Turing Test: State Coercion or Punishment

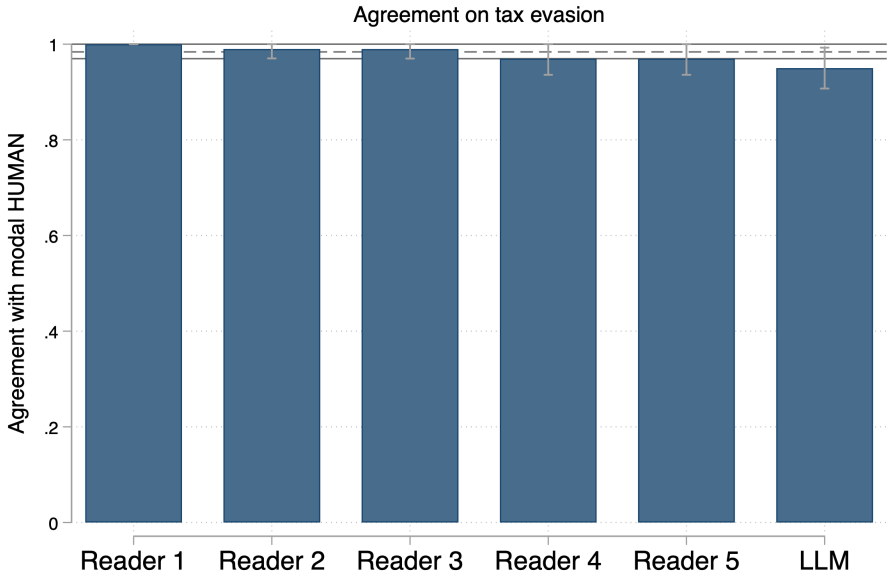


Figure C2: Turing Test: Tax Evasion

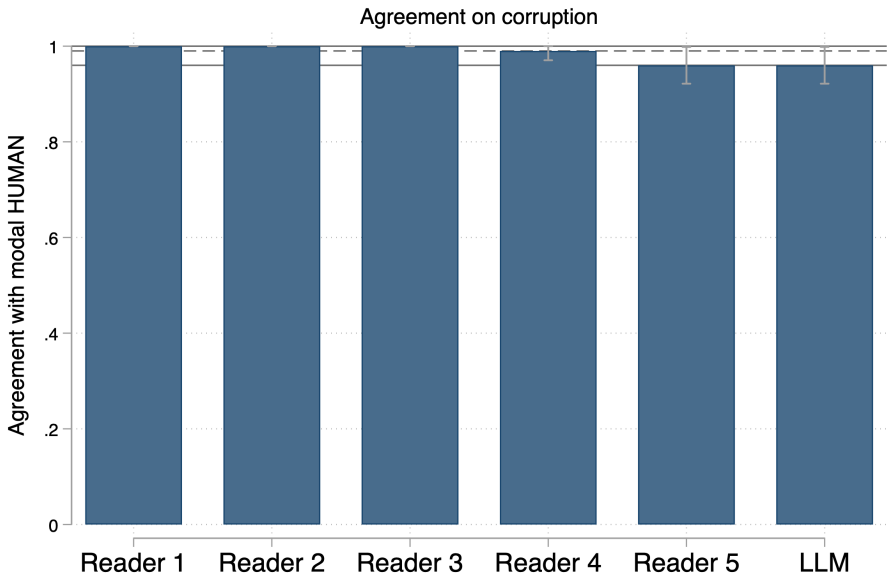


Figure C3: Turing Test: Corruption

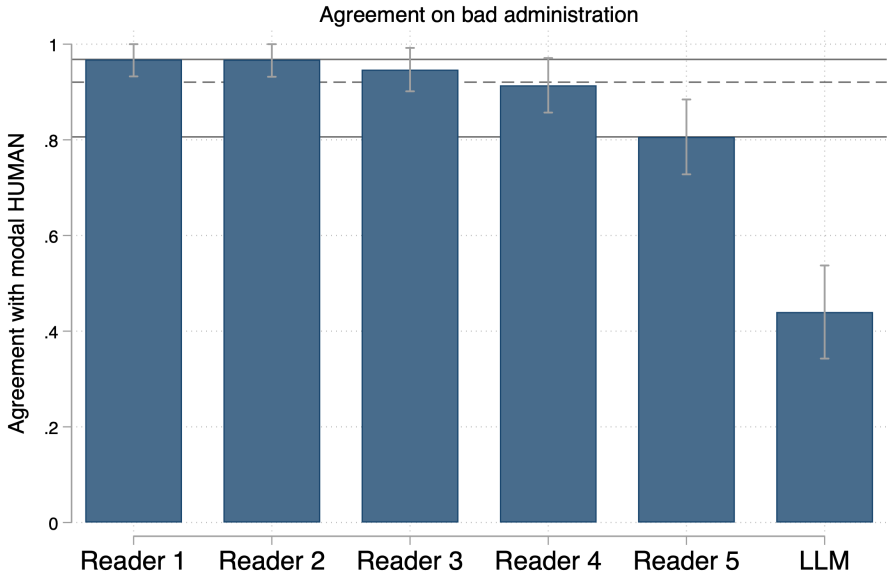
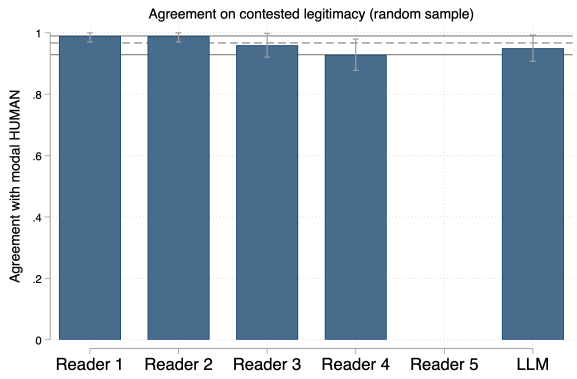
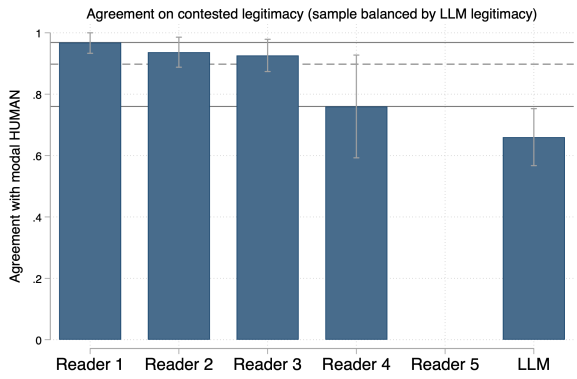


Figure C4: Turing Test: Administrative Failure



(a) All Documents



(b) Balanced Sample

Figure C5: Turing Test: Contested Legitimacy

D.1 Balance at the Boundary

Appendix Table D1 reports balance tests for the geographic and pre-Sokoto characteristics of Section **??**, discussed in Section **??**. Appendix Table D2 reports balance tests for colonial-era characteristics: religion, occupation, and literacy shares from the 1931 and 1952/53 Censuses, civil servants per 1,000 population, public goods from the 1914 and 1915 Blue Books, and transportation infrastructure from the georeferenced 1924, 1934, and 1949 maps.

D.2 Effects over Time

Table D3 estimates equation (1) separately for 1900–1939, 1940–1949, and 1950–1979 at the 100-kilometer bandwidth. For both negative tax sentiment and formal complaints, the estimated discontinuity is substantially larger in the later periods, with the largest effects observed after 1950.

D.3 Ethnic Homelands and Protectorates

Panel (a) of Appendix Figure D1 maps the ethnic homelands of the *Ethnographic Atlas* (Murdock 1967), with the Sokoto Caliphate boundary overlaid; panel (b) shows the boundary together with the colonial protectorate boundaries. Appendix Table D4 shows the RD results with ethnic-homeland fixed effects.

Table D1: Balance at the Boundary: Geographic and Pre-Sokoto Characteristics

	Bandwidth			
	(1) 50km	(2) 100km	(3) 150km	(4) 200km
Panel A: Geographic Controls				
Distance to Nearest River (km)	1.107	13.630	10.478	9.050
	(64.417)	(61.000)	(59.259)	(60.953)
Distance to Coast (km)	-2.705	18.359	10.625	22.134
	(142.111)	(141.532)	(135.604)	(132.796)
Cassava Suitability	52.764	198.141	301.977	-669.333
	(805.072)	(624.413)	(588.611)	(613.348)
Cocoa Suitability	-100.447	-129.202	-555.630	-692.948
	(900.548)	(918.490)	(910.394)	(927.591)
Maize Suitability	-329.726	33.528	217.607	109.466
	(586.693)	(500.268)	(555.240)	(599.674)
Millet Suitability	-164.035	-28.289	18.380	69.270
	(804.021)	(991.306)	(938.630)	(899.615)
Palm Oil Suitability	68.873	136.318	-155.619	-215.294
	(619.704)	(641.575)	(586.596)	(579.512)
Yam Suitability	35.164	114.291	150.452	-350.451
	(823.443)	(702.773)	(678.596)	(676.080)
Mean Length of Growing Period	0.594	3.375	-4.830	-5.139
	(32.452)	(33.930)	(32.752)	(31.894)
Mean Precipitation	23.805	26.718	-6.848	-8.567
	(193.446)	(199.463)	(193.360)	(185.665)
Mean Temperature	0.216	0.264	0.238	0.298
	(0.646)	(0.806)	(0.766)	(0.702)
Mean Elevation	6.614	-0.471	3.504	5.776
	(18.281)	(12.743)	(12.770)	(12.874)
Mean Ruggedness	5.539	9.826	8.738	11.338
	(32.533)	(36.947)	(34.404)	(35.450)
Malaria Suitability	-0.342	-0.244	-0.050	0.010
	(0.875)	(0.928)	(0.972)	(1.012)
Tsetse Suitability	0.004	0.018	0.015	0.001
	(0.215)	(0.206)	(0.199)	(0.192)
Panel B: Pre-Sokoto Controls				
Population Density (1800)	1.561	-0.357	-1.176	-0.484
	(4.760)	(5.058)	(5.342)	(5.829)
Slave Exports (000s)	-8.342	19.338	13.734	11.472
	(132.068)	(111.227)	(102.807)	(100.873)
Observations	362	700	1,000	1,265

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Each cell reports the discontinuity at the Sokoto Caliphate boundary in the predetermined characteristic named in the row, at the fixed bandwidth indicated in the column. Panel A covers geographic characteristics; Panel B covers pre-Sokoto historical characteristics. Data sources are described in Section **??**. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

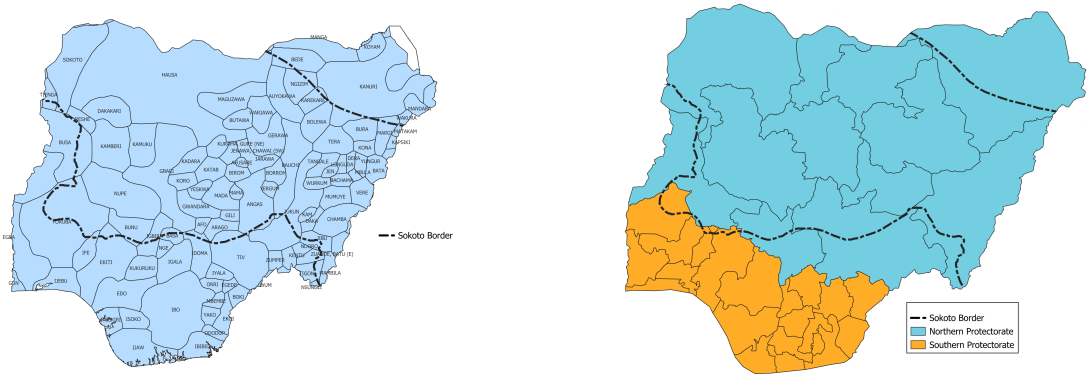
Table D2: Balance at the Boundary: Colonial Controls

	Bandwidth			
	(1) 50km	(2) 100km	(3) 150km	(4) 200km
Panel C: Colonial Controls (1 of 2)				
Religion				
% Christian^{C}	-3.297	-3.275	-3.854	-4.739
	(3.959)	(3.808)	(3.629)	(3.519)
% Muslim^{C}	-1.272	2.542	2.726	2.244
	(6.106)	(5.998)	(5.665)	(5.837)
Occupation				
% Agriculture^{C}	2.209	0.925	0.175	-0.491
	(2.525)	(2.482)	(2.631)	(2.611)
% Crafts / Trade^{C}	-1.919	-1.060	-0.529	-0.009
	(2.180)	(2.140)	(2.279)	(2.267)
% Admin / Professional^{C}	-0.144	-0.003	0.085	0.089
	(0.212)	(0.198)	(0.197)	(0.202)
Civil Servants per 1,000 (1931)^{C}	0.456	0.364	0.271	0.268
	(1.333)	(1.297)	(1.390)	(1.447)
Native Admin per 1,000 (1931)^{C}	-0.181	-0.130	-0.043	-0.002
	(0.683)	(0.723)	(0.764)	(0.788)
Literacy				
% Literate^{C}	-0.490	-0.511	-0.447	-0.478
	(0.571)	(0.590)	(0.574)	(0.545)
^{C} 1931 & 1952/53 Census.				

	Bandwidth			
	(1) 50km	(2) 100km	(3) 150km	(4) 200km
Panel C: Colonial Controls (2 of 2)				
Public Goods				
Health Facilities^{B}	-0.021	0.043	0.011	-0.027
	(0.302)	(0.310)	(0.317)	(0.342)
Medical Facilities^{M}	-0.004	-0.058	-0.079	-0.052
	(0.091)	(0.087)	(0.084)	(0.078)
Education Facilities^{B}	-0.285	0.183	0.436	0.604
	(0.708)	(0.736)	(0.780)	(0.830)
Transport Facilities^{B}	-0.095	0.055	0.015	-0.035
	(0.322)	(0.341)	(0.384)	(0.452)
Water/Sanitation/Electricity Utilities^{B}	-0.101	0.038	0.027	0.017
	(0.325)	(0.345)	(0.366)	(0.406)
British Administrative Offices^{B}	0.057	-0.072	-0.220	-0.344
	(0.335)	(0.321)	(0.315)	(0.323)
British Police Posts^{B}	0.058	-0.025	-0.136	-0.209
	(0.292)	(0.281)	(0.287)	(0.290)
Government Infrastructure^{B}	0.004	0.014	0.065	0.093
	(0.431)	(0.422)	(0.427)	(0.450)
Post Office/Telegraph/Telephone^{M}	0.043	0.021	-0.012	-0.011
	(0.032)	(0.035)	(0.034)	(0.032)
Transportation Infrastructure				
Road Density (km)^{M}	2.131	4.121	1.195	-2.480
	(8.024)	(5.692)	(4.836)	(4.406)
Railroad Density (km)^{M}	1.186	0.410	0.213	0.383
	(0.569)	(0.607)	(0.546)	(0.571)
Observations	344	672	970	1,233

Standard errors in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
^{B} 1914 & 1915 Blue Books. ^{M} Georeferenced 1924, 1934, and 1949 maps.

Notes: Each cell reports the discontinuity at the Sokoto Caliphate boundary in the colonial-era characteristic named in the row, at the fixed bandwidth indicated in the column. Superscript C denotes the 1931 and 1952/53 Censuses, B the 1914 and 1915 Blue Books, and M the georeferenced 1924, 1934, and 1949 maps. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Ethnic Homelands (b) Colonial Protectorates

Figure D1: The Sokoto Boundary, Ethnic Homelands, and Colonial Protectorates

Notes: Panel (a) shows the ethnic homelands of the *Ethnographic Atlas* (Murdock 1967) with the Sokoto Caliphate boundary overlaid. Panel (b) shows the Sokoto Caliphate boundary together with the colonial protectorate boundaries.

Table D3: Growing Tax Resistance over the Colonial Period

	100km Bandwidth		
	(1) 1900–1939	(2) 1940–1949	(3) 1950–1979
Panel A: Negative Sentiment			
In Sokoto	-0.032	0.257***	0.388***
	(0.029)	(0.067)	(0.019)
Observations	2016	1618	5295
Control Mean	0.322	0.198	0.238
Panel B: Formal Tax Complaint			
In Sokoto	0.032**	0.095***	0.469***
	(0.014)	(0.015)	(0.019)
Observations	2016	1618	5295
Control Mean	0.074	0.075	0.075
Geographic Controls	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) at the 100 km bandwidth, estimated separately for documents dated 1900–1939, 1940–1949, and 1950–1979. Outcomes, unit of observation, controls, and fixed effects are as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D4: RD Results with Ethnic Homeland Fixed Effects

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	-0.126***	-0.089**	-0.016	0.063*	0.112***
	(0.042)	(0.044)	(0.038)	(0.034)	(0.031)
Observations	7156	5778	8931	15995	22714
Control Mean	0.231	0.233	0.251	0.284	0.300
Panel B: Formal Tax Complaint					
In Sokoto	0.041***	0.006	0.122***	0.079***	0.101***
	(0.012)	(0.005)	(0.007)	(0.011)	(0.012)
Observations	6924	5778	8931	15995	22714
Control Mean	0.063	0.055	0.075	0.091	0.103
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Ethnicity FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) adding ethnic-homeland fixed effects from the *Ethnographic Atlas* (Murdock 1967), mapped in Appendix Figure D1. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.4 Results Within the Northern Protectorate

As panel (b) of Appendix Figure D1 shows, part of the Caliphate boundary lies entirely within the Northern Protectorate. Appendix Table D5 shows that the results hold within the Northern Protectorate, where the colonial tax code was uniform.

Table D5: RD Results Within the Northern Protectorate

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.124***	0.124***	0.251***	0.184***	0.140***
	(0.044)	(0.044)	(0.040)	(0.037)	(0.036)
Observations	5735	5735	8882	15593	20757
Control Mean	0.218	0.218	0.245	0.256	0.257
Panel B: Formal Tax Complaint					
In Sokoto	1.563***	0.102***	0.264***	0.216***	0.165***
	(0.232)	(0.007)	(0.020)	(0.015)	(0.018)
Observations	2921	5735	8882	15593	20757
Control Mean	0.019	0.034	0.067	0.070	0.070
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

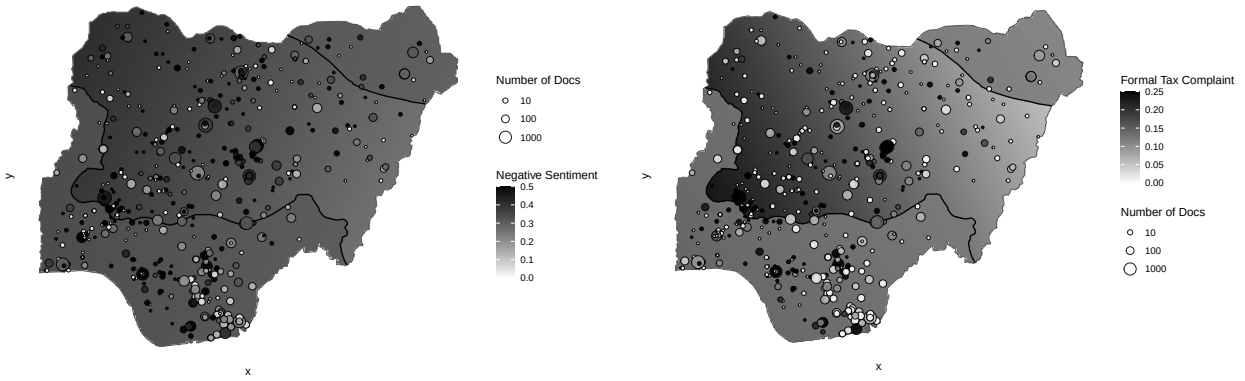
Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) restricting the sample to the Northern Protectorate, which keeps the colonial tax code fixed. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.5 RD Figures in Latitude and Longitude

Appendix Figure D2 reproduces Figure 6 with the RD surface estimated linearly in latitude and longitude.



(a) Negative Tax Sentiment (b) Formal Tax Complaints

Figure D2: Discontinuities at the Sokoto Boundary, Linear in Latitude and Longitude

D.6 Archival Coverage

Appendix Tables D6, D7, and D8 show that the main results are robust to document-type fixed effects, archive-by-decade fixed effects, and archive-by-year fixed effects, respectively.

D.7 Spatial and Population Coverage

Appendix Table D9 reports RD estimates for the number of documents in 25 km grid cells, in 10 km grid cells, and per 1,000 residents at the division level. Appendix Table D10 reweights each observation by the inverse of its location's document count, so that every location receives equal weight; Appendix Table D11 reweights by inverse division-level population from the 1952 Census, so that every person receives equal weight.

D.8 Geocoding Uncertainty

Appendix Table D12 reports a donut RD that drops potentially ambiguous locations by excluding a 10-kilometer buffer around the border.

D.9 Border Specificity

Appendix Table D13 addresses the segment of the boundary adjacent to Kanem–Bornu, excluding Kanem–Bornu observations in Panel A and treating Kanem–Bornu as part of the Sokoto Caliphate in Panel B. Appendix Table D14 addresses the Jos Plateau, excluding Jos Plateau observations in Panel A and carving the plateau out of the Sokoto boundary in

Table D6: RD Results with Document-Type Fixed Effects

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.036*	0.093***	0.110***	0.063***	0.069***
	(0.021)	(0.020)	(0.019)	(0.021)	(0.026)
Observations	7156	5778	8931	15995	22714
Control Mean	0.231	0.233	0.251	0.284	0.300
Panel B: Formal Tax Complaint					
In Sokoto	0.080***	0.074***	0.064***	0.025	0.035**
	(0.014)	(0.015)	(0.015)	(0.016)	(0.017)
Observations	6263	5778	8931	15995	22714
Control Mean	0.065	0.055	0.075	0.091	0.103
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) adding document-type fixed effects. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D7: RD Results with Archive-by-Decade Fixed Effects

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.164***	0.185***	0.227***	0.133***	0.113***
	(0.034)	(0.037)	(0.033)	(0.030)	(0.032)
Observations	8183	5776	8929	15993	22712
Control Mean	0.250	0.233	0.251	0.284	0.300
Panel B: Formal Tax Complaint					
In Sokoto	0.098***	0.134***	0.215***	0.118***	0.092***
	(0.020)	(0.019)	(0.027)	(0.021)	(0.021)
Observations	6205	5776	8929	15993	22712
Control Mean	0.065	0.055	0.075	0.091	0.103
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) adding archive-by-decade fixed effects. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D8: RD Results with Archive-by-Year Fixed Effects

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.071**	0.197***	0.239***	0.155***	0.123***
	(0.036)	(0.036)	(0.030)	(0.026)	(0.031)
Observations	7164	5776	8929	15993	22712
Control Mean	0.232	0.233	0.251	0.284	0.300
Panel B: Formal Tax Complaint					
In Sokoto	0.098***	0.091***	0.205***	0.123***	0.089***
Observations	6213	5776	8929	15993	22712
Control Mean	0.065	0.055	0.075	0.091	0.103
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) adding archive-by-year fixed effects. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D9: No Discontinuity in Document Counts at the Boundary

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Number of Documents (25 km grid)					
In Sokoto	44.168	55.630	47.986	48.123	36.175
	(33.966)	(46.870)	(36.812)	(30.751)	(25.966)
Observations	618	360	706	1012	1279
Control Mean	11.312	9.924	9.978	9.577	14.205
Panel B: Number of Documents (10 km grid)					
In Sokoto	10.758	17.868	12.229	10.316*	7.550
	(6.957)	(13.375)	(8.005)	(6.028)	(4.877)
Observations	3566	2182	4292	6157	7737
Control Mean	1.452	1.167	1.418	1.595	2.554
Panel C: Number of Documents per 1000 Residents (Division-Level)					
In Sokoto	-0.248	28.079	1.920	9.144**	2.216
	(11.646)	(96.063)	(5.402)	(3.587)	(3.923)
Observations	10	13	32	45	63
Control Mean	0.147	0.147	0.740	0.593	0.572
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates for the number of documents in 25 km grid cells (Panel A), in 10 km grid cells (Panel B), and per 1,000 residents at the division level (Panel C). Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. All specifications include geographic controls and border-segment fixed effects. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D10: RD Results with Inverse Document-Count Weights

	Bandwidth			
	(1) 50km	(2) 100km	(3) 150km	(4) 200km
Panel A: Negative Sentiment				
In Sokoto	-0.201	0.022	0.168*	0.184**
	(0.160)	(0.121)	(0.099)	(0.088)
Observations	6,016	9,278	16,810	24,609
Control Mean	0.246	0.256	0.313	0.362
Panel B: Formal Tax Complaint				
In Sokoto	0.061	0.229**	0.227***	0.202***
	(0.147)	(0.106)	(0.081)	(0.071)
Observations	6,016	9,278	16,810	24,609
Control Mean	0.071	0.081	0.105	0.151
Geographic Controls	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) weighting each document by the inverse of its location's document count, so that every location receives equal weight. Columns use fixed bandwidths of 50, 100, 150, and 200 km. Outcomes, unit of observation, controls, and fixed effects are as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D11: RD Results with Inverse 1952 Population Weights

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.146***	0.068	0.200***	0.138***	0.127***
	(0.043)	(0.045)	(0.042)	(0.042)	(0.042)
Observations	8521	6016	9278	16810	24609
Control Mean	0.227	0.220	0.228	0.259	0.315
Panel B: Formal Tax Complaint					
In Sokoto	0.123***	0.104***	0.186***	0.076***	0.081***
	(0.009)	(0.008)	(0.015)	(0.014)	(0.019)
Observations	6336	6016	9278	16810	24609
Control Mean	0.047	0.046	0.056	0.073	0.123
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) weighting each document by inverse division-level population from the 1952 Census, so that every person receives equal weight. Outcomes, unit of observation, bandwidths, controls, and fixed effects are as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D12: Donut RD: Excluding a 10 km Buffer Around the Border

	Bandwidth			
	10km Buffer			
	(1)	(2)	(3)	(4)
	50km	100km	150km	200km
Panel A: Negative Sentiment				
In Sokoto	0.070	0.218**	0.253***	0.226***
	(0.505)	(0.098)	(0.062)	(0.061)
Observations	3084	6346	13878	21677
Control Mean	0.256	0.261	0.322	0.370
Panel B: Formal Tax Complaint				
In Sokoto	-0.435	0.129	0.199***	0.180***
	(0.374)	(0.095)	(0.063)	(0.057)
Observations	3084	6346	13878	21677
Control Mean	0.086	0.088	0.112	0.157
Geographic Controls	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) excluding all locations within a 10 km buffer of the Sokoto Caliphate boundary. Columns use fixed bandwidths of 50, 100, 150, and 200 km. Outcomes, unit of observation, controls, and fixed effects are as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Panel B.

Table D13: Border Specificity: Kanem–Bornu

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Excluding Kanem-Bornu Observations					
Negative Sentiment	0.120***	0.086**	0.182***	0.161***	0.166***
	(0.040)	(0.043)	(0.041)	(0.038)	(0.037)
Observations	7,263	5,926	8,408	15,377	23,171
Control Mean	0.237	0.239	0.236	0.319	0.378
Formal Tax Complaint	0.122***	0.139***	0.154***	0.082***	0.085***
	(0.020)	(0.013)	(0.021)	(0.026)	(0.029)
Observations	6,029	5,926	8,408	15,377	23,171
Control Mean	0.074	0.075	0.072	0.111	0.166
Panel B: Kanem-Bornu Treated as Part of Sokoto Caliphate					
Negative Sentiment	0.152***	0.128***	0.177***	0.160***	0.146***
	(0.041)	(0.043)	(0.041)	(0.038)	(0.037)
Observations	7,364	6,026	9,288	16,820	24,619
Control Mean	0.237	0.239	0.236	0.319	0.378
Formal Tax Complaint	0.123***	0.140***	0.136***	0.076***	0.054*
	(0.022)	(0.013)	(0.021)	(0.025)	(0.029)
Observations	6,555	6,026	9,288	16,820	24,619
Control Mean	0.078	0.075	0.072	0.111	0.166
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) for negative tax sentiment and formal tax complaints, excluding Kanem–Bornu observations (Panel A) and treating Kanem–Bornu as part of the Sokoto Caliphate (Panel B). Bandwidths, controls, and fixed effects are as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.10 Functional Form

Appendix Table D15 replaces the local linear function of distance with a quadratic; Appendix Table D16 uses an RD polynomial linear in latitude and longitude.

Table D14: Border Specificity: Jos Plateau

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Excluding Jos Plateau Observations					
Negative Sentiment	0.131***	0.172***	0.209***	0.153***	0.138***
	(0.040)	(0.043)	(0.039)	(0.034)	(0.035)
Observations	7,290	6,026	9,071	16,292	21,908
Control Mean	0.244	0.249	0.257	0.314	0.362
Formal Tax Complaint	0.187***	0.199***	0.216***	0.093***	0.073**
	(0.018)	(0.014)	(0.024)	(0.023)	(0.028)
Observations	6,126	6,026	9,071	16,292	21,908
Control Mean	0.075	0.075	0.083	0.106	0.151
Panel B: Modified Sokoto Boundary with Jos Plateau Carved Out					
Negative Sentiment	0.125**	0.084	0.131***	0.060*	0.079**
	(0.051)	(0.056)	(0.048)	(0.037)	(0.033)
Observations	10,123	9,423	12,600	20,513	27,100
Control Mean	0.278	0.280	0.275	0.307	0.341
Formal Tax Complaint	0.041	0.064*	0.113***	0.081***	0.100***
	(0.034)	(0.034)	(0.031)	(0.028)	(0.027)
Observations	9,437	9,423	12,600	20,513	27,100
Control Mean	0.095	0.095	0.093	0.106	0.132
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) for negative tax sentiment and formal tax complaints, excluding Jos Plateau observations (Panel A) and using a modified Sokoto boundary with the Jos Plateau carved out (Panel B). Bandwidths, controls, and fixed effects are as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D15: RD Polynomial Quadratic in Distance to the Boundary

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.137***	0.236***	0.159***	0.130***	0.131***
	(0.042)	(0.050)	(0.043)	(0.041)	(0.042)
Observations	10401	6016	9278	16810	24609
Control Mean	0.297	0.246	0.256	0.313	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.187***	0.370***	0.121***	0.147***	0.133***
	(0.014)	(0.018)	(0.018)	(0.024)	(0.033)
Observations	7375	6016	9278	16810	24609
Control Mean	0.073	0.071	0.081	0.105	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) with a quadratic polynomial in distance to the boundary, estimated separately on each side. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D16: RD Polynomial Linear in Latitude and Longitude

	Bandwidth			
	(1) 50km	(2) 100km	(3) 150km	(4) 200km
Panel A: Negative Sentiment				
In Sokoto	0.291***	0.071*	0.073**	0.053
	(0.066)	(0.040)	(0.033)	(0.034)
Observations	6016	9278	16810	24609
Control Mean	0.246	0.256	0.313	0.362
Panel B: Formal Tax Complaint				
In Sokoto	0.279***	0.072**	0.038	0.041
	(0.072)	(0.031)	(0.026)	(0.031)
Observations	6016	9278	16810	24609
Control Mean	0.071	0.081	0.105	0.151
Geographic Controls	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) with an RD polynomial linear in latitude and longitude. Columns use fixed bandwidths of 50, 100, 150, and 200 km. Outcomes, unit of observation, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.11 Inference

Appendix Table D17 clusters standard errors at the province level; Appendix Table D18 uses spatially correlated standard errors following Conley (1999) with a 25-kilometer cutoff.

Table D17: RD Results with Province-Clustered Standard Errors

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.119***	0.165***	0.194***	0.157***	0.143***
	(0.040)	(0.022)	(0.032)	(0.023)	(0.021)
Observations	6590	6016	9278	16810	24609
Control Mean	0.248	0.246	0.256	0.313	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.153***	0.190***	0.199***	0.097***	0.081***
	(0.025)	(0.017)	(0.031)	(0.023)	(0.026)
Observations	6015	6016	9278	16810	24609
Control Mean	0.071	0.071	0.081	0.105	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1). Outcomes, unit of observation, bandwidths, controls, and fixed effects are as in Table 1. Standard errors, clustered at the province level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

E Within-Caliphate Exposure

This appendix presents the maps underlying the within-Caliphate exposure measures of Section 5.3 and the ribat placebo results.

Appendix Figure E1 shows the movements of Usman dan Fodio during 1804–1817, from Last (1967). Appendix Figure E2 shows the elevation-based least-cost path connecting the locations on this route, which defines the consolidation corridor used in Panels A and B of

Table D18: RD Results with Conley Standard Errors

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.135*	0.261***	0.161**	0.163***	0.103**
	(0.081)	(0.101)	(0.071)	(0.045)	(0.049)
Observations	7739	6016	9278	16810	24609
Control Mean	0.240	0.246	0.256	0.313	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.183***	0.265***	0.172**	0.085*	0.056
	(0.068)	(0.085)	(0.069)	(0.044)	(0.045)
Observations	6121	6016	9278	16810	24609
Control Mean	0.070	0.071	0.081	0.105	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1). Outcomes, unit of observation, bandwidths, controls, and fixed effects are as in Table 1. Standard errors, spatially correlated following Conley (1999) with a 25 km cutoff, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D19: More Contested Legitimacy in Sokoto

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Contested Legitimacy (all observations)					
In Sokoto	0.083***	0.189***	0.194***	0.101***	0.076***
	(0.015)	(0.014)	(0.019)	(0.017)	(0.017)
Observations	6613	5778	8931	15995	22714
Control Mean	0.053	0.055	0.062	0.068	0.079
Panel B: Contested Legitimacy (high confidence only)					
In Sokoto	0.123***	0.307***	0.270***	0.134***	0.099***
	(0.022)	(0.015)	(0.022)	(0.019)	(0.021)
Observations	5444	4769	7382	12570	17730
Control Mean	0.058	0.060	0.069	0.081	0.098
Panel C: Contested Legitimacy (very high confidence only)					
In Sokoto	0.205***	0.284***	0.400***	0.190***	0.185***
	(0.039)	(0.014)	(0.057)	(0.049)	(0.042)
Observations	1617	1518	2269	3750	4992
Control Mean	0.054	0.049	0.073	0.098	0.131
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) for an indicator of contested legitimacy, among all observations (Panel A), among documents whose legitimacy classification has high confidence (Panel B), and very high confidence (Panel C). The unit of observation is a document. Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. All specifications include geographic controls, archive fixed effects, and border-segment fixed effects. Standard errors, clustered by document location, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 2. Appendix Figure E3 shows the major locations of the Sokoto Caliphate, including the emirates and the territories under the Vizier, which define the proximity measure used in Panel C of Table 2.

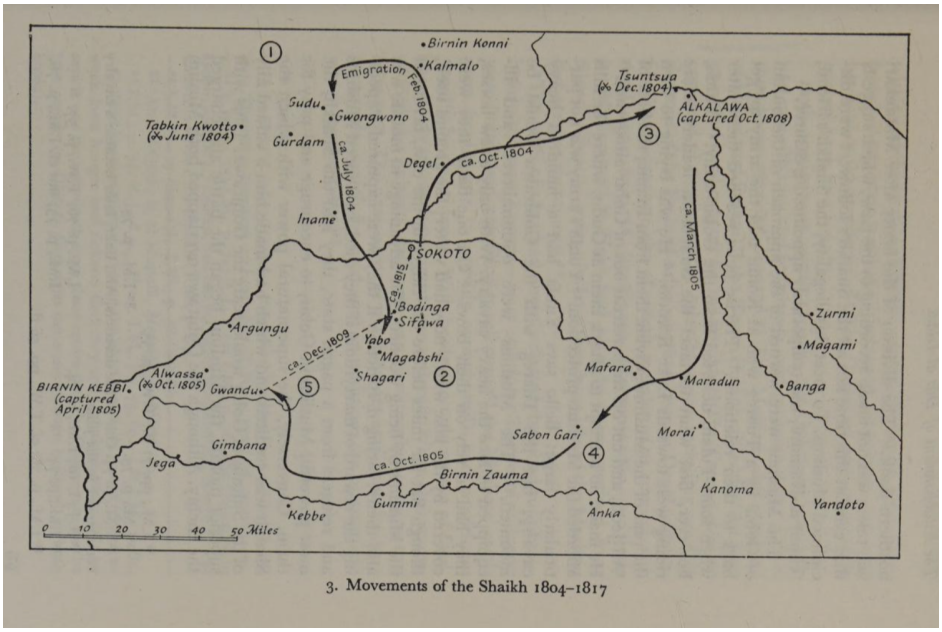


Figure E1: Movements of Usman dan Fodio, 1804–1817

Notes: Source: Last (1967).

Appendix Figure E4 shows the royal ribats of 1839–1845, from Last (1967), which capture Sokoto's military and defensive footprint; Appendix Table E1 reports the placebo results discussed in Section 5.3.

F Additional Mechanism Evidence

This appendix presents additional evidence on the fiscal legitimacy channel of Section 6.

Appendix Table F1 reports the within-region test of Section 6.1, relating the document-level outcomes to the division-level Muslim population share from the 1952 Census, within and outside the former Caliphate; Appendix Figure G4 maps the underlying population shares.

The remaining exhibits concern the direct LLM-based measure of contested tax legitimacy (Appendix B.9 reproduces the prompt; Appendix C its validation). Appendix Figure F1 shows the words most distinctive of documents coded as contesting legitimacy. Appendix Table F2 shows that contested legitimacy is strongly associated with negative sentiment and formal complaints at the document level, with larger associations in the samples restricted to

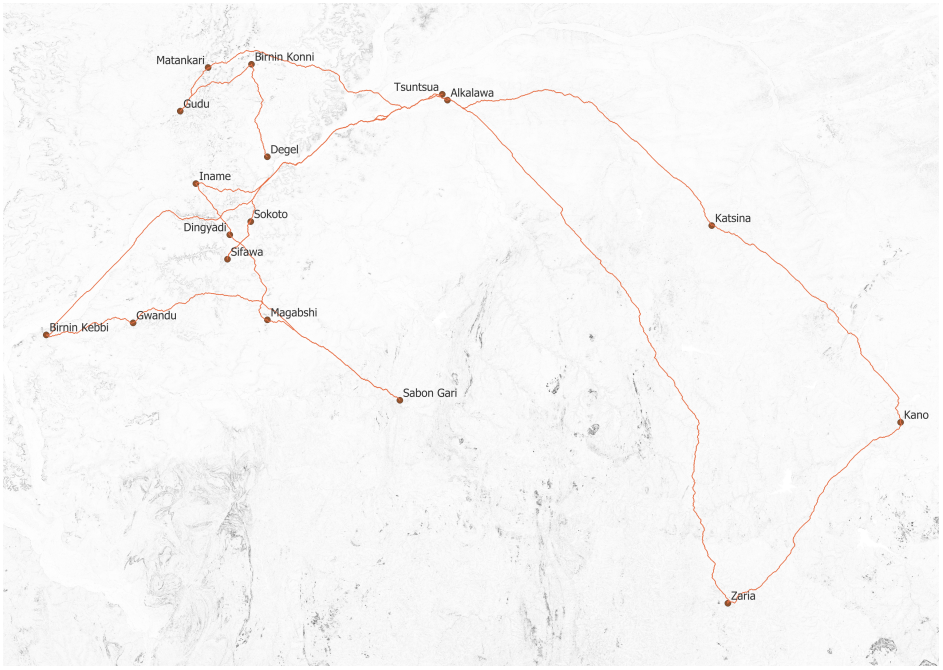


Figure E2: Movements of Usman dan Fodio: Least-Cost Path

Notes: The figure shows the elevation-based least-cost path connecting the locations of Usman dan Fodio's movements listed and mapped in Last (1967).

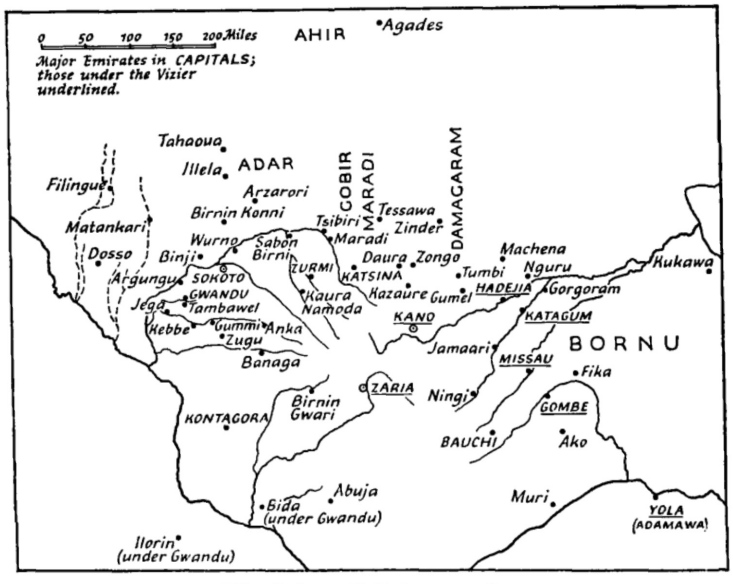


Figure E3: Major Locations in the Sokoto Caliphate

Notes: Source: Last (1967), *The Sokoto Caliphate*. All administrative centres (major emirates, vizier seats, and settlements) were digitized from this map to construct the proximity variable used in Panel C of Table 2.

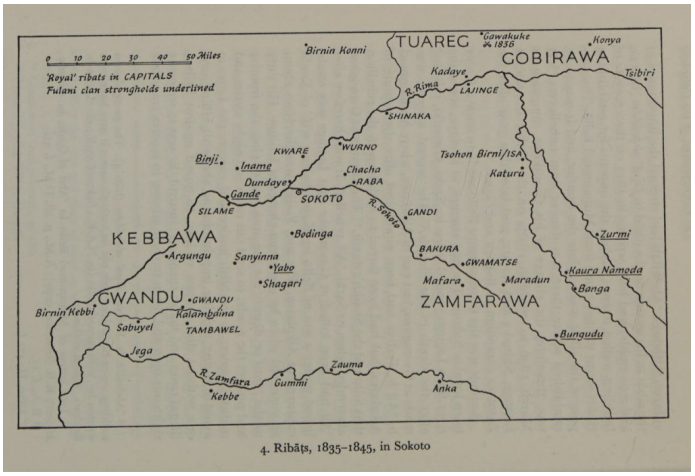


Figure E4: Royal Ribats, 1839–1845

Notes: Source: Last (1967).

Table E1: Placebo: Proximity to Royal Ribats

	Neg. Sentiment	Formal Tax Complaint	Cont. Legitimacy	Agg. Relational Index
	(1)	(2)	(3)	(4)
Royal Ribats	-0.033*	-0.033*	-0.023	-0.072
	(0.026)	(0.019)	(0.019)	(0.054)
Obs.	18,172	18,172	17,906	17,906
Control mean	0.353	0.176	0.094	
Geo controls	Yes	Yes	Yes	Yes
Pre-Sokoto controls	Yes	Yes	Yes	Yes
Archive FE	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: The table relates colonial-era outcomes to proximity to the royal ribats of 1839–1845, mapped in Last (1967), within the former Caliphate. The unit of observation is a document; outcomes are indicators for negative tax sentiment, a formal tax complaint, contested legitimacy (introduced in Section 6), and an aggregate relational index. All specifications include geographic controls, pre-Sokoto controls, and archive fixed effects. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table F1: Muslim Population Share and Tax Resistance, Within and Outside Sokoto

	Negative sentiment (1)	Formal complaint (2)
Panel A: Within Sokoto		
Percent Muslims	0.002	0.003***
	(0.001)	(0.001)
Observations	17,149	17,149
Panel B: Outside Sokoto		
Percent Muslims	-0.006*	-0.005***
	(0.003)	(0.002)
Observations	11,540	11,540

Notes: The table relates the document-level outcomes to the division-level Muslim population share from the 1952 Census, within the former Sokoto Caliphate (Panel A) and outside it (Panel B). The unit of observation is a document; outcomes are indicators for negative tax sentiment, column (1), and a formal tax complaint, column (2). Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

high- and very-high-confidence legitimacy classifications; Table D19 in the main text reports the RD estimates.

Table F2: Contested Legitimacy, Negative Sentiment, and Formal Complaints

	All Observations		Legitimacy Score Confidence: High		Legitimacy Score Confidence: Very High	
	Negative sentiment (1)	Formal complaint (2)	Negative sentiment (3)	Formal complaint (4)	Negative sentiment (5)	Formal complaint (6)
Contested legitimacy	0.711***	0.677***	0.742***	0.703***	0.843***	0.867***
	(0.004)	(0.010)	(0.004)	(0.010)	(0.006)	(0.011)
Mean dep. var.	0.342	0.162	0.329	0.162	0.282	0.183
Mean indep. var.	0.087	0.087	0.107	0.107	0.153	0.153
Observations	21,902	21,902	17,279	17,279	4,847	4,847

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table relates the document-level outcomes, indicators for negative tax sentiment and a formal tax complaint, to the indicator for contested legitimacy. The unit of observation is a document. Columns (1)–(2) use all observations; columns (3)–(4) restrict to documents whose legitimacy classification has high confidence; columns (5)–(6) restrict to very high confidence. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G Alternative Explanations, Supporting Evidence

This appendix presents the maps, figures, and tables supporting the examination of alternative explanations in Section 6.3.

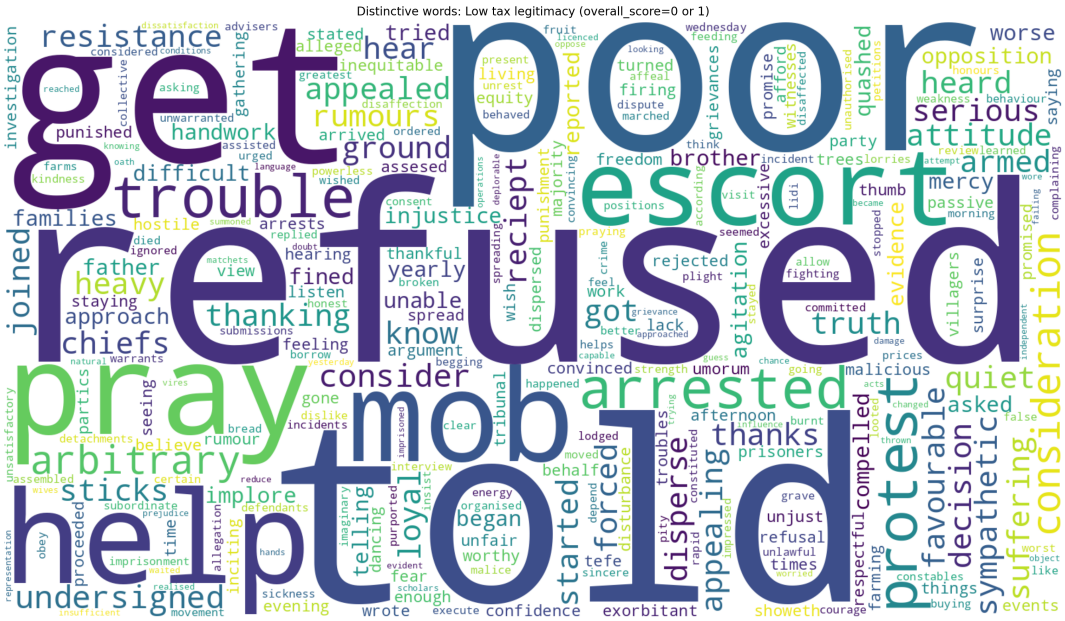
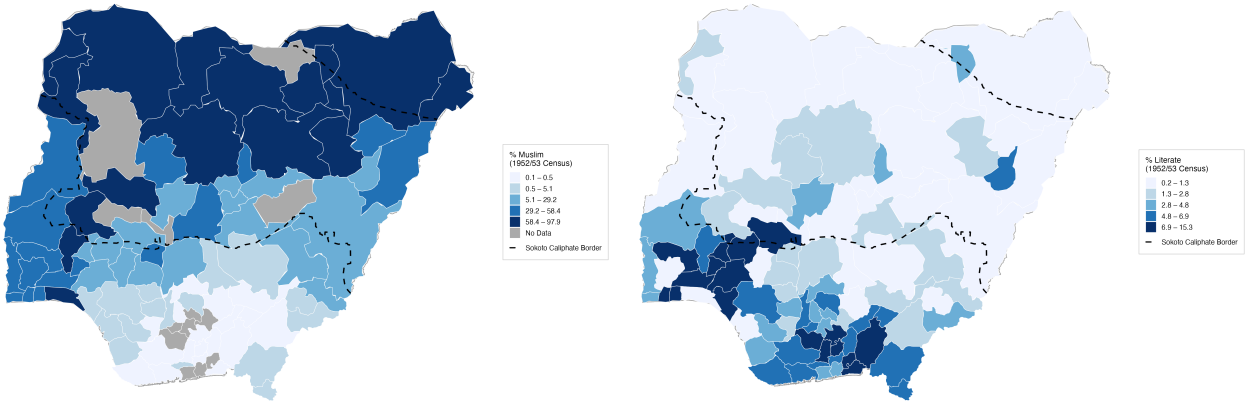


Figure F1: Word Cloud: Contested Legitimacy

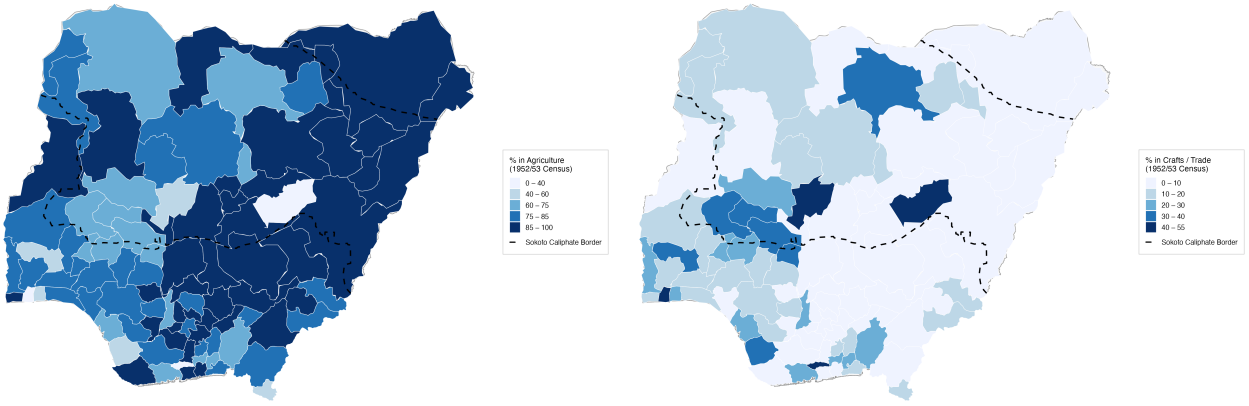
G.1 Demography

Appendix Figure G1 maps Muslim population shares and literacy from the 1952/53 Census; Appendix Figures G2 and G3 map occupation shares from the same census; and Appendix Figure G4 maps division-level Muslim population shares and Roman-script literacy from the 1952 Census around the Sokoto boundary. Appendix Tables G1 and G2 show that the main results are robust to controlling for literacy and for occupation shares, respectively.



(a) Muslim Population (b) Literacy

Figure G1: Religion and Literacy (1952/53 Census)



(a) Agriculture (b) Crafts / Trade

Figure G2: Occupation Shares (1952/53 Census)

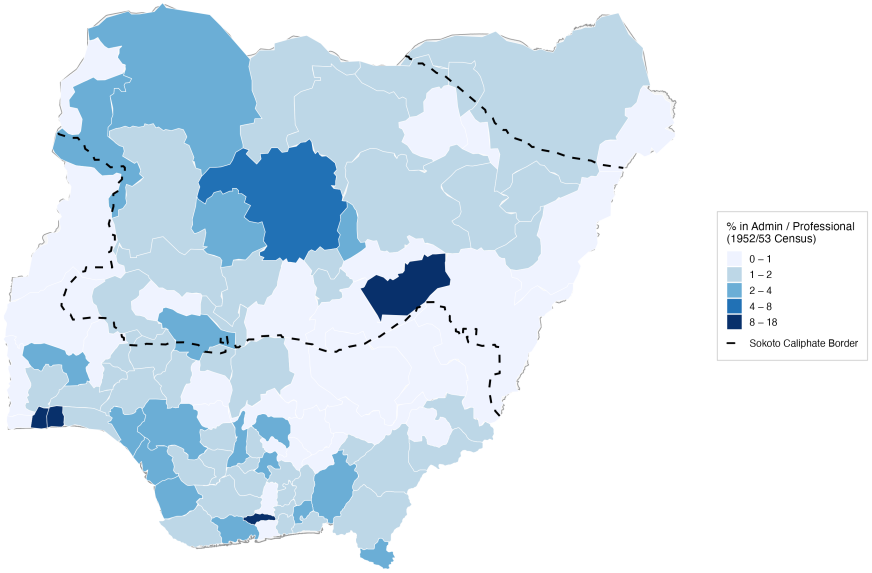
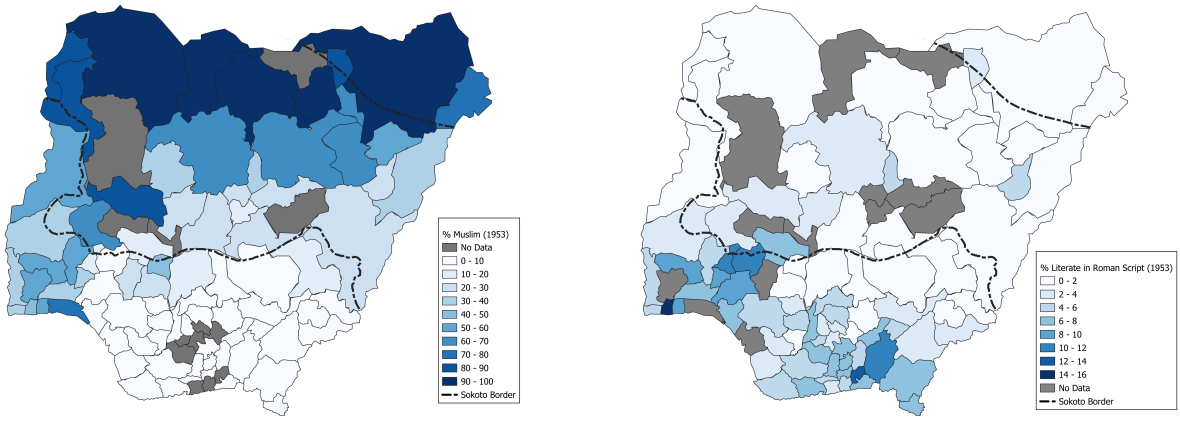


Figure G3: Occupation Shares (1952/53 Census), Administration



(a) Muslim Population (%, Division-Level 1952 Census)

(b) Roman-Script Literacy (%, Division-Level 1952 Census)

Figure G4: Religion and Literacy by Division (1952 Census)

Table G1: Robustness: Literacy Controls

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.108***	0.173***	0.177***	0.144***	0.154***
	(0.041)	(0.043)	(0.040)	(0.034)	(0.034)
Observations	8,531	6,026	9,288	16,820	24,619
Control Mean	0.257	0.249	0.257	0.314	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.133***	0.199***	0.180***	0.079***	0.092***
	(0.027)	(0.014)	(0.026)	(0.024)	(0.030)
Observations	7,364	6,026	9,288	16,820	24,619
Control Mean	0.076	0.075	0.083	0.106	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Literacy Control	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses, clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) adding a control for literacy. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors, clustered at the city level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G2: Robustness: Occupation Share Controls

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.050	0.073*	0.202***	0.137***	0.118***
	(0.040)	(0.043)	(0.039)	(0.034)	(0.033)
Observations	6,558	6,026	9,288	16,820	24,619
Control Mean	0.249	0.249	0.257	0.314	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.108***	0.151***	0.212***	0.070***	0.050*
	(0.023)	(0.013)	(0.023)	(0.022)	(0.027)
Observations	6,602	6,026	9,288	16,820	24,619
Control Mean	0.080	0.075	0.083	0.106	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Occupation Shares	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) adding controls for occupation shares. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.2 Tax Instruments

Appendix Table G3 restricts the sample to the years after 1927, when tax instruments were essentially identical across protectorates.

Table G3: Post-1927 Sample

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.182***	0.226***	0.249***	0.129***	0.114***
	(0.042)	(0.048)	(0.042)	(0.039)	(0.037)
Observations	7,768	5,588	8,490	15,406	21,786
Control Mean	0.235	0.229	0.236	0.274	0.294
Panel B: Formal Tax Complaint					
In Sokoto	0.169***	0.204***	0.224***	0.105***	0.091***
	(0.013)	(0.010)	(0.022)	(0.025)	(0.027)
Observations	6,009	5,588	8,490	15,406	21,786
Control Mean	0.071	0.062	0.075	0.092	0.104
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Historical Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses, clustered at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) restricting the sample to documents dated after 1927, when tax instruments were essentially identical across protectorates. Outcomes, unit of observation, and bandwidths are as in Table 1; specifications include geographic controls, historical controls, archive fixed effects, and border-segment fixed effects. Standard errors, clustered at the city level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

G.3 Tax Burden

Appendix Figure G5 shows the regional split of colonial tax revenue: the bulk of colonial revenue was raised in the South.

G.4 Tax Administration

Appendix Table G4 reports the RD estimates for bad administration and corruption discussed in Section 6.3. Appendix Figure G6 maps civil servants per 1,000 population from

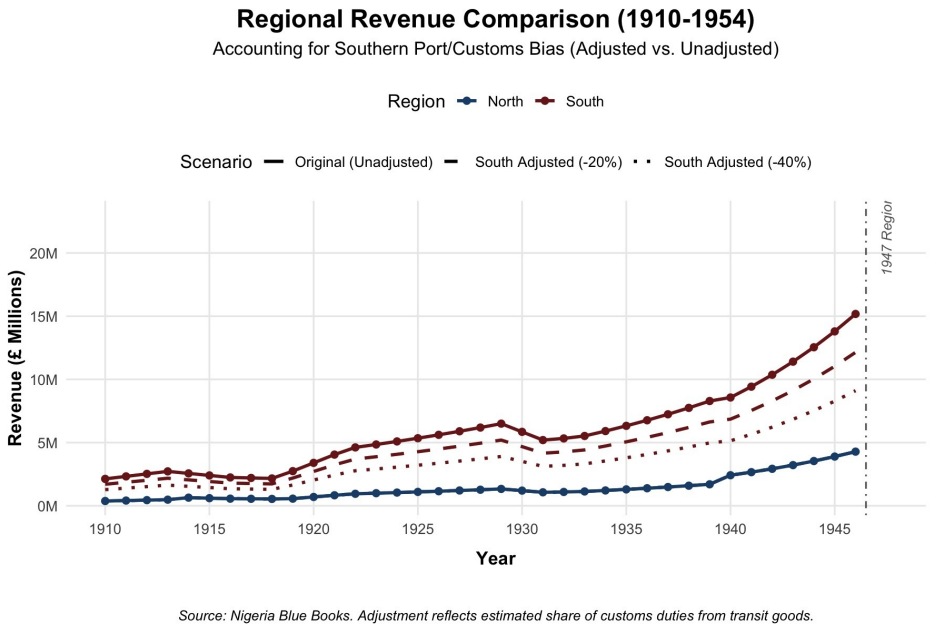


Figure G5: Tax Revenue: Northern vs. Southern Protectorate

Notes: The figure shows aggregate colonial tax revenue in the Northern and Southern Protectorates. Data: *Blue Books* and *Reports from the Committee of Taxation.*

the 1931 Census, and Appendix Table G5 shows that the main results are robust to controlling for them.

G.5 Tax Enforcement

Appendix Table G6 reports the RD estimates for state coercion or punishment and tax evasion discussed in Section 6.3.

G.6 Public Goods Provision

Appendix Figures G7, G8, and G9 show the georeferenced maps of colonial infrastructure: roads and railroads in 1924, medical facilities in 1934, and airfields, ferries, and post office, telegraph, and telephone locations in 1949. Appendix Table G7 shows that the main results are unaffected by controlling for colonial public goods.

H Tax Revolts and Riots

This appendix describes the construction of the tax revolt and riot data referenced in Section 5.1, and reports the associated results.

Table G4: Tax Administration at the Boundary

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Bad Administration					
In Sokoto	-0.091***	-0.113***	-0.085***	-0.102***	-0.101***
	(0.030)	(0.032)	(0.029)	(0.030)	(0.029)
Observations	8195	5778	8931	15995	22714
Panel B: Corruption					
In Sokoto	-0.006	-0.003	0.006	0.024***	0.030***
	(0.005)	(0.005)	(0.006)	(0.008)	(0.009)
Observations	7405	5778	8931	15995	22714
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Historical Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates at the Sokoto Caliphate boundary for indicators that a document describes bad administration (Panel A) or corruption (Panel B). The unit of observation is a document. Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. All specifications include geographic controls, historical controls, and border-segment fixed effects. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

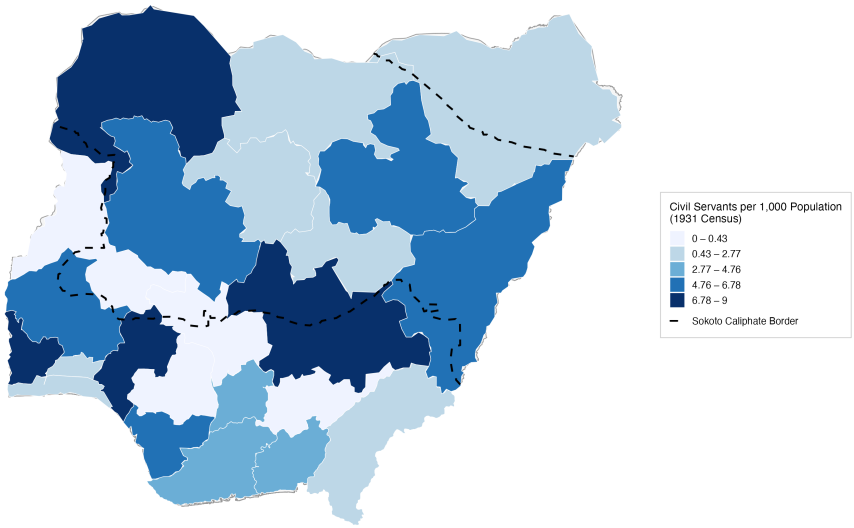


Figure G6: Civil Servants per 1,000 Population (1931 Census)

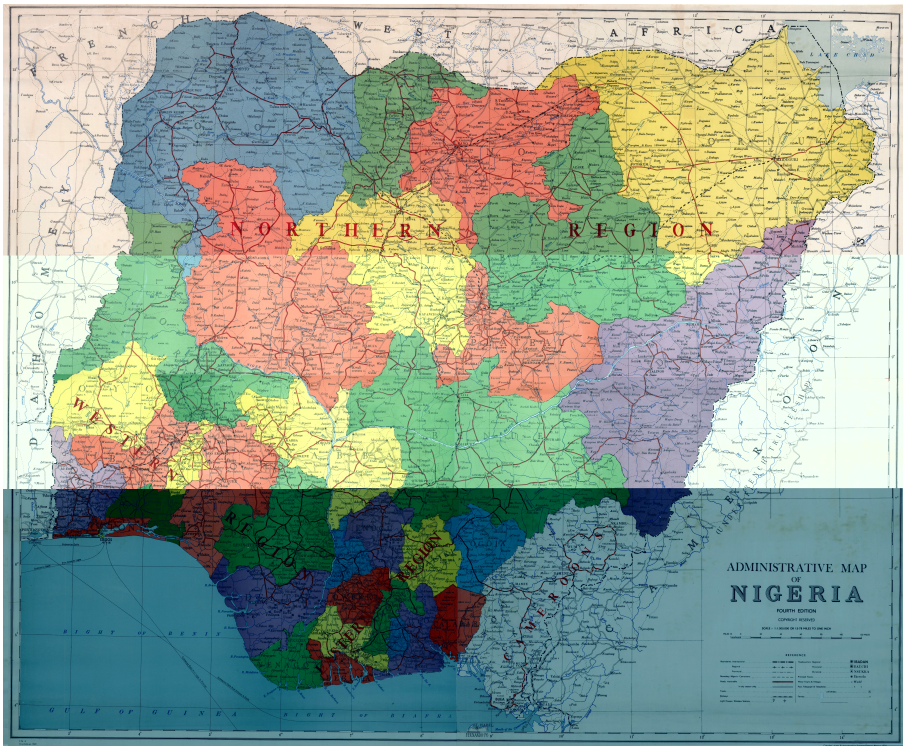


Figure G7: Roads and Railroads, 1924

Notes: Source: Surveyor-General Nigeria, 1924; Okoye et al. (2019).

Table G5: Robustness: Civil Servants per 1,000 Control

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.183***	0.169***	0.180***	0.148***	0.142***
	(0.035)	(0.043)	(0.040)	(0.035)	(0.034)
Observations	9,556	6,026	9,288	16,820	24,619
Control Mean	0.257	0.249	0.257	0.314	0.362
Panel B: Formal Tax Complaint					
In Sokoto	0.132***	0.197***	0.184***	0.085***	0.074**
	(0.029)	(0.013)	(0.027)	(0.024)	(0.030)
Observations	7,270	6,026	9,288	16,820	24,619
Control Mean	0.076	0.075	0.083	0.106	0.151
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes
Civil Servants per 1,000	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table reports spatial RD estimates of equation (1) adding a control for civil servants per 1,000 population from the 1931 Census. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table G6: No Difference in Enforcement Efforts

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: State Coercion or Punishment					
In Sokoto	-0.011	0.035***	-0.034*	-0.026	-0.026
	(0.014)	(0.012)	(0.020)	(0.033)	(0.039)
Observations	5645	5778	8931	15995	22714
Panel B: Tax Evasion					
In Sokoto	-0.001	0.014	0.004	-0.007	-0.002
	(0.019)	(0.020)	(0.019)	(0.016)	(0.015)
Observations	8193	5778	8931	15995	22714
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Historical Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates at the Sokoto Caliphate boundary for indicators that a document describes state coercion or punishment (Panel A) or tax evasion (Panel B). The unit of observation is a document. Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. All specifications include geographic controls, historical controls, and border-segment fixed effects. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

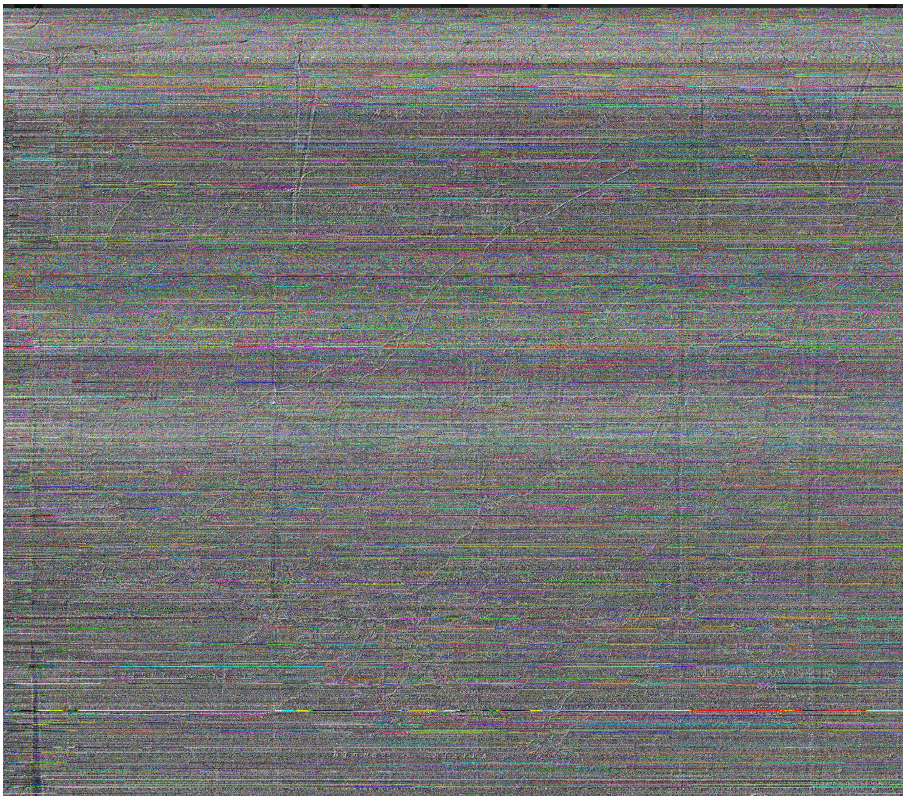


Figure G8: Medical Facilities, 1934

Notes: Source: Nigeria Surveys, 1934/1935.



Figure G9: Airfields, Ferries, and Post Office/Telegraph/Telephone, 1949

Notes: Source: Survey Department Nigeria, 1949.

Table G7: Robustness: Colonial Public Goods Controls

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Negative Sentiment					
In Sokoto	0.179***	0.131***	0.282***	0.201***	0.131***
	(0.040)	(0.044)	(0.039)	(0.034)	(0.033)
Observations	7138	5774	8904	15874	22590
Control Mean	0.251	0.233	0.251	0.284	0.300
Panel B: Formal Tax Complaint					
In Sokoto	0.145***	0.152***	0.318***	0.206***	0.123***
	(0.012)	(0.007)	(0.018)	(0.019)	(0.018)
Observations	5788	5774	8904	15874	22590
Control Mean	0.055	0.055	0.075	0.091	0.103
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Archive FEs	Yes	Yes	Yes	Yes	Yes
Public Good Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates of equation (1) adding controls for colonial public goods. Outcomes, unit of observation, bandwidths, controls, and fixed effects are otherwise as in Table 1. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

H.1 Data Construction

We combine two sources of information. The first is the colonial tax documents themselves: we use an LLM-based procedure to identify tax revolts and riots, where *revolts* are organized, large scale challenges to state authority centered on taxation, and *riots* are localized outbreaks of violence triggered by a tax-related event. Because the number of documents with a revolt or riot flagged is small ($N = 122$), we verified them individually and kept only those deemed valid ($N = 38$). The second source is the historical literature: 62 additional tax protests and riots were identified and geolocated from historical sources.

The final dataset contains 100 protests and riots, of which 22 to 73 fall within the spatial RD region, depending on the bandwidth (50 to 200 km). We construct an indicator for, and the number of, revolts and riots in 25 by 25 km grid cells.

H.2 Historical Sources

The historical events are drawn from the following sources. Books: Falola (2009), *Colonialism and Violence in Nigeria*; Afigbo (1972), *The Warrant Chiefs: Indirect Rule in Southeastern Nigeria*; Thomas (2012), *Violence and Colonial Order: Police, Workers, and Protests in the European Colonial Empires, 1918–1940*; **?**, *An Economic History of West Africa*; and Usman (1987), *The Manipulation of Religion in Nigeria.* Articles: Afolabi (2023) on tax revolts in Yorubaland, 1925–1955, and Byfield (2003) on the Egba women's tax revolt. Colonial reports: the Reports of the Native Courts (Eastern Region) Commission of Inquiry (1950), the Reports of the Native Courts (Northern Provinces) Commission of Inquiry (1952), the Aba Commission of Inquiry (1929), and the Report of a Committee on Trade and Taxation for British West Africa (1929).

H.3 Results

Appendix Table H1 shows that the protest and riot data validate the document-level outcomes: tax protests and riots are positively and significantly associated with both negative sentiment and formal tax complaints. Appendix Table H2 reports grid-level RD estimates at the Sokoto boundary for the protest indicator and the number of protests; none of the estimates is statistically significant.

Table H1: Tax Protests and Riots, Negative Sentiment, and Formal Complaints

	Tax Protests/Riots	
	(1)	(2)
Negative Sentiment	0.007***	
	(0.001)	
Formal Tax Complaint		0.004***
		(0.001)
Mean dep. var.	0.003	0.003
Observations	32,166	32,166

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Each column reports a regression of an indicator for a tax protest or riot on the outcome named in the column: negative tax sentiment in column (1) and a formal tax complaint in column (2). Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table H2: Effects on Tax Protests and Riots (Grid-Cell Level)

	Bandwidth				
	(1) MSE-Optimal	(2) 50km	(3) 100km	(4) 150km	(5) 200km
Panel A: Protest/Riot Indicator					
In Sokoto	-0.041	-0.142	-0.069	-0.044	-0.039
	(0.041)	(0.092)	(0.059)	(0.045)	(0.038)
Control Mean	0.024	0.041	0.025	0.026	0.029
Observations	770	360	706	1012	1279
Panel B: Number of Protests/Riots					
In Sokoto	-0.034	-0.148	-0.059	-0.032	-0.028
	(0.071)	(0.111)	(0.076)	(0.066)	(0.059)
Control Mean	0.039	0.053	0.035	0.035	0.052
Observations	629	360	706	1012	1279
Geographic Controls	Yes	Yes	Yes	Yes	Yes
Border Segment FEs	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

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

Notes: The table reports spatial RD estimates at the Sokoto Caliphate boundary for an indicator of any tax protest or riot (Panel A) and the number of tax protests or riots (Panel B). The unit of observation is a 25 by 25 km grid cell. Column (1) uses the MSE-optimal bandwidth; columns (2)–(5) use fixed bandwidths of 50, 100, 150, and 200 km. All specifications include geographic controls and border-segment fixed effects. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.