DISRUPTING STATE LEGITIMACY: Evidence from 50,000 Colonial Documents*

Augustin Bergeron and Abdoulaye Ndiaye

August 27, 2026

Abstract

Does state legitimacy shape citizens' resistance to state demands, and thereby the state's capacity to implement policy? We study this question through a transformation of fiscal authority in colonial Nigeria. Before British conquest, taxation in the Sokoto Caliphate was embedded in an Islamic fiscal order defining citizens' obligations and rulers' authority to tax. Colonial rule preserved much of the machinery of collection while disrupting this institutional basis: religiously sanctioned obligations were curtailed, religious distinctions in tax liability removed, and previously rejected levies reintroduced. Nearby areas were incorporated into the same colonial state, but colonial taxation did not displace a comparable fiscal 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 more negative tax sentiment and formal complaints where colonial taxation displaced Sokoto's fiscal order. The form of resistance matches predictions specific to this disruption: additional complaints are disproportionately associated with Muslims and partly concentrate on a tax explicitly rejected under Sokoto rule. A direct text-based measure likewise shows greater contestation of the state's right to tax. The results are reinforced by within-Caliphate variation in exposure to Sokoto institutions and are not explained by differences in colonial tax policy, administration, enforcement, public goods, or observable population composition. The findings identify legitimacy as a distinct constraint on state capacity: the effectiveness of new institutions depends not only on what states can administer or enforce, but also on the sources of authority and obligation they displace.

JEL Codes: D74, H20, H26, N47, Z12.

Keywords: fiscal legitimacy; state capacity; taxation; colonial institutions; Sokoto Caliphate; Nigeria; large language models; regression discontinuity.

*Bergeron (abergeron@fas.harvard.edu) - Harvard University; 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, 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, and Jonathan Weigel 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). Most existing work studies the state's ability to design, administer, and enforce policy (Pomeranz 2015; Finan et al. 2017; Naritomi 2019; Basri et al. 2021; Besley et al. 2021). Yet effective policy may depend on more than the state's administrative and coercive capabilities. Citizens may comply with a state demand because they recognize an obligation to do so (Levi 1988; Tyler 2006), or resist the same demand because they reject the authority behind it. This paper asks how state legitimacy shapes citizens' responses to state demands, and thus the state's effective capacity to implement policy.

We study this question through taxation, where the state's claim on citizens is especially explicit. Our setting is a major transformation of fiscal authority in colonial Nigeria. Before British conquest, much of northern Nigeria belonged to the Sokoto Caliphate (1804–1903), a centralized Islamic polity founded by Usman dan Fodio's reform movement. Taxation in the Caliphate was embedded in an Islamic fiscal order: obligations were defined through Islamic law, administered through emirate institutions, and understood in part as religious duties (Last 1967; Hiskett 1973; Lovejoy 2016). These rules governed both what subjects were obliged to pay and what rulers were entitled to demand, so taxes viewed as inconsistent with Islamic law could themselves provoke opposition.

British conquest transformed this fiscal order without fully replacing the institutions through which taxes were collected. Colonial authorities continued to rely heavily on emirate institutions and local intermediaries, while changing the rules and authority underlying taxation. Official collection of *zakat* as a religious obligation was curtailed; religious distinctions governing liability for other taxes were weakened or removed; and *jangali*, a cattle tax rejected by Sokoto reformers on Islamic grounds, was reintroduced. For many Muslims, colonial rule therefore changed not only who claimed the authority to tax, but also which taxes they were required to pay. Elsewhere in Nigeria, British rule likewise expanded and reorganized taxation, but generally did not displace a comparable centralized, religiously grounded fiscal order.

This setting provides an unusual opportunity to isolate a source of legitimacy rooted in recognized authority and obligation rather than in state performance, which has been the focus of much of the literature on legitimacy (Levi 1988; Tyler 2006; Besley 2020; Besley et al. 2026). In most settings, these dimensions are difficult to separate because governments whose authority is more widely accepted may also pursue different policies, administer them more effectively, enforce them differently, or provide more public goods. Colonial Nigeria

instead holds the new fiscal regime comparatively fixed while varying the institutional order it displaced. Nearby communities were incorporated into the same colonial state and increasingly subjected to similar fiscal institutions, but in former Sokoto those institutions displaced a recognized religious fiscal order and altered Muslims' fiscal obligations in ways they generally did not elsewhere. This yields a simple prediction: if recognized fiscal authority matters for citizens' responses to taxation, resistance should be greater where colonial rule displaced Sokoto's fiscal order.

Testing this prediction requires observing resistance to taxation systematically. We therefore 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. The corpus contains the routine records of colonial fiscal administration, including correspondence, tax and financial records, administrative and investigative reports, petitions, and formal grievances. These documents allow us to observe how taxation was discussed and contested across locations and over much of the colonial period.

We use large language models (LLMs) to recover structured information from these records. Our two main outcomes capture complementary dimensions of tax resistance. The first measures 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. We also use LLMs to recover document locations and dates, standardize historical place names, and match them to historical maps.

Because LLMs may misread complex historical text or hallucinate, we validate the main outcomes against independent human judgments (Korinek 2023; Amirizaniani et al. 2024; Dell 2025). Five research assistants independently code a random sample of 100 documents for tax sentiment and formal tax complaints. We then conduct a Turing-style validation exercise (Turing 1950; Lagakos et al. 2025): each human coder and the LLM is benchmarked against the modal human classification, allowing us to ask whether the LLM performs within the range of human coders. For tax sentiment, we compare absolute distance from the modal human score; for formal complaints, agreement with the modal human classification. On both outcomes, the LLM falls within the range of human performance.

Our main empirical strategy exploits the former boundary of the Sokoto Caliphate in a spatial regression discontinuity design. Because each document is geolocated, we compare nearby locations on opposite sides of the historical boundary that were exposed to different pre-colonial fiscal institutions but subsequently incorporated into the same colonial state. We estimate local linear specifications in distance to the boundary, allowing slopes to differ across sides, and include border-segment and archive fixed effects together with a rich set

of geographic and pre-Sokoto controls. We report estimates at 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. While historical accounts emphasize military reach and ecological constraints in shaping the limits of Sokoto expansion (Kirk-Greene 1965; Last 1967; Lovejoy 2016), we find no discontinuities in a broad set of geographic and pre-Sokoto characteristics. The boundary also cuts through several ethnic homelands (Murdock 1967), and the estimates remain similar with ethnic-territory fixed effects. A separate concern is that the former Sokoto boundary may also have become a boundary between different colonial tax regimes. In practice, direct-tax institutions converged substantially across Nigeria, and the historical Sokoto frontier cut across the colonial protectorates. The results also remain similar when we restrict the sample to the Northern Protectorate, where the colonial fiscal framework is held much more tightly fixed.

We find substantially greater tax resistance inside the former Sokoto Caliphate. At the MSE-optimal bandwidth, documents from locations just inside the boundary are 9.3 percentage points more likely to express negative tax sentiment, approximately 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 difference therefore appears both in the broader language surrounding taxation and in a concrete form of resistance directed toward the colonial state, and 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. The first is that the result could reflect how the archive was assembled rather than differences in the content of the records. The three archives contribute different numbers and types of documents, and their composition may also vary over time. We therefore show that the estimates are robust to document-type fixed effects and with archive-by-decade and archive-by-year fixed effects, so that the comparison is made among similar records and within the same archive and time period. A related concern is uneven spatial coverage: some locations generate many more surviving documents than others, which could cause heavily represented places to receive disproportionate weight. We find no discontinuity in document counts at the Sokoto boundary, and the results remain significant when observations are reweighted to give equal weight to locations. They are also robust to reweighting by local population from the 1952 Census, addressing the possibility that more populous areas are mechanically more likely to generate surviving records.

The second concern is uncertainty over the location and historical interpretation of the

Sokoto boundary. Small geocoding errors matter most for observations closest to the frontier, yet the estimates remain similar in donut specifications that exclude these locations. Nor do the results depend on how historically ambiguous portions of the boundary are treated: they are robust to alternative treatment of areas adjacent to Kanem–Bornu, where some Islamic fiscal institutions were also present, and to excluding the Jos Plateau, whose incorporation into Sokoto is debated in the historical literature. The third concern is the econometric implementation of the spatial RD. The discontinuity is stable under alternative spatial functional forms, and the conclusions are unchanged when inference is based on province-level clustering or spatially correlated standard errors.

A boundary design cannot rule out every unobserved characteristic that may change at the historical frontier. We therefore ask whether the same pattern appears using a distinct source of variation that does not rely on the Sokoto boundary. Within the former Caliphate, exposure to Sokoto institutions varied considerably: some places were more closely connected to the process of state-building and to the administrative structures through which Sokoto rule was exercised. If greater colonial tax resistance reflects the institutional order established under Sokoto, rather than an unrelated feature of the historical boundary, resistance should also be greater in these more intensively exposed locations. We construct two measures from Last (1967). The first is proximity to Usman dan Fodio's 1804–1817 consolidation corridor, which captures exposure to the formative process of Sokoto state-building. The second is proximity to the Caliphate's major administrative centers, which captures connection to its subsequent administrative structure. Both measures are associated with greater colonial-era tax resistance within former Sokoto. We then ask whether these gradients simply reflect proximity to places where Sokoto's historical presence was generally stronger. As a placebo, we use proximity to the royal ribats, fortified settlements associated primarily with the Caliphate's military and defensive presence. Proximity to the ribats does not predict greater resistance. The within-Caliphate evidence therefore points to a more specific pattern: resistance is greater in places more closely connected to Sokoto's state-building and administrative institutions, but not near sites associated primarily with its military presence.

We next ask whether the form of resistance bears the imprint of the fiscal order that colonial rule displaced. Sokoto's institutional history yields two predictions specific to the fiscal-legitimacy interpretation. First, the additional resistance should be disproportionately associated with Muslims. This prediction follows not only because Sokoto's fiscal order was explicitly Islamic, but because 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 tax grievance, excluding officials and individuals merely mentioned in the document, and classify extracted names as Muslim, Christian, or unclear. Formal complaints inside former Sokoto are substantially more likely to involve Muslim complainants; conditional on observing a formal complaint, complainants are also more likely to be Muslim across bandwidths. This pattern is unlikely to be mechanically explained by later population composition: Muslim population share is smooth at the former Sokoto boundary in the 1952 Census.[1] Consistent with the same interpretation, formal complaints increase with Muslim 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, *jangali* was rejected by Usman dan Fodio's reform movement as a levy without a legitimate basis in Islamic law and replaced by the legally prescribed *zakat* (Last 1967); colonial authorities later reintroduced the tax. If resistance reflects the displacement of Sokoto's fiscal order, complaints inside former Sokoto should therefore be especially likely to target *jangali*. We use a separate LLM classification to identify formal complaints centered specifically on the tax, distinguishing them from documents in which *jangali* 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, part of the excess resistance is concentrated on a levy whose legitimacy had been explicitly rejected under the pre-colonial fiscal order.

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 is also discontinuously higher inside former Sokoto. Taken together, the three results line up with distinct implications of the same historical mechanism: the additional resistance is disproportionately associated with the population whose fiscal obligations changed most sharply, partly concentrated on a tax the displaced order had explicitly deemed illegitimate, and more likely to contest the rightfulness of taxation itself.

The composition of resistance already narrows the set of plausible alternatives. Differences in colonial policy, administrative capacity, enforcement, or public-goods provision could generate a higher overall level of dissatisfaction on one side of the boundary, but they do not naturally predict the historically specific pattern documented above. We nevertheless

[1] Because the Census postdates the fall of the Caliphate by nearly half a century, we do not interpret this as evidence of pre-colonial balance in religious composition. Rather, it shows that the discontinuity in Muslim complainants is not simply a reflection of a corresponding difference in Muslim population shares later in the colonial period.

examine these alternatives directly. We nevertheless examine these alternatives directly. Observable population composition also 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 the estimates are robust to controlling for these characteristics. Nor does the gap appear to reflect differences in the taxes imposed or their burden. Direct-tax instruments were broadly similar across Nigeria, and the results persist both after 1927, when tax systems had largely converged across the former protectorates, and within the Northern Protectorate, where nearby locations faced the same colonial tax regime. Aggregate colonial accounts likewise provide no indication of a heavier fiscal burden in former Sokoto; the bulk of colonial revenue was raised in southern Nigeria.

A second class of explanations concerns how the colonial state implemented taxation. We find little evidence that former Sokoto was governed by a weaker or more coercive fiscal administration. Documents from former Sokoto are, if anything, less likely to describe poor tax administration, while civil servants per capita in the 1931 Census vary smoothly at the boundary and controlling for them leaves the main estimates intact. References to coercion and punishment show no consistent discontinuity, and neither does tax evasion.

A final possibility is that greater resistance reflected receiving fewer public goods in return for taxation. We assemble new measures of roads and railroads, medical facilities, administrative and other public facilities, and post, telegraph, and telephone infrastructure from colonial maps and Blue Books. These measures vary smoothly across the historical boundary, and controlling for them leaves the main estimates essentially unchanged. Taken together, these exercises provide little evidence that the resistance gap is explained by differences in population composition, tax instruments or burden, administrative capacity, enforcement, or public-goods provision.

We finally ask whether the historical legitimacy gap remains visible today. Two channels could plausibly generate persistence. Post-independence taxation largely inherited the colonial fiscal system rather than restoring the pre-colonial Islamic fiscal order, and beliefs about the state's rightful authority to tax may themselves persist across generations. 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. Other outcomes are directionally consistent but less precisely estimated: trust in the tax department is lower, and downstream measures of tax morale also point toward weaker acceptance of taxation. The contemporary evidence is therefore consistent with a persistent difference in fiscal legitimacy,

with the clearest result appearing for the outcome most directly tied to the state's perceived right to tax.

The paper makes three contributions. First, it contributes to the literature on state capacity and legitimacy. A large body of work studies the investments and institutions that allow states to raise revenue and implement policy (Besley and Persson 2009, 2011; Finan et al. 2017; Besley et al. 2021). Much less is known empirically about whether state effectiveness also depends on citizens recognizing the authority behind state demands (Levi 1988; Tyler 2006; Besley 2020; Besley and Dray 2024). The central empirical difficulty is that legitimacy typically covaries with policy, administrative capacity, enforcement, and public-goods provision. Our setting provides unusually sharp variation in the institutional basis of fiscal authority while holding the colonial policy environment comparatively fixed. We show that a broadly similar fiscal regime generated substantially greater resistance where it displaced an established source of legitimate obligation. The results therefore identify legitimacy as a distinct constraint on policy implementation: effective state capacity depends not only on what states can administer or enforce, but also on whether citizens recognize the authority underlying what states demand.

Second, the paper contributes to the literature on historical institutions and state formation. Much of this literature studies how colonial and precolonial institutions generate persistent economic and political effects (Acemoglu et al. 2001, 2002; Banerjee and Iyer 2005; Dell 2010; Michalopoulos and Papaioannou 2013; Lowes et al. 2017; Dell and Olken 2020; Lowes and Montero 2021; Dupas et al. 2024), while related work examines how state interventions interact with religious institutions to shape political organization (Bazzi et al. 2020, 2026). We study a different but complementary question: how the consequences of state-building depend on the institutional order that new state institutions encounter and displace. Colonial rule disrupted existing institutions across Nigeria, but the nature of that disruption differed across settings. In former Sokoto, colonial taxation displaced a centralized, religiously grounded fiscal order that had defined both citizens' obligations and rulers' authority to tax; elsewhere, it generally did not displace a comparable source of fiscal legitimacy. More broadly, our results suggest that the effects of institutional change depend not only on the institutions states introduce, but also on how those institutions interact with pre-existing sources of authority and obligation.

Third, 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 unusually rich evidence on state–citizen interactions, but much of that information has remained inaccessible to systematic quantitative analysis because it is dispersed across large collections

of heterogeneous documents. We use LLMs to recover tax attitudes, formal grievances, locations, dates, complainant characteristics, and the content of disputes from 43,648 colonial records. Because these measures require interpretation rather than simple extraction, we benchmark the LLM against five independent human coders using a common modal-human classification and show that its performance falls within the human range. The contribution is therefore not simply computational scale: it is a framework for turning large historical archives into systematic measures of political and economic behavior that can be linked to modern empirical designs.

The remainder of the paper proceeds as follows. Section 2 presents the historical setting; Section 3 the archival corpus, the outcomes, and their validation; and Section 4 the spatial RD design. Section 5 reports the main results, robustness checks, and the within-Caliphate evidence; Section 6 tests the fiscal legitimacy channel against the alternative explanations; Section 7 documents persistence in contemporary survey data; and Section 8 concludes.

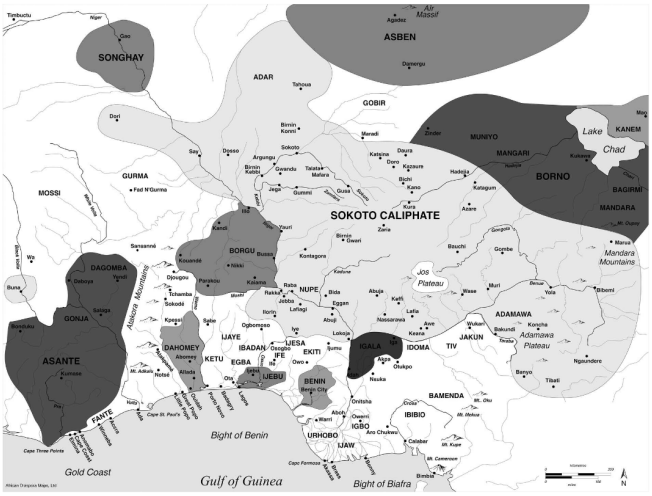
2 Historical Context

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

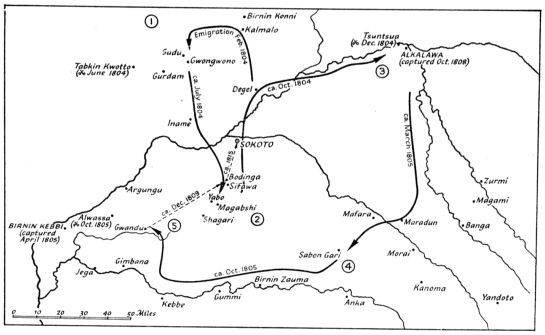
The Sokoto Caliphate (1804–1903) was a centralized Islamic polity founded by Usman dan Fodio's reform movement. Its emirates were governed under Islamic law through a system of courts, administration, and taxation (Last 1967; Hiskett 1973; Lovejoy 2016). Figure 1 shows the territorial extent of the Caliphate, from Lovejoy (2016), alongside 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



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) Territorial Extent of the Sokoto Caliphate



3. Movements of the Shaikh 1804–1817

(b) Usman dan Fodio's Movements, 1804–1817

Figure 1: The Sokoto Caliphate

Notes: Panel (a) shows the territorial extent of the Sokoto Caliphate, from Lovejoy (2016). Panel (b) shows the movements of Usman dan Fodio during 1804–1817, from Last (1967).

obligations to religious status and to rules governing who could be taxed and on what basis. Fiscal obligations were therefore not simply demands for revenue; they were embedded in a recognized Islamic legal framework and administered through emirate institutions.

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, was rejected by the reform movement 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 later 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 presence 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 taxation operated within a recognized normative framework. The distinction was not simply between high and low taxation: Islamic law specified which fiscal obligations were legitimate and constrained the ability of rulers to introduce levies outside that framework. Fiscal authority therefore rested not only on an obligation to pay, but also on accepted rules governing what could legitimately be demanded.

Outside the Caliphate, taxation took heterogeneous forms, including tribute, tolls, levies, and labor obligations. These arrangements could themselves rest on recognized political or customary authority. What was generally absent was a comparable centralized system in which fiscal obligations across a large territory were anchored in a common religious legal order.[2]

2.2 Colonial Taxation and Fiscal Authority

British control over Nigeria was established gradually. Lagos was annexed in 1861, while British authority expanded elsewhere through the Niger Coast Protectorate and the Royal Niger Company during the late nineteenth century. In 1900, Britain established the Protectorates of Northern and Southern Nigeria; Kano and Sokoto were conquered in 1903. The territories were amalgamated into the Colony and Protectorate of Nigeria in 1914, although the broad north–south administrative division persisted throughout the colonial period. Appendix A.4 provides a detailed chronology.

As colonial control expanded, taxation developed within an increasingly common fiscal framework. Fiscal institutions differed across regions and evolved over time, but the major instruments of direct taxation increasingly converged. Figure 2, constructed from the

[2]The Caliphate's frontier illustrates the distinction. Ningi resisted incorporation into Sokoto throughout much of the nineteenth century. In 1846, Mallam Hamza, an Islamic scholar and founder of the Ningi state, left Kano amid opposition to the land tax (*kurdin kasa*) and allied with communities that had resisted Caliphate taxation and slave raiding. Sokoto's fiscal claims in these areas remained contested rather than becoming integrated into the Caliphate's regular fiscal structure (Patton Jr 1975, pp. 235–236). Similar tensions characterized parts of Zuru and other frontier areas. Geographic proximity to Sokoto therefore did not itself imply acceptance of Sokoto's fiscal authority.

Blue Books and the *Reports from the Committee of Taxation*, documents this convergence across the Northern and Southern Protectorates. Important differences remained—notably *jangali* in the North and export taxation in the South—and we examine these directly in Section 6.4. Crucially for our design, the historical Sokoto boundary did not become a colonial fiscal boundary: nearby locations on opposite sides were increasingly subject to similar tax institutions, and substantial portions of the frontier lay within the same colonial protectorate (Figure D1, Panel b).

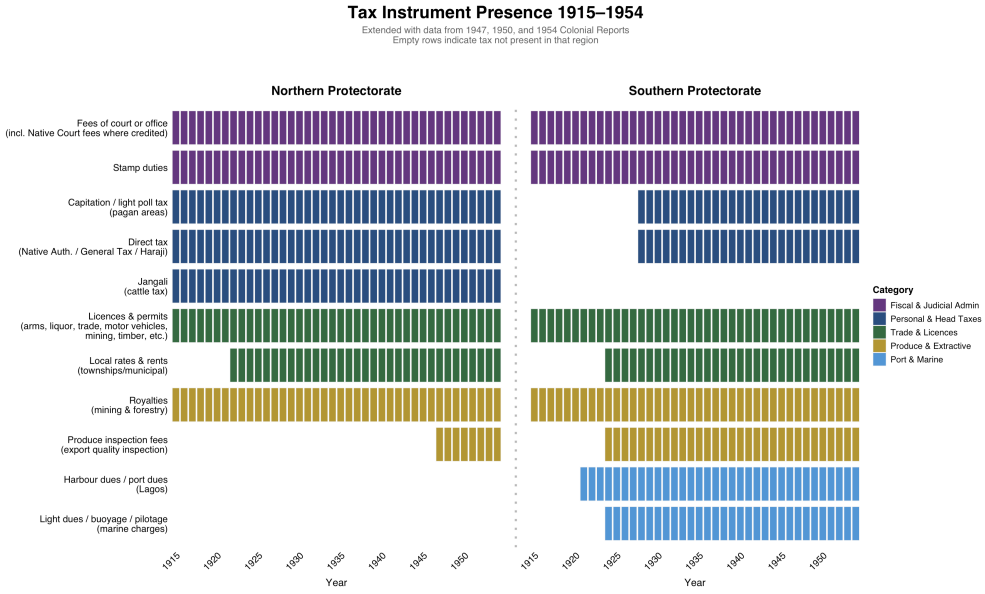


Figure 2: Tax Instruments in the Northern and Southern Protectorates

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, however, the same colonial expansion transformed the institutional basis of taxation without fully replacing the machinery through which taxes were collected. British authorities continued to rely heavily on emirate institutions and local intermediaries, but substantially altered the rules governing fiscal obligations (Kirk-Greene 1965; Watts 1983; Lovejoy 1986; Mamdani 1996). Official collection of *zakat* as a religious obligation was curtailed, while the religious distinctions that had governed liability for other taxes were weakened or removed. Under Sokoto rule, Muslims in the Caliphate's core territories had generally been exempt from poll and land taxes analogous to *jizya* and *kharaj*; under colonial rule, such taxes were imposed without the same religious distinction. Colonial authorities also reintroduced *jangali*, restoring a cattle levy that the Sokoto reform movement had rejected as lacking a legitimate basis in Islamic law. Historical accounts indicate that these

changes in fiscal obligations themselves became sources of resistance.[3] Colonial rule therefore preserved substantial continuity in the institutions of tax collection while changing both the authority under which taxes were owed and, for many Muslims, the taxes to which they were subject.

The resulting institutional change differed sharply across the former Sokoto boundary. Within the Caliphate, colonial taxation displaced a centralized religious fiscal order that had defined both taxpayers' obligations and rulers' authority to tax. Nearby areas were also subjected to colonial taxation, but generally without the displacement of a comparable religiously grounded fiscal order. This contrast yields predictions that go beyond a higher average level of resistance. If the disruption of fiscal authority is the relevant historical difference, resistance should be greater where Sokoto institutions had been more deeply established; it should be disproportionately associated with Muslims, whose fiscal obligations changed especially sharply under colonial rule; and it should concentrate on taxes, such as *jangali*, whose legitimacy the Sokoto order had explicitly contested. The empirical analysis uses each of these predictions to distinguish the fiscal-legitimacy interpretation from a broader legacy of Sokoto rule.

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. Figure 3 overlays the resulting boundary on a 1924 map of Nigeria, which we also use to geolocate the archival records.

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

[3] In the core Sokoto and Gwandu areas, Muslims had not previously been subject to poll taxation or *jangali*, while *zakat* retained its character as a religious obligation. The imposition of poll tax and *jangali* under colonial rule became explicit sources of grievance (Lovejoy 1990; Jumare 1998).

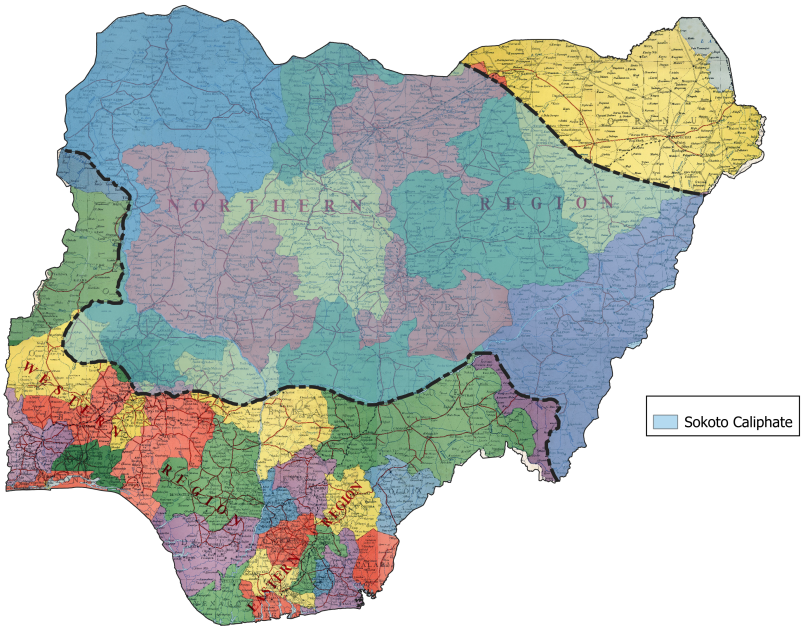


Figure 3: The Sokoto Caliphate Boundary on the 1924 Map of Nigeria

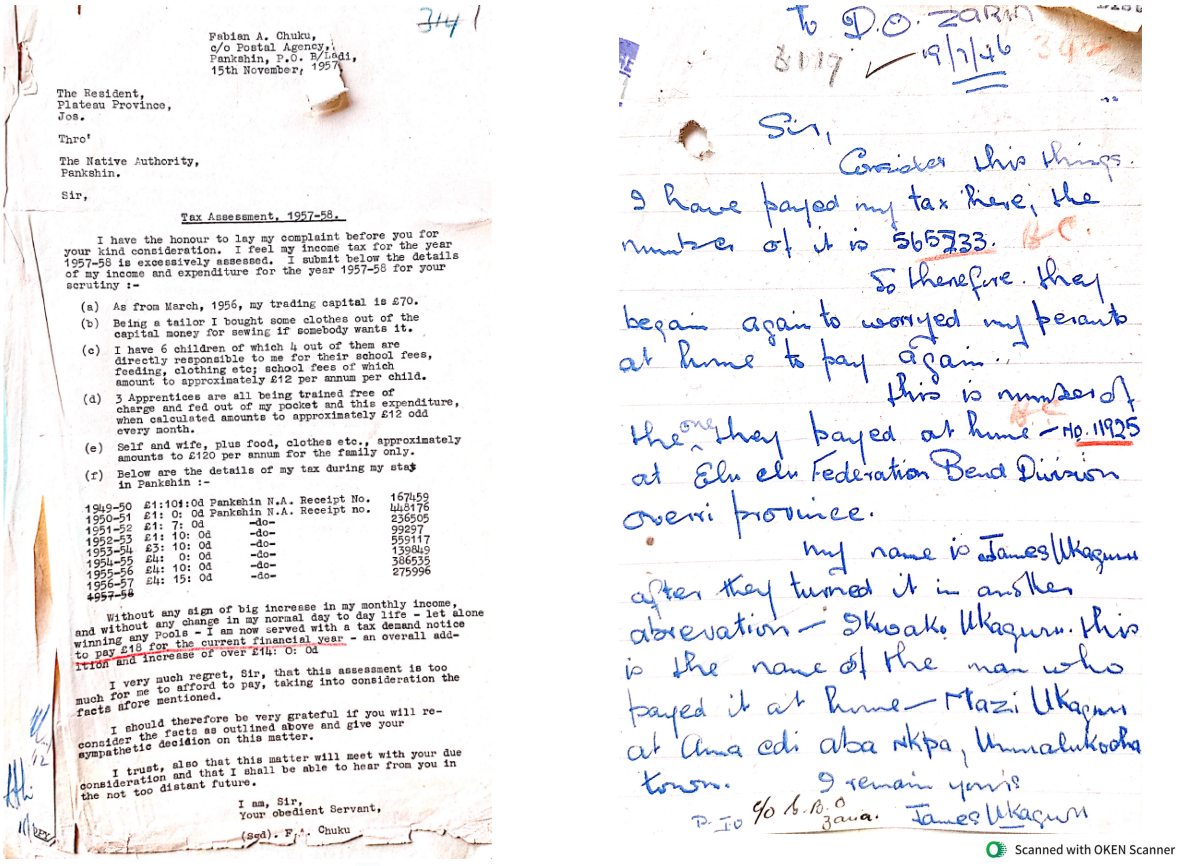
Notes: The figure shows the boundary of the Sokoto Caliphate, digitized from Lovejoy (2016) and georeferenced using stable landmarks (e.g., named settlements, rivers), overlaid on the 1924 map of Nigeria.

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

Our main data consist of 43,648 tax-related documents collected from the three zonal offices of the National Archives of Nigeria. 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. Collection in Ibadan remains ongoing.

The documents record the day-to-day operation of colonial fiscal administration, including correspondence, financial and tax records, reports, petitions, and complaints. Figure 4 reproduces two examples, including a page from a tax-complaints file; Appendix Figure B1 shows the physical archive and additional examples.



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 — 19/7/46

Sir,
Consider this things.
I have payed my tax there, the number of it is 565733.
So therefore they began again to worryed my peasants at home to pay again.
This is number of the one they payed at home — HD. 11925 at Edu edu Federation Bend Division Ovwri province.
my name is James Ukaguru after they turned it in another abrevation — Okwako Ukaguru. this is the name of the man who payed it at home — Nazi Ukaguru at Ama edi aba Nkpa, Umuahukoha town.
I remain yours
P.IO c/o D.O. Zaria. James Ukaguru

Scanned with OKEN Scanner

Figure 4: Examples of Documents

Importantly, the corpus is not selected for episodes of tax resistance. We use an LLM to classify each document by type, allowing documents to belong to multiple categories. Correspondence between authorities is by far the most common category (71%), followed

by financial and tax records (32%) and taxation reports (20%). Correspondence between authorities and civilians accounts for 10% of documents, investigative or special reports for 9%, historical records for 8%, and complaints, requests, or petitions for 7%. The remaining categories are 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%). Expressions of tax resistance therefore appear within a much broader body of routine administrative records.

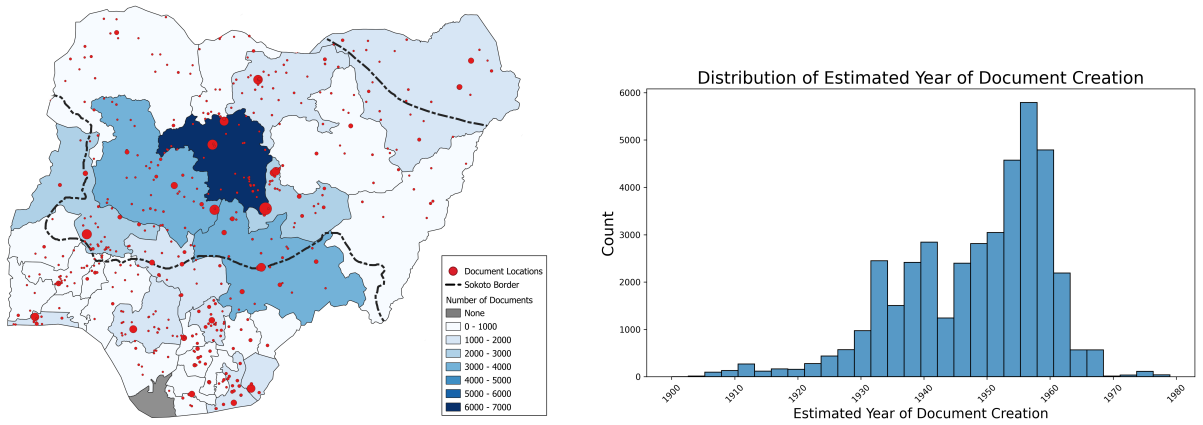
3.2.1 Extracting Location and Date

Each document is assigned a geographic location and year. Geolocation is central to our spatial RD design, which compares documents from locations on opposite sides of the former Sokoto boundary. We first use an LLM to extract the place associated with each document from the text. This step requires resolving substantial variation in historical place names, including alternative spellings, OCR errors, abbreviations, and changes in naming conventions. We standardize these references to canonical place names and match them to localities on a 1924 map of Nigeria, the same historical map onto which we georeference the Sokoto boundary. This places the archival records and the historical boundary in a common geographic space. Appendix Figure B2 shows the resulting document locations on the 1924 map.

We similarly recover the year of each document from the underlying text rather than relying exclusively on archival folder labels. The LLM uses all available dating information—including document headers, dates mentioned in the body, and archival stamps—and extracts the most precise date supported by the document. We retain the corresponding year and use the archival folder year when the extracted date conflicts with it. Figure 5 summarizes the resulting spatial and temporal coverage. The corpus contains documents from locations on both sides of the former Sokoto boundary and spans much of the colonial period, providing the spatial variation used in the RD and allowing us to examine how tax resistance evolved over time.

3.2.2 Measuring Tax Sentiment

Our first outcome measures the sentiment toward taxation expressed or reported in each document. This measure is intended to capture the broader climate surrounding taxation in the colonial record, including opposition expressed by taxpayers as well as problems and resistance recorded by colonial administrators. We ask an LLM to classify each document on an ordinal scale from 0 to 4, ranging from very negative to very positive; documents



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

Figure 5: 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.

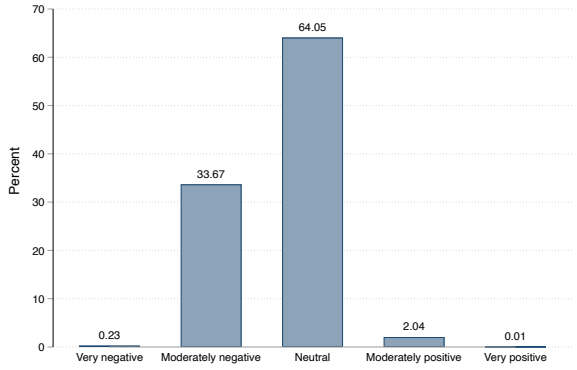
that cannot be reliably interpreted receive a separate score of −1. Very negative documents describe acute conflict or distress surrounding taxation, including riots, violent clashes, or punitive actions. Negative documents describe nonviolent opposition or problems, such as petitions, tax evasion or refusal, corruption allegations, 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 instructions.

Because individual LLM responses can vary, we classify each document multiple times and aggregate the resulting scores using the mode, with the median used to break ties. In the resulting data, 64% of documents are classified as neutral and 34% as negative or very negative; only 2% are classified as positive or very positive. Figure 6 reports the full distribution and the evolution of negative tax sentiment over the colonial period.

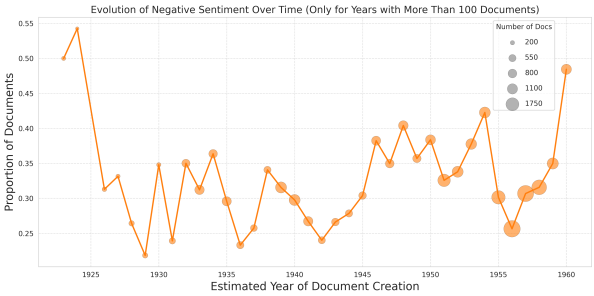
3.2.3 Measuring Formal Tax Complaints

Our second outcome captures a more direct expression of tax resistance. 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. Unlike the sentiment measure, which can reflect resistance or problems described by either taxpayers or administrators, this outcome identifies an explicit grievance directed toward the state. It therefore provides a narrower but more concrete measure of contestation.

We identify formal complaints using a separate LLM classification of tax-related issues.



(a) Distribution of Tax Sentiment

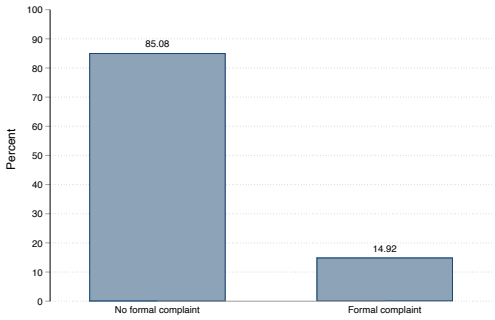


(b) Negative Sentiment over Time

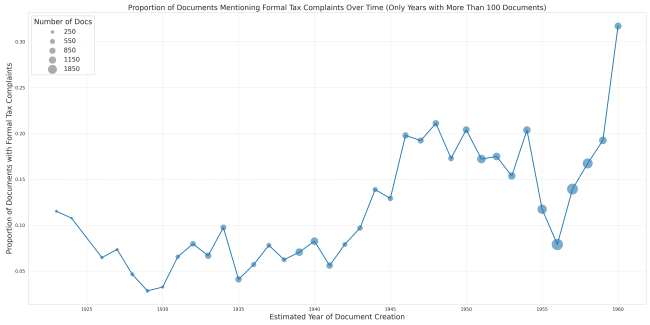
Figure 6: Tax Sentiment

Notes: Panel (a) shows the distribution of the LLM-coded tax sentiment score across documents. Panel (b) shows the share of documents expressing negative tax sentiment by year.

A document is classified as a formal complaint only when it contains an explicit grievance about taxation submitted to or reported to an authority; merely discussing taxation or describing an administrative problem is insufficient. Appendix B.5 reproduces the classification instructions.[4] Formal tax complaints appear in 15% of the corpus. Figure 7 reports their overall frequency and evolution over the colonial period.



(a) Share of Formal Tax Complaints



(b) Formal Tax Complaints over Time

Figure 7: Formal Tax Complaints

Notes: Panel (a) shows the share of documents classified as formal tax complaints. Panel (b) shows the share of documents containing a formal tax complaint by year.

The two outcomes capture complementary dimensions of tax resistance. Negative sentiment is broad: it measures adverse discussion of taxation whether voiced by taxpayers or

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

recorded by administrators. Formal complaints are narrower and identify explicit grievances brought before an authority.

Both measures become more prevalent over the colonial period, particularly in its later decades. This common pattern is reassuring: it is consistent with historical accounts documenting the expansion of organized 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. Rather, their correspondence with the historical record provides a first check that the measures recover meaningful variation in the underlying documents.

3.3 Validating the LLM Measures

We next subject the measures to a more direct validation exercise, comparing the LLM classifications with independent human readings of the same documents. This exercise addresses two distinct questions: whether tax sentiment and formal complaints can be identified consistently from the historical record, and whether the LLM can recover those classifications as reliably as a human reader. Both questions matter because the outcomes require interpretation rather than simple keyword matching. Assessing tax sentiment requires distinguishing routine reports of collection from descriptions of administrative problems or resistance, while identifying a formal complaint requires determining whether a document contains an actual grievance submitted to an authority rather than merely discussing a tax dispute. LLMs may misinterpret such distinctions or produce classifications unsupported by the underlying text (Korinek 2023; Amirizaniani et al. 2024; Dell 2025).

Neither outcome has an independently observable ground-truth label, but agreement across human readers provides a direct indication of whether the classifications are sufficiently well defined to be measured reliably. Following the Turing-test approach of Lagakos et al. (2025), five research assistants independently code a random audit sample of 100 documents using the same substantive definitions as the LLM. For each document, we construct the modal human classification and use it as a common benchmark. We then evaluate each of the five human coders and the LLM against this same benchmark. This approach allows for occasional human disagreement without treating any individual coder as ground truth.

Because the two outcomes differ in their structure, we use different performance measures. Tax sentiment is ordinal, so we calculate each rater's absolute distance from the modal human score. Formal complaints are binary, so we calculate the fraction of documents for which each rater agrees with the modal human classification. The Turing test asks whether the LLM's performance falls within the range observed among the five human readers. If the

LLM disagrees with the human consensus no more than individual humans disagree with that same consensus, its classifications do not stand out from human readings (Lagakos et al. 2025).

Figure 8 reports the results. Human coders exhibit a high degree of agreement on both outcomes, indicating that the classifications can be identified consistently from the underlying text. For formal tax complaints, individual research assistants agree with the modal human classification in approximately 94–98% of documents. For the five-category sentiment measure, the average absolute deviation from the modal human score ranges from approximately 0.16 to 0.34 points on a 0–4 scale. Against these relatively tight human benchmarks, the LLM falls within the human range on both outcomes: its average absolute deviation is approximately 0.25 for tax sentiment, while its agreement rate is approximately 94% for formal complaints. Thus, the primary outcomes are sufficiently well defined to be coded consistently by independent human readers, and the LLM reproduces those classifications at a level comparable to an additional human coder.

The scale of the archive makes this comparison consequential. Based on the time required by the research assistants to read and classify the 100-document audit sample, having a single research assistant code all 43,648 documents would require more than one year of full-time work. The LLM allows us to apply a common classification procedure to the entire corpus while producing classifications that fall within the range of independent human judgments on the audit sample. This combination of scalability and human-level agreement makes it possible to construct systematic measures of tax attitudes and resistance from an archival corpus of this size.

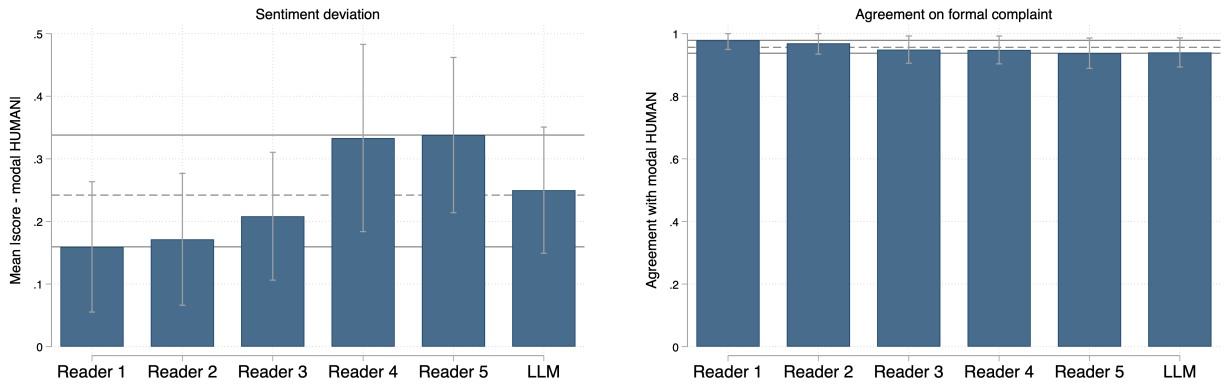
We conduct analogous validation exercises for the secondary outcomes used later in the paper, including state coercion or punishment, tax evasion, corruption, and administrative failure. The LLM again performs within the range of human coders (Appendix Figures C1–C5).[5] Appendix B reproduces the classification prompts and provides additional implementation details.

4 Empirical Strategy

4.1 Spatial Regression Discontinuity Specification

Our empirical strategy exploits the former boundary of the Sokoto Caliphate. Colonial conquest eliminated this boundary as a fiscal jurisdiction: locations on either side were incorporated into the British colonial state and increasingly subjected to similar tax insti-

[5] We discuss the more interpretive direct measure of fiscal legitimacy separately in Section 6.



(a) Tax Sentiment (b) Formal Tax Complaints

Figure 8: Turing Test: LLM Performance Falls Within the Human Range

Notes: The figure benchmarks each rater, five human research assistants and the LLM, against the modal human score on a random audit sample of 100 documents. Panel (a) reports the absolute distance to the modal human tax sentiment score. Panel (b) reports the agreement rate with the modal human formal-complaint classification.

tutions. The boundary nevertheless continued to separate places with different pre-colonial fiscal histories. Inside it, colonial taxation displaced the Islamic fiscal order described in Section 2; immediately outside it, no comparable Sokoto fiscal order had governed.

We exploit this historical discontinuity in a spatial regression discontinuity design, comparing tax resistance in nearby locations on opposite sides of the former Caliphate boundary. 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 Sokoto Caliphate. The function $f(location_l)$ is linear in distance to the historical boundary, with separate slopes on the two sides. 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 boundary, while archive fixed effects $\zeta_{a(d)}$ absorb level differences across the Kaduna, Enugu, and Ibadan collections. The coefficient β captures the discontinuity in tax resistance at the former Sokoto boundary. 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. Our baseline specification is local linear; Section 5 examines alternative spatial functional forms and approaches to inference.

Identification comes from comparisons among nearby locations facing the same segments of the former Sokoto boundary, which no longer determined fiscal jurisdiction under colonial rule. The key identifying assumption is that, absent historical exposure to Sokoto, other determinants of colonial tax resistance would vary smoothly across that boundary. Two threats are especially important. First, Sokoto may have expanded into locations that already differed systematically in ways related to later tax resistance. Second, the historical frontier may subsequently have become aligned with differences in colonial fiscal policy. We examine each 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).[6] These accounts generate a first set of observable implications. We examine terrain, vegetation and agricultural ecology, and tsetse suitability, and then broaden the balance exercise to a much wider 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 implicated by the historical accounts of Sokoto expansion or in a substantially broader set of observable pre-Sokoto determinants of settlement and economic activity.

A related concern is that Sokoto expansion may have followed pre-existing ethnic or livelihood patterns—for example, if the Caliphate disproportionately incorporated Fulani or other pastoral populations that were independently more resistant to taxation.[7] The *Ethnographic Atlas* of Murdock (1967) provides a useful check: the Sokoto boundary cuts through several ethnic homelands rather than coinciding with their borders (Panel (a) of Appendix Figure D1). The main estimates also remain similar when we include ethnic-territory fixed effects, so that comparisons are made within the same ethnic homeland on

[6] 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. Historians therefore emphasize the geography of cavalry warfare in explaining the limits of Sokoto expansion.

[7] This concern is particularly relevant because Usman dan Fodio and many of the jihad's leaders were Fulani, and members of this leadership subsequently became prominent in the emirate structures established after the conquest of the Hausa states. The Caliphate was not, however, ethnically homogeneous: the jihad unfolded principally in Hausaland, drew support from a broader population, and the resulting state incorporated numerous ethnic groups (Last 1967).

opposite sides of the Sokoto boundary (Appendix Table D3). This specification also absorbs fixed differences across ethnic groups in pre-colonial characteristics that could plausibly affect later tax resistance.[8]

Common Colonial Fiscal Regime. A second concern is that the former Sokoto boundary may later have become aligned with differences in colonial fiscal policy or state capacity. If locations on the Sokoto side faced systematically different taxes, heavier fiscal demands, weaker administration, harsher enforcement, or lower public-goods provision, the resistance discontinuity could reflect the colonial state itself rather than the pre-colonial fiscal order it encountered. Several features of the setting help address this concern. As documented in Section 2.2, direct-tax institutions became increasingly similar across colonial Nigeria, and the Sokoto boundary did not determine colonial administration. Panel (b) of Appendix Figure D1 shows that the historical frontier cuts across the colonial protectorates, so substantial portions of the boundary compare locations governed within the same protectorate. We exploit this feature directly by re-estimating the spatial RD within the Northern Protectorate, where the colonial tax regime is held substantially more tightly fixed; the resulting estimates remain similar to the baseline results (Appendix Table D4). The discontinuity also persists after 1927, when direct-tax instruments had largely converged across the former protectorates (Appendix Table G3). Section 6.4 examines the remaining dimensions of the colonial state directly. We find little evidence of discontinuities in fiscal burden, administrative capacity, enforcement, or public-goods provision that could account for the result. Taken together, these results provide little support for the view that the Sokoto discontinuity reflects a systematically different colonial fiscal environment rather than the pre-colonial fiscal order encountered by colonial rule.

5 Results

5.1 Tax Resistance at the Sokoto Boundary

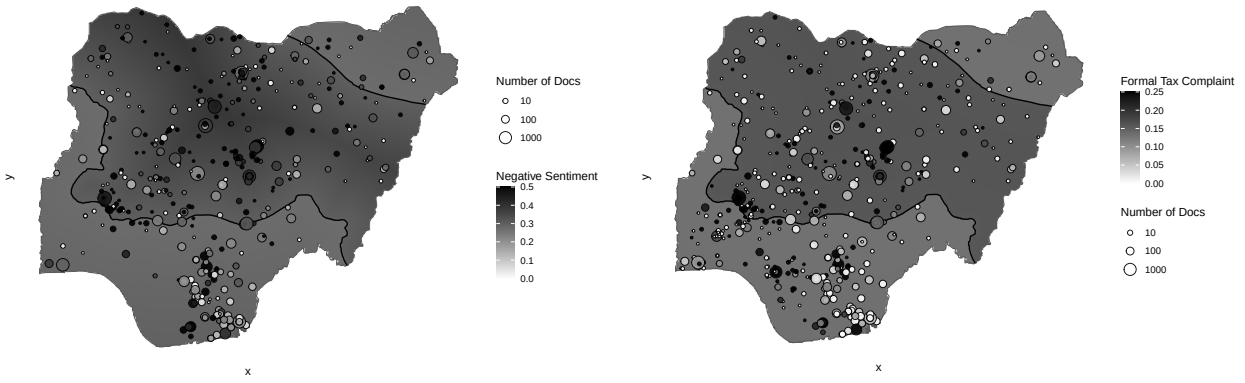
Table 1 reports the main estimates of equation (1). Colonial tax resistance is substantially higher just inside the former Sokoto Caliphate. At the MSE-optimal bandwidth, documents from locations inside the boundary are 9.3 percentage points more likely to express negative tax sentiment, relative to a control mean of 24 percent just outside the boundary. The

[8] Pastoralism is one relevant example. Ethnic groups differed substantially in their reliance on pastoral livelihoods, and Fulani populations that became prominent in the emirate structures were historically more pastoral. The ethnic-territory fixed-effects specification therefore helps address the possibility that the Sokoto discontinuity reflects greater tax resistance among more pastoral populations rather than differences in fiscal legitimacy.

estimated increase is therefore roughly 40 percent of the outside-Sokoto mean. The effect on formal complaints is even larger relative to its baseline: documents inside former Sokoto are 18.2 percentage points more likely to contain a written tax grievance submitted to an authority, compared with a control mean of only 7 percent. Crossing the historical boundary therefore more than triples the probability that the archival record contains a formal tax complaint.

The pattern is not specific to the bandwidth selected by the MSE criterion. Across fixed bandwidths from 50 to 200 kilometers, the estimated discontinuity in negative sentiment ranges from 14.3 to 19.4 percentage points, while the discontinuity in formal complaints ranges from 8.1 to 19.9 percentage points. The estimates are positive at every bandwidth and statistically significant throughout. Thus, the central result appears in both of our main outcomes and over a broad range of local comparisons around the historical boundary.

The discontinuity is also visible geographically. Figure 9 plots the raw location-level outcomes without covariate adjustment. Both negative sentiment and formal complaints rise sharply on the Sokoto side of the historical boundary. The maps also make clear that the result is not generated by a single isolated cluster of locations: elevated resistance appears along substantial portions of the former Caliphate frontier.



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

Figure 9: 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).

The Sokoto resistance gap becomes more pronounced over the colonial period. Table 2 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

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

1950. We do not place strong interpretation on this temporal pattern, since the archival composition and broader colonial environment also changed over time. The main point is that the Sokoto discontinuity is not confined to a narrow episode of colonial rule and, if anything, becomes more pronounced in the later colonial period.

We also ask whether the resistance captured in the archival record is related to more visible forms of collective action. Appendix H constructs a separate dataset of 100 tax protests, riots, and revolts, combining events identified in the colonial documents with those documented in historical sources. These events are strongly positively associated with both negative tax sentiment and formal complaints (Appendix Table H1). This correspondence provides an additional validation of our archival measures: locations and periods in which the documents record greater dissatisfaction and more formal grievances are also those in which collective tax resistance is more likely to occur.

The spatial RD for protests and riots themselves is considerably less precise, and we do not detect a significant discontinuity at the Sokoto boundary (Appendix Table H2). This contrast highlights what the archival measures capture beyond overt collective action. Riots and revolts are salient but high-threshold manifestations of resistance, whereas negative sentiment and formal complaints capture lower-intensity forms of contestation that need not escalate into collective action. This distinction is especially important in a colonial setting, where overt resistance could carry severe costs. The Satiru revolt of 1906 provides a stark illustration of these costs.[9] Focusing only on rare episodes of protest may therefore miss substantial variation in how state demands are received. Consistent with this distinction, the Sokoto boundary generates a large discontinuity in lower-intensity expressions of tax resistance, while any discontinuity in overt collective action is estimated too imprecisely to distinguish from zero.

5.2 Robustness

We assess the robustness of the spatial discontinuity to three broad concerns: the construction 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.

Archival composition and coverage. A first concern is that the discontinuity could reflect differences in which records survive or enter the archival sample rather than differences

[9]The Satiru revolt formed part of a broader Mahdist challenge to British colonial rule and the political order established after the conquest of the Sokoto Caliphate. British forces violently suppressed the uprising, killing or imprisoning many of its participants and leaders and sharply curtailing subsequent revolutionary resistance (Lovejoy 1990).

Table 2: Growing Tax Resistance over the Colonial Period

	100km Bandwidth		
	(1)	(2)	(3)
	1900–1939	1940–1949	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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in the content of those records. The three archives may differ in the number and types of documents they preserve, and their composition might also change over time. Our baseline specification includes archive fixed effects, which absorb average differences across archives. We then address finer differences in composition in two ways. First, we add document-type fixed effects, so that identification comes from comparisons among the same kinds of records on opposite sides of the Sokoto boundary. Second, we interact archive fixed effects with decade and then with year, so that comparisons are made among documents drawn from the same archive in the same period. The estimated discontinuities remain positive and statistically significant across these specifications, although the magnitudes are attenuated in some cases (Appendix Tables D5, D6, and D7). These results make it unlikely that the baseline discontinuity is generated solely by systematic differences in the composition of archival holdings across space or time.

A separate concern is uneven archival coverage across locations. Some places contribute many more surviving documents than others, so a document-level specification gives greater weight to locations that are more intensively represented in the archive. We first test whether archival intensity itself changes at the Sokoto boundary and find no discontinuity in the number of surviving documents (Appendix Table D8). We then reweight documents by the inverse number of records associated with each location, so that each location receives equal weight regardless of how many documents survive from it; the resulting estimates are very similar to the baseline (Appendix Table D9). Finally, because archival intensity may also scale with underlying population size, we reweight documents by the inverse population of their division using the 1952 Census. The estimates are again similar (Appendix Table D10). Taken together, these exercises indicate that the results are not driven by differences in what types of records survive, by a discontinuity in archival coverage, or by disproportionate weight on locations that are either more heavily represented in the archive or more populous.

Boundary measurement and historical ambiguity. A second concern is that the estimated discontinuity may be sensitive to measurement error in the mapped Sokoto boundary. The boundary is digitized from a historical map drawn by an expert historian and, like any reconstruction of a nineteenth-century frontier, is not perfectly precise. At the same time, the map identifies numerous named settlements and other stable geographic features, which allow us to georeference the frontier closely. Based on the remaining uncertainty around these reference points, we treat locations within 10 kilometers of the boundary as the area in which treatment assignment is most plausibly affected by mapping error. We therefore estimate a donut RD excluding this 10-kilometer band. The resulting estimates are similar to the baseline despite removing precisely the observations for which treatment classification is most sensitive to boundary uncertainty (Appendix Table D11).

We then examine whether the results depend on how historically ambiguous areas are classified. Kanem–Bornu was a neighboring Islamic polity in which some institutions relevant to our setting, including Islamic forms of taxation, were also present. We therefore both exclude observations adjacent to Kanem–Bornu and, alternatively, classify Kanem–Bornu as part of the treated area.[10] Neither exercise materially changes the results (Appendix Table D12). We conduct an analogous exercise for the Jos Plateau, whose relationship to the Sokoto Caliphate is less clear-cut in the historical literature.[11] Excluding the plateau leaves the main conclusions unchanged, while redefining the boundary to carve it out yields estimates of similar sign and magnitude, though with less precision in some specifications (Appendix Table D13). These exercises indicate that the main discontinuity is not driven by plausible mapping error or by a particular coding of historically contested portions of the Sokoto 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 D14 and D15), 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 D16). 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 D17). Thus, neither the estimated discontinuities nor the statistical conclusions depend materially on the baseline spatial functional form or error structure.

[10]Kanem–Bornu was itself a long-standing Islamic polity with established Islamic legal and fiscal institutions (Lavers 1971). Its fiscal system was nevertheless less fully organized around the reformist Islamic principles that characterized nineteenth-century Sokoto: Islamic taxation coexisted with other, more discretionary levies and fiscal practices inherited from the older political order (Cohen 1967). Bornu therefore provides an intermediate case—closer to Sokoto than most neighboring areas in its Islamic institutions, but distinct from Sokoto's reformist fiscal order.

[11]The Jos Plateau lay within the broader sphere of Sokoto influence, but the extent of effective incorporation was limited and contested. Communities on the Plateau repeatedly resisted nineteenth-century incursions associated with the Caliphate and maintained substantial political independence until British conquest (Tambo 1978). We therefore treat the Plateau's classification as historically ambiguous.

5.3 Within-Caliphate Variation

A boundary design cannot rule out every unobserved characteristic that may change at the historical frontier. We therefore ask whether the same pattern appears using a distinct source of variation that does not rely on the Sokoto boundary. Within the former Caliphate, exposure to Sokoto institutions varied considerably: some places were more closely connected to the process of state formation and to the administrative structure through which Sokoto rule was subsequently exercised. If greater colonial tax resistance reflects the institutional order established under Sokoto, rather than an unrelated feature of the historical boundary, resistance should also be greater in these more intensively exposed locations.

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 the consolidation of the new political order.[12] We digitize the locations identified by Last (1967) and connect them using an elevation-based least-cost path, producing the consolidation corridor shown in Figure 1 and Appendix Figures E1 and E2. Proximity to this corridor therefore captures connection to the areas most directly associated with the formation and early consolidation of the Caliphate.

The second measure captures proximity to the Caliphate's subsequent administrative structure. Using the historical maps in Last (1967), we digitize major administrative locations, including emirate centers and territories administered under the Vizier (Appendix Figure E3).[13] These locations capture connection to the institutions through which Sokoto rule was exercised after consolidation. The two measures therefore capture complementary dimensions of institutional exposure: proximity to the formation of the Sokoto political order and proximity to the administrative network through which that order was subsequently governed.

Table 3 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. Panel A relates tax resistance to proximity to the consolidation corridor among locations within 100 kilometers of the reconstructed path. Greater proximity predicts significantly more negative tax sentiment; the estimate for formal com-

[12] The locations include places 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).

[13] 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). These locations therefore capture the principal nodes through which the Caliphate's political and fiscal institutions were exercised.

plaints is also positive but less precisely estimated. Because distance to a reconstructed route may capture direct exposure imperfectly, Panel B instead uses the locations that Usman dan Fodio is recorded as having visited. We match these locations to geographically and historically comparable locations elsewhere within the Caliphate and estimate document-level regressions with matched-pair fixed effects. Locations along the Shaikh's route exhibit significantly more formal complaints and, at the 10 percent level, more negative tax sentiment than their matched counterparts. Panel C turns to the Caliphate's administrative network. Among locations within 100 kilometers of the major administrative centers, greater proximity predicts significantly more negative tax sentiment and more formal complaints.

The pattern is therefore not tied to a single construction of exposure or empirical approach. Both proximity to the formative process of Sokoto state-building and proximity to its subsequent administrative structure are associated with greater colonial-era tax resistance. Most importantly, these comparisons are made entirely within the former Caliphate and therefore do not rely on the location of the Sokoto boundary. They provide a distinct piece of evidence consistent with the interpretation that the border discontinuity reflects exposure to Sokoto institutions rather than an unrelated characteristic of the historical frontier.

The within-Caliphate gradients could nevertheless reflect proximity to places where Sokoto's historical presence was generally stronger, rather than exposure to the state-building and administrative institutions emphasized above. We assess this alternative using the royal ribats mapped by Last (1967). Established between 1839 and 1845, the ribats were fortified settlements associated primarily with the Caliphate's military and defensive presence (Appendix Figure E4). They therefore capture a distinct dimension of Sokoto's geographic presence—military control and frontier defense—rather than the state-building and administrative exposure measured above. If colonial tax resistance were simply greater in places where Sokoto power or influence had been more concentrated, proximity to the ribats should predict greater resistance as well. It does not: proximity to the royal ribats is not positively associated with either negative tax sentiment or formal complaints; if anything, the relationships are negative (Appendix Table E1).

The placebo therefore helps distinguish exposure to Sokoto's state-building and administrative institutions from Sokoto presence more generally. Resistance is greater in locations more closely connected to the Caliphate's formation and administrative structure, but not near sites associated primarily with its military and defensive presence. These comparisons do not eliminate every source of within-Sokoto heterogeneity, but they are difficult to reconcile with an interpretation in which the boundary effect simply reflects proximity to Sokoto power or historical activity in general. Instead, the within-Caliphate evidence points to a

Table 3: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

more specific pattern: colonial tax resistance is strongest along dimensions of Sokoto exposure most closely associated with the institutions through which its political and fiscal order was established and administered.

6 Mechanisms

The results so far show that colonial taxation encountered greater resistance in areas formerly governed by the Sokoto Caliphate. Our interpretation is that colonial rule disrupted an established source of fiscal legitimacy. The colonial regime preserved important elements of the machinery of tax collection but altered both the authority underlying fiscal obligations and, for many Muslims, the obligations themselves. 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* that Sokoto reformers had rejected on Islamic grounds were reintroduced.

This interpretation yields predictions about the form of resistance, not simply its level. First, the additional resistance should be disproportionately associated with Muslims, whose fiscal status changed especially sharply under colonial rule. Second, resistance should be concentrated on taxes whose legitimacy had been explicitly contested under the Sokoto order. We test these two predictions in turn and then provide complementary evidence using a direct text-based measure of contested fiscal legitimacy.

6.1 Testing the Fiscal Legitimacy Channel (I): Who Complains?

Our first test asks who is responsible for the formal complaints in the archival record. This is informative because 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 later imposed by the colonial state had previously been restricted to other populations or rejected altogether. If the disruption of this fiscal order contributed to resistance, the increase in 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 extracts the individual actually making the complaint, excluding officials and individuals merely mentioned in the document. Thus, when an official reports another person's grievance, the procedure identifies that individual rather than the official as the complainant. We then match the extracted names to curated repositories of Muslim and Christian names, with

unmatched names classified by the LLM using linguistic and cultural naming conventions. Appendix B.7 describes the procedure in detail.

Formal complaints on the Sokoto side of the boundary are substantially more likely to involve Muslim complainants. Figure 10 shows the composition of complainants on either side of the boundary, while Table 4 estimates 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, and the point estimates are positive at every reported bandwidth. Panel B asks whether the composition of formal complaints also shifts toward Muslim complainants. The estimates are positive throughout, though less precisely estimated at the MSE-optimal bandwidth; at 50 kilometers, the Muslim share among complainants rises by 23 percentage points relative to a control mean of 11 percent. The additional formal resistance inside former Sokoto is therefore disproportionately associated with Muslim complainants.

This pattern is unlikely to mechanically explained by a larger Muslim population on the Sokoto side. Muslim population share varies smoothly at the historical boundary in the 1952 Census (Appendix Table D2), even though complaints involving Muslim complainants change sharply. Because the Census postdates the fall of the Caliphate by nearly half a century, we do not interpret this as evidence that religious composition was balanced before colonization. Rather, it shows that the discontinuity in Muslim complainants is not simply a reflection of a corresponding difference in Muslim population shares later in the colonial period.

A complementary test exploits variation in religious composition away from the boundary. The fiscal-legitimacy interpretation implies that Muslim population share should matter differently depending on the pre-colonial fiscal order: a larger Muslim population should predict greater resistance within former Sokoto, where colonial taxation displaced an Islamic fiscal system, but not necessarily outside it. Consistent with this prediction, formal complaints increase significantly with the division-level Muslim population share within former Sokoto, whereas outside the Caliphate the corresponding relationship is negative (Appendix Table F1). The boundary comparison and the within-Sokoto heterogeneity therefore point in the same direction: the additional resistance is concentrated among the population whose fiscal position was most directly affected by the displacement of Sokoto's Islamic fiscal order.

Table 4: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

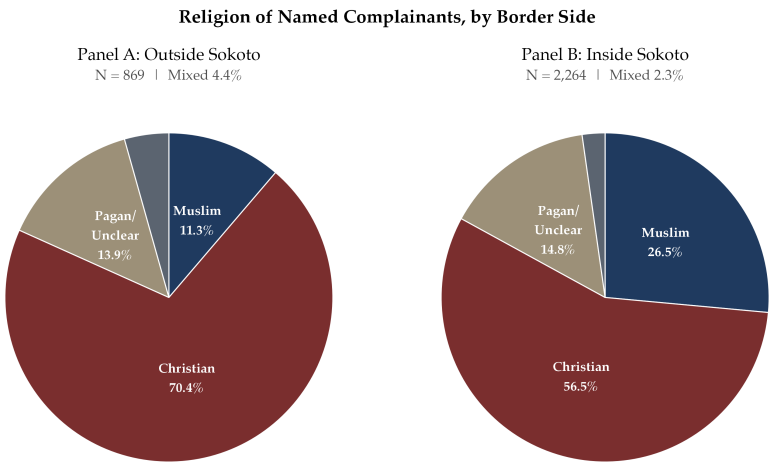


Figure 10: Complainant Religion on Either Side of the Sokoto Boundary

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 Testing the Fiscal Legitimacy Channel (II): Which Taxes Are Contested?

Our second test asks whether resistance concentrates on taxes whose legitimacy had been explicitly contested under Sokoto rule. *Jangali*, a tax on cattle, provides the sharpest case. As discussed in Section 2.1, *jangali* predated the jihad but was rejected by Usman dan Fodio's reform movement as a levy without a legitimate basis in Islamic law and replaced by the legally prescribed *zakat*. Colonial authorities later reintroduced the tax, and historical accounts identify its reimposition as a source of grievance in areas whose Muslim populations had previously been exempt. A complaint against *jangali* therefore provides an unusually direct test of whether colonial resistance concentrated on a levy that the displaced fiscal order had explicitly deemed illegitimate.

We use an LLM to identify formal complaints centered specifically on *jangali*, distinguishing substantive grievances about the tax from documents in which it is merely listed or mentioned alongside unrelated taxes. Appendix B.8 reproduces the classification prompt. Similar 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, such taxes were imposed without the same religious distinction. These levies therefore also represented a sharp change in Muslims' fiscal obligations. In practice, however, colonial records use several overlapping terms for poll and land taxation, making them substantially harder to identify consistently. We therefore focus on

jangali, for which both the historical prediction and the textual classification are unusually sharp.

Table 5 shows a pronounced discontinuity in *jangali*-related grievances. Such complaints are essentially absent immediately outside the former Caliphate but become more common on the Sokoto side. Panel A considers the unconditional probability that a document contains a formal complaint centered on *jangali*. The estimated discontinuities are positive at every bandwidth and statistically significant at four of the five reported bandwidths. Panel B instead asks whether, conditional on a formal complaint, the composition of grievances shifts specifically toward *jangali*. These estimates are less stable across bandwidths, though generally positive and economically large.[14] The unconditional evidence is therefore the cleaner result: part of the excess formal resistance inside former Sokoto is concentrated on a levy that the displaced fiscal order had explicitly deemed illegitimate.

6.3 Testing the Fiscal Legitimacy Channel (III): LLM Measure of Contested Legitimacy

The preceding tests infer the legitimacy mechanism from two features of resistance—who complains and which taxes are contested—without requiring us to classify legitimacy claims directly. As complementary evidence, we construct a text-based measure that does so explicitly. The measure identifies whether a document treats taxation as rightful or instead challenges the authority, religious principle, custom, or law on which the tax rests. Importantly, the classification distinguishes such challenges from negative sentiment, complaints about administration, and objections to enforcement. Appendix B.9 reproduces the classification prompt, and Appendix Figure C5 reports its validation.

The contested-legitimacy measure is strongly associated with both negative sentiment and formal complaints at the document level (Appendix Table F2). More importantly for our argument, it also exhibits a sharp discontinuity at the Sokoto boundary. Table 6 shows that documents from former Sokoto are significantly more likely to contest the legitimacy of taxation at the MSE-optimal bandwidth. The magnitude of the discontinuity increases as we restrict attention to classifications in which the LLM expresses greater confidence, from 0.083 in the full sample to 0.205 among the most confidently classified documents (Panels A–C). The archival language therefore reinforces the two preceding tests: resistance on the Sokoto side was more likely to challenge the rightfulness of taxation itself, rather than simply expressing dissatisfaction with how taxes were administered or enforced.

[14] The exception is the 100-kilometer bandwidth, for which the estimated discontinuity is negative and statistically significant.

Table 5: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The three exercises draw on different features of the archival record but test implications of the same historical change. The Muslim-complainant results show that the additional resistance is disproportionately associated with the population whose fiscal obligations changed most sharply under colonial rule. The *jangali* results show that part of this resistance concentrates on a tax that the Sokoto order had explicitly rejected. The text-based measure, in turn, shows that the language of the dispute itself more often challenges the rightfulness of taxation. No single exercise is decisive on its own. Taken together, however, they indicate that the Sokoto discontinuity is not simply an increase in generalized dissatisfaction with colonial taxation: the composition and content of resistance closely match the institutional changes produced by the displacement of Sokoto's fiscal order.

6.4 Alternative Explanations

The composition of resistance already narrows the set of plausible alternatives. Differences in colonial policy, administrative capacity, enforcement, or public-goods provision could in principle generate a higher overall level of dissatisfaction on one side of the boundary, but they do not naturally predict that the additional complaints should come disproportionately from Muslims, concentrate partly on a tax explicitly rejected under Sokoto's Islamic fiscal order, and more often challenge the legitimacy of taxation itself. We nevertheless examine these alternatives directly, together with differences in population composition. Appendix G reports the corresponding data, maps, and estimates.

Demographic Composition. A first possibility is that the Sokoto boundary separates populations that were systematically different in ways that could independently generate more formal tax resistance. As a large and institutionally developed pre-colonial polity, Sokoto may have contained more educated, wealthier, or more commercially engaged populations, who may have been more able to submit written grievances or more exposed to taxation. Its population was also more heavily Muslim, raising the possibility that the complaint discontinuity simply reflects religious composition rather than the displacement of an Islamic fiscal order. The 1952 Census provides measures of Muslim population share, Latin-script literacy, and occupational structure. None exhibits a discontinuity at the historical boundary (Appendix Table D2).[15] The main results are also robust to conditioning directly on these characteristics: controlling for literacy leaves both discontinuities precisely

[15]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.

estimated, while controlling for occupational shares produces the same qualitative pattern (Appendix Tables G1 and G2). These results provide little evidence that observable differences in population composition account for the resistance gap.

Tax Instruments. A second possibility is that former Sokoto simply faced a different set of colonial taxes. We reconstruct colonial tax instruments from the *Blue Books* and the *Reports from the Committee of Taxation*. The broad structure of direct taxation was similar across Nigeria, although there were notable exceptions: *jangali* was levied in the North, while export taxes were concentrated in the South (Figure 2). We address these residual differences in two ways. First, the Sokoto discontinuity persists in the post-1927 period, when direct-tax instruments had largely converged across the former protectorates (Appendix Table G3). Second, it remains when we restrict the sample to the Northern Protectorate, comparing locations subject to the same colonial tax regime on either side of the Sokoto boundary (Appendix Table D4). Differences in statutory tax instruments therefore cannot readily account for the main resistance gap.

Tax Burden. Greater resistance might instead reflect heavier taxation. The aggregate fiscal accounts point in the opposite direction. Revenue data from the same colonial sources show that the bulk of colonial revenue was raised in southern Nigeria rather than in the areas formerly governed by Sokoto (Appendix Figure G5). These aggregate accounts do not measure the effective burden faced by every locality, but they provide no indication that the Sokoto side was systematically more heavily taxed.

Tax Administration. Another possibility is that residents of former Sokoto encountered a weaker or more abusive colonial tax administration. We examine both the content of the archival record and an independent measure of administrative presence. 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). Administrative presence measured independently from the archive yields a similar conclusion: 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). The resistance gap therefore does not appear to reflect a discontinuously weaker colonial bureaucracy.

Tax Enforcement. Resistance could also respond to how aggressively taxes were enforced rather than to their perceived legitimacy. We use the archival text to identify references to state coercion and punishment as well as tax evasion. There is no consistent discontinuity in references to coercion or punishment: the estimates change sign across bandwidths. Ref-

erences to tax evasion likewise show no significant discontinuity at any reported bandwidth (Appendix Table G6). These patterns provide little evidence that systematically harsher enforcement on the Sokoto side accounts for the increase in resistance.

Public Goods Provision. Finally, greater resistance could reflect receiving fewer services in return for taxation. We assemble several independent measures of colonial public investment: road and railroad density from the Surveyor-General of Nigeria (1924) and Okoye et al. (2019); province-level counts of health and education facilities, transport infrastructure, water, sanitation and electricity utilities, administrative offices, police posts, and other government infrastructure from the 1914/16 *Nigeria Blue Book*; medical facilities from Nigeria Surveys (1934/35); and post offices, telegraph stations, and telephone facilities from the Survey Department of Nigeria (1949). Appendix Figures G7–G9 map these measures. We find no discontinuity in public-goods provision at the historical boundary, and controlling for these measures leaves the estimated resistance gap intact (Appendix Tables D2 and G7).

These exercises weaken a broad class of explanations for why colonial taxation might have generated more resistance in former Sokoto. We find little evidence that locations across the boundary faced systematically different populations, statutory tax regimes, fiscal burdens, administrative capacity, enforcement, or public-goods provision. More importantly, such explanations do not naturally account for the composition of the resistance we observe: the additional complaints are disproportionately associated with Muslims, partly concentrated on a tax explicitly rejected under Sokoto's Islamic fiscal order, and more likely to contest the rightfulness of taxation itself. The evidence therefore favors an interpretation in which the relevant historical difference was not simply exposure to Sokoto rule, but the displacement of the fiscal authority that rule had established.

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 7 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.

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

Table 7: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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

Afigbo, A. E., *The Warrant Chiefs: Indirect Rule in Southeastern Nigeria, 1891–1929*, Longman, 1972. 100

Afolabi, Abiodun, "Tax Revolts in Yorubaland, 1925–1955," *Working Paper*, 2023. 100

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, 18

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

Basri, M. Chatib, Mayara Felix, Rema Hanna, and Benjamin A. Olken, "Tax Administration versus Tax Rates: Evidence from Corporate Taxation in Indonesia," *American Economic Review*, 2021, *111* (12), 3827–3871. 1

Bazzi, Samuel, Gabriel Koehler-Derrick, and Benjamin Marx, "The Institutional Foundations of Religious Politics: Evidence from Indonesia," *The Quarterly Journal of Economics*, 2020, *135* (2), 845–911. 7

_ **, Masyhur Hilmy, and Benjamin Marx**, "Religion, Education, and the State," *The Review of Economic Studies*, 2026, *93* (3), 1494–1535. 7

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

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

_ **and Torsten Persson**, "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 2009, *99* (4), 1218–1244. 1, 7

_ **and** _ , *Pillars of prosperity: The political economics of development clusters*, Princeton University Press, 2011. 7

_ **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

Besley, Timothy J, Torsten Persson, and Christopher Dann, "Pillars of prosperity: A ten-year update," CEPR Discussion Paper DP16256, CEPR 2021. 1, 7

Byfield, Judith, "Taxation, Women, and the Colonial State: Egba Women's Tax Revolt," *Meridians: feminism, race, transnationalism*, 2003, *3* (2), 250–277. 100

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. 3, 20, 24

Cohen, Ronald, *The Kanuri of Bornu* Case Studies in Cultural Anthropology, New York: Holt, Rinehart and Winston, 1967. 28

Conley, Timothy G., "GMM Estimation with Cross Sectional Dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45. 81, 82

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, 18

_ **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

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. 100

_ **and Matthew M. Heaton**, *A History of Nigeria*, Cambridge: Cambridge University Press, 2008. 18

Finan, Frederico, Benjamin A Olken, and Rohini Pande, "The personnel economics of the developing state," *Handbook of economic field experiments*, 2017, *2*, 467–514. 1, 7

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

Hiskett, Mervyn, *The Sword of Truth: The Life and Times of the Shehu Usuman dan Fodio*, Oxford University Press, 1973. 1, 8

Hopkins, A. G., *An Economic History of West Africa*, Columbia University Press, 1973. 100

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. 10, 53

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. 3, 11, 12, 21, 54

Korinek, Anton, "Language Models and Cognitive Automation for Economic Research," NBER Working Paper 30957, National Bureau of Economic Research 2023. 2, 7, 18

Lagakos, David, Stelios Michalopoulos, and Hans-Joachim Voth, "American life histories," Technical Report, National Bureau of Economic Research 2025. 2, 7, 18, 19

Last, Murray, *The Sokoto Caliphate*, London: Longmans, 1967. 1, 3, 4, 5, 8, 9, 12, 21, 29, 30, 31, 52, 54, 81, 83, 84, 85

Lavers, John E., "Islam in the Bornu Caliphate: A Survey," *Odù: Journal of Yoruba and Related Studies*, 1971, *5*, 27–53. 28

Levi, Margaret, *Of Rule and Revenue*, University of California Press, 1988. 1, 7

Lovejoy, Paul E., "Fugitive slaves : resistance to slavery in the Sokoto Caliphate.," 1986. 11

_ , "Revolutionary Mahdism and Resistance to Colonial Rule in the Sokoto Caliphate, 1905–1906," *The Journal of African History*, 1990, *31* (2), 217–244. 12, 25

_ , *Jihad in West Africa during the Age of Revolutions*, Ohio University Press, 2016. 1, 3, 8, 9, 12, 13, 21, 54

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

_ **, Nathan Nunn, James A. Robinson, and Jonathan L. Weigel**, "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom," *Econometrica*, 2017, *85* (4), 1065–1091. 7

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

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

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

Murdock, George Peter, *Ethnographic Atlas*, University of Pittsburgh Press, 1967. 3, 21, 64, 66, 67

Naritomi, Joana, "Consumers as Tax Auditors," *American Economic Review*, 2019, *109* (9), 3031–3072. 1

Nunn, Nathan, "The Long-Term Effects of Africa's Slave Trades," *Quarterly Journal of Economics*, 2008, *123* (1), 139–176. 14

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. 41, 94

Pomeranz, Dina, "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 2015, *105* (8), 2539–2569. 1

Tambo, David C., "The "Hill Refuges" of the Jos Plateau: A Historiographical Examination," *History in Africa*, 1978, *5*, 201–223. 28

Thomas, Martin, *Violence and Colonial Order: Police, Workers, and Protests in the European Colonial Empires, 1918–1940*, Cambridge University Press, 2012. 100

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, 7

Usman, Yusuf Bala, *The Manipulation of Religion in Nigeria, 1977–1987*, Vanguard Publishers, 1987. 100

Watts, Michael, *Silent Violence: Food, Famine, and Peasantry in Northern Nigeria*, University of California Press, 1983. 11

Part I

Appendix

A Additional Historical Background **52**
- A.1 Islamic Taxes Before and During British Rule . . . 52
- A.2 Tax Legitimacy in Pre-Colonial Sokoto . . . 52
- A.3 Frontier Evidence: Ningi and Zuru . . . 53
- A.4 Colonial Nigeria: Key Dates . . . 53
- A.5 The Sokoto Caliphate Cavalry . . . 54
- A.6 The Satiru Revolt, 1904–06 . . . 54

B Data and Measurement Details **54**
- B.1 The National Archives . . . 54
- B.2 Historical Map of Locations . . . 54
- B.3 Document Topics . . . 54
- B.4 Tax Sentiment Prompt . . . 57
- B.5 Tax Issue Prompt . . . 58
- B.6 Distinctive Vocabulary of the Outcomes . . . 58
- B.7 Measuring Complainant Religion . . . 58
- B.8 Jangali Complaint Prompt . . . 60
- B.9 Tax Legitimacy Prompt . . . 60

C Additional Validation **61**

D RD Validity and Robustness **61**
- D.1 Balance at the Boundary . . . 64
- D.2 Ethnic Homelands and Protectorates . . . 64
- D.3 Results Within the Northern Protectorate . . . 64
- D.4 RD Figures in Latitude and Longitude . . . 64
- D.5 Archival Coverage . . . 69
- D.6 Spatial and Population Coverage . . . 69
- D.7 Geocoding Uncertainty . . . 69
- D.8 Border Specificity . . . 69
- D.9 Functional Form . . . 77
- D.10 Inference . . . 81

E Within-Caliphate Exposure **81**

F Additional Mechanism Evidence **83**

G Alternative Explanations, Supporting Evidence **86**
- G.1 Demography . . . 87
- G.2 Tax Instruments . . . 91

G.3 Tax Burden 91
G.4 Tax Administration 91
G.5 Tax Enforcement 92
G.6 Public Goods Provision 92

H Tax Revolts and Riots **92**
H.1 Data Construction 100
H.2 Historical Sources 100
H.3 Results 100

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 6.3 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.

B.2 Historical Map of Locations

Appendix Figure B2 shows the geolocated document locations on the 1924 map of Nigeria, together with the Sokoto Caliphate boundary.

B.3 Document Topics

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

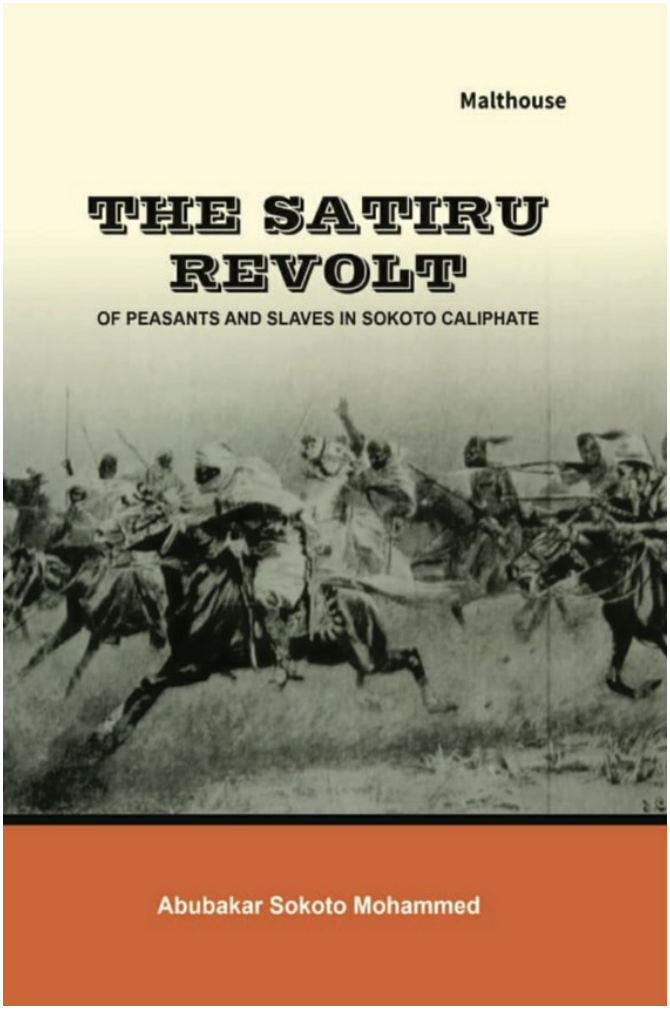






Figure A2: The Satiru Revolt, 1904–06







Figure B1: The National Archives, Kaduna

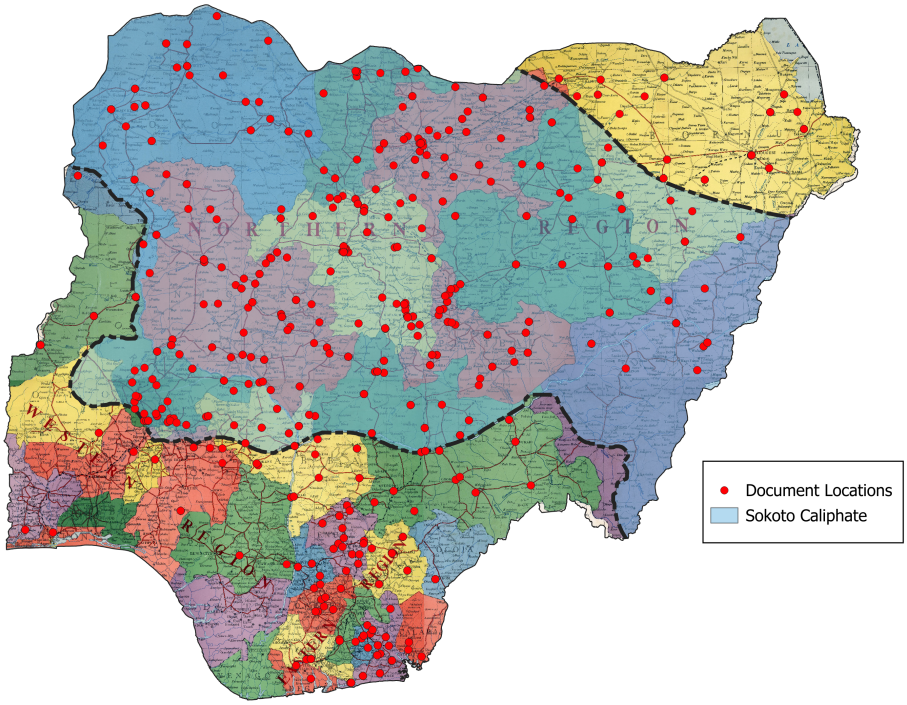


Figure B2: Document Locations on the 1924 Map of Nigeria

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

- 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 3.2.3 is produced by the following prompt, reproduced verbatim.

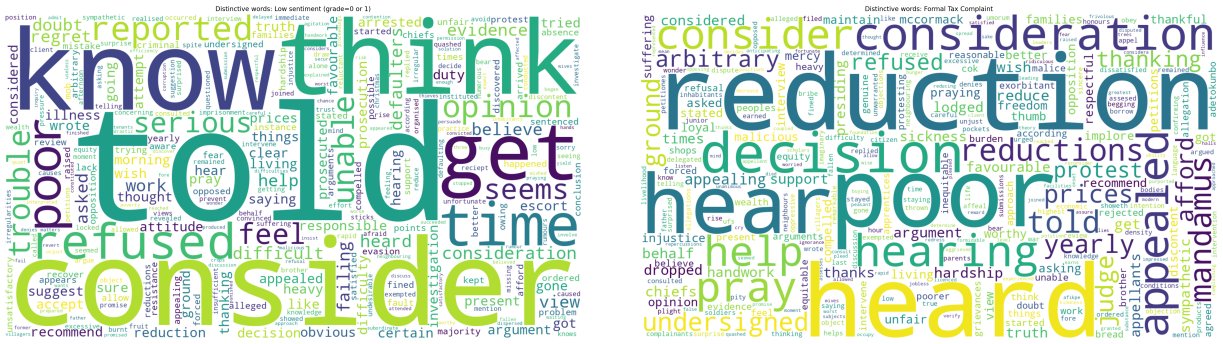
> You are a meticulous historical analyst. Your task is to read a document from colonial Nigeria and identify all described issues related to taxation, from administrative problems to open rebellion. A conflict over a chieftaincy dispute or a report on road building are NOT tax issues. You must link 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 B3 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 and 3.2.3.

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 B3: 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 6.3 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 6.3 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 6.3. 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.

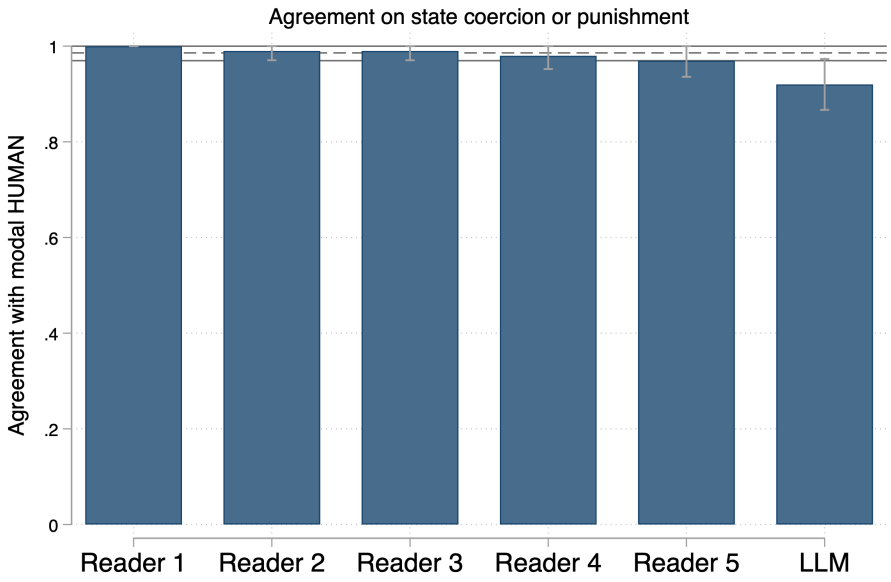


Figure C1: Turing Test: State Coercion or Punishment

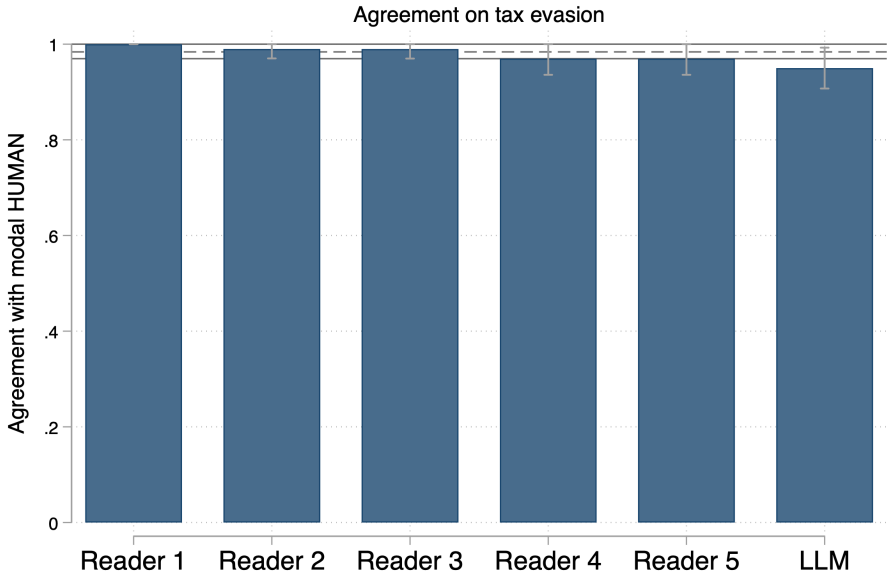


Figure C2: Turing Test: Tax Evasion

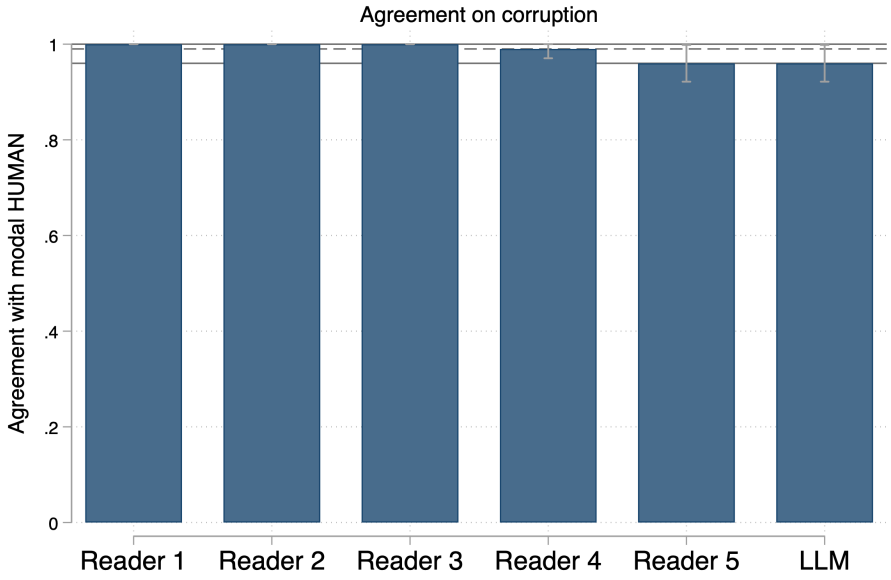


Figure C3: Turing Test: Corruption

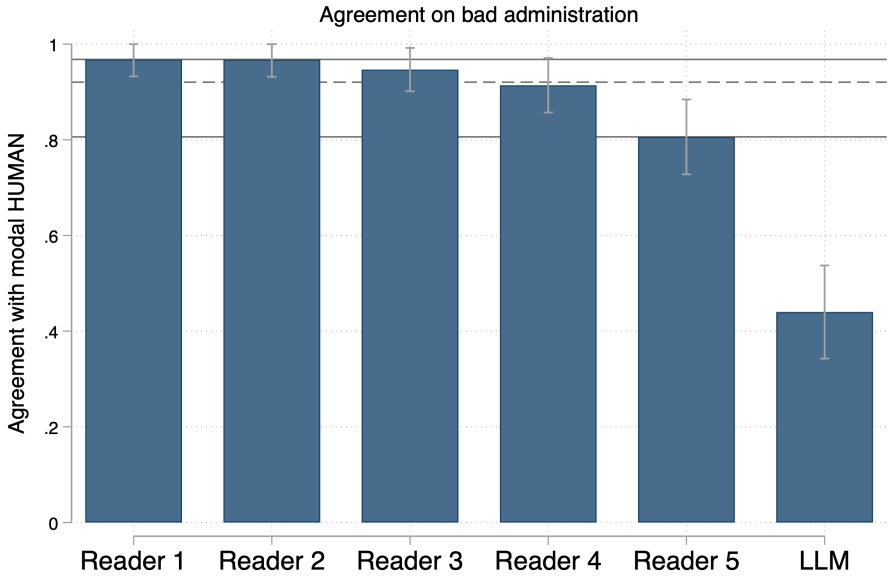
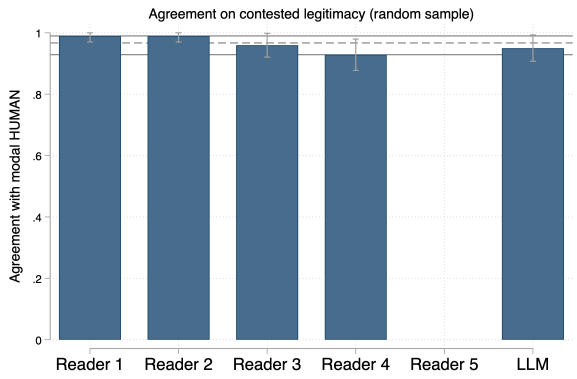
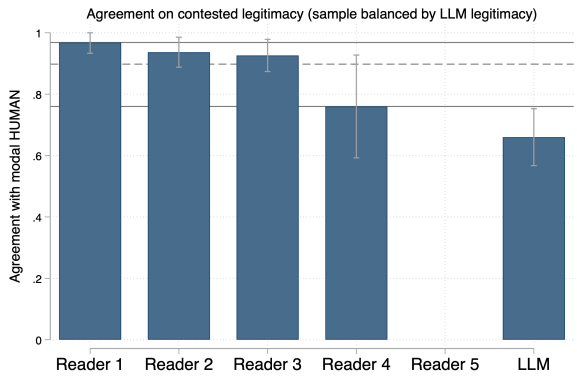


Figure C4: Turing Test: Administrative Failure



(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 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 D3 shows the RD results with ethnic-homeland fixed effects.

D.3 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 D4 shows that the results hold within the Northern Protectorate, where the colonial tax code was uniform.

D.4 RD Figures in Latitude and Longitude

Appendix Figure D2 reproduces Figure 9 with the RD surface estimated linearly in latitude and longitude.

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

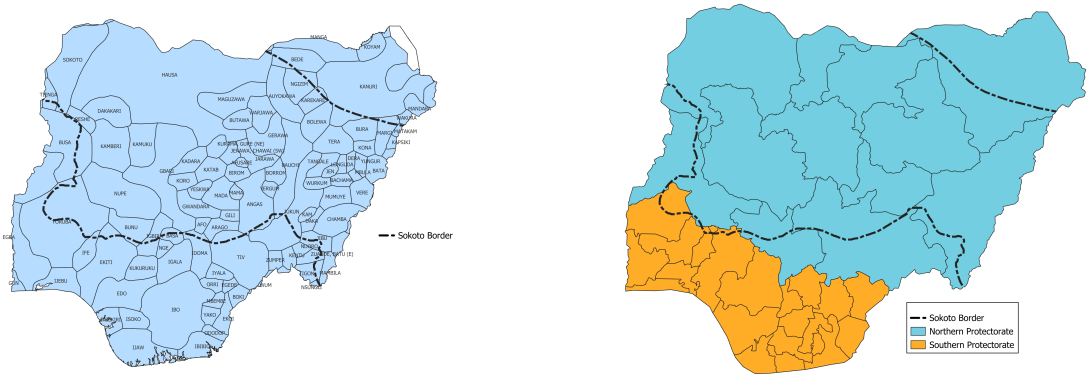
Table D2: Balance at the Boundary: Colonial Controls

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

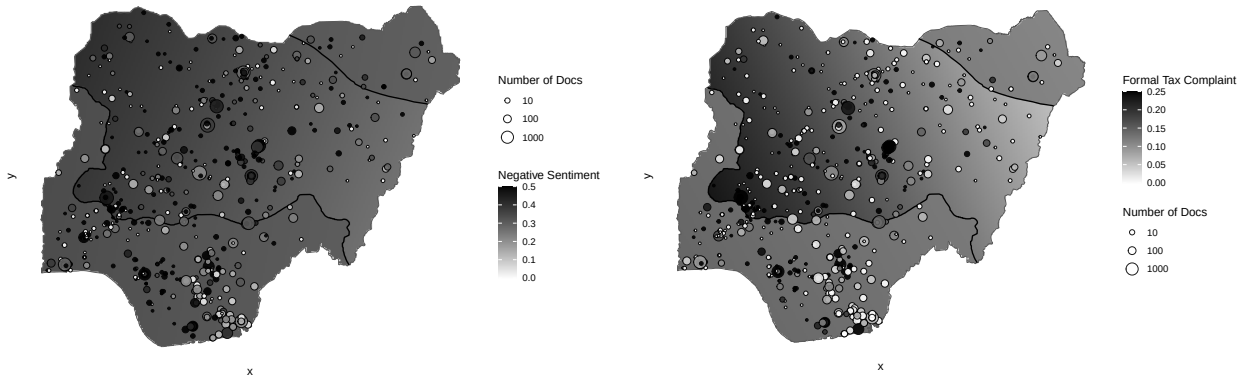
Table D4: RD Results Within the Northern Protectorate

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



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

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

D.5 Archival Coverage

Appendix Tables D5, D6, and D7 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.6 Spatial and Population Coverage

Appendix Table D8 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 D9 reweights each observation by the inverse of its location's document count, so that every location receives equal weight; Appendix Table D10 reweights by inverse division-level population from the 1952 Census, so that every person receives equal weight.

D.7 Geocoding Uncertainty

Appendix Table D11 reports a donut RD that drops potentially ambiguous locations by excluding a 10-kilometer buffer around the border.

D.8 Border Specificity

Appendix Table D12 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 D13 addresses the Jos Plateau, excluding Jos Plateau observations in Panel A and carving the plateau out of the Sokoto boundary in

Table D5: 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 D6: 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 D7: 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 D8: 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 D9: 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 D10: 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 D11: 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 D12: 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.9 Functional Form

Appendix Table D14 replaces the local linear function of distance with a quadratic; Appendix Table D15 uses an RD polynomial linear in latitude and longitude.

Table D13: 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 D14: 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 D15: 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.10 Inference

Appendix Table D16 clusters standard errors at the province level; Appendix Table D17 uses spatially correlated standard errors following Conley (1999) with a 25-kilometer cutoff.

Table D16: 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 D17: 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 3. 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 3.

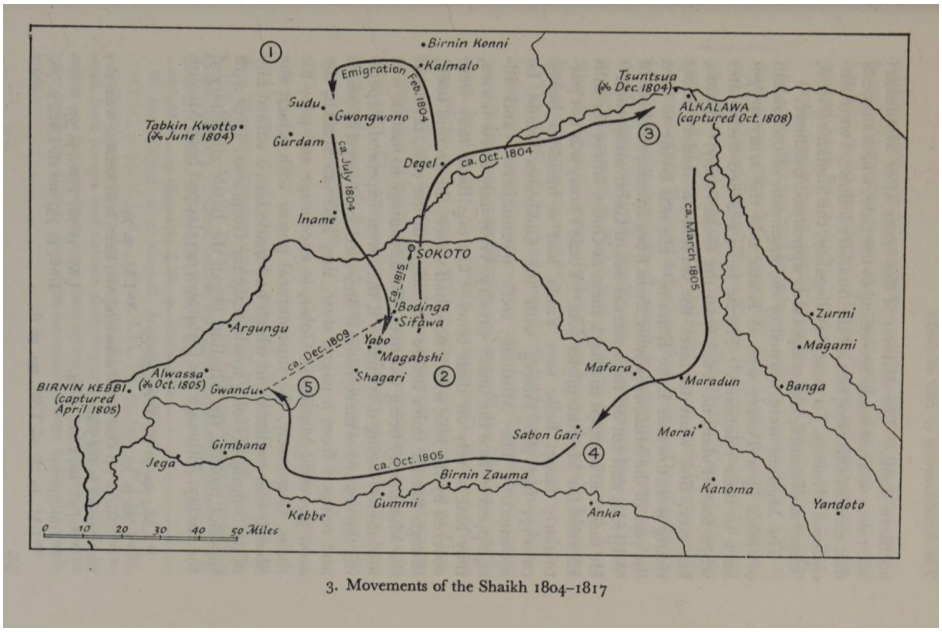


Figure E1: Movements of Usman dan Fodio, 1804–1817

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

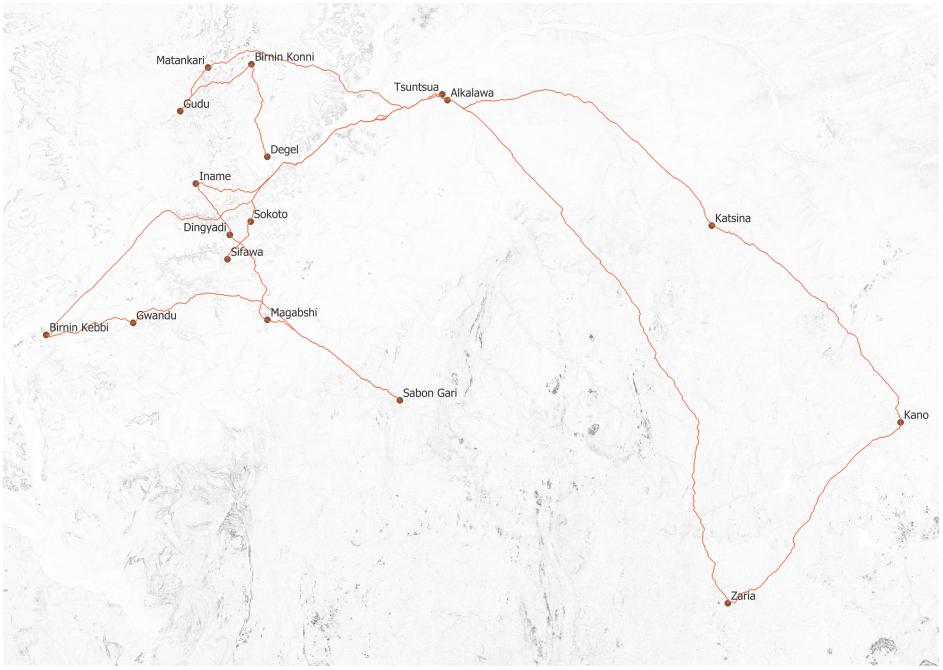
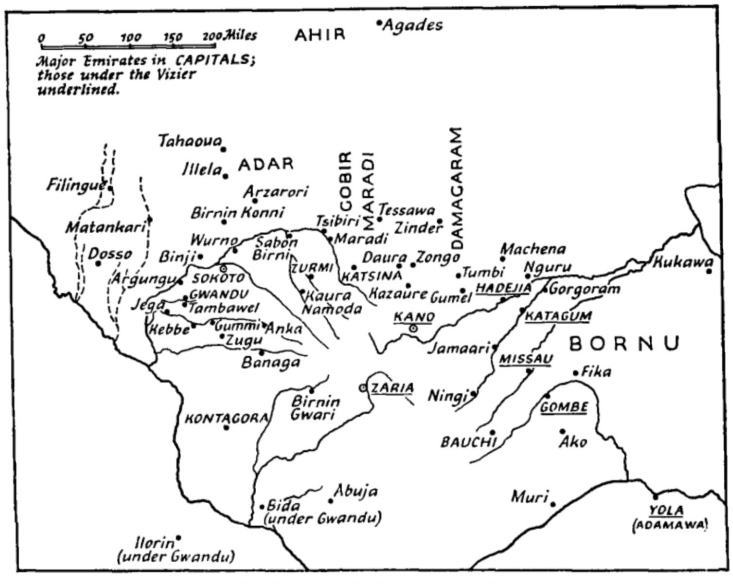


Figure E2: Movements of Usman dan Fodio: Least-Cost Path

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).



9. The Sokoto Caliphate *c.* 1890

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 3.

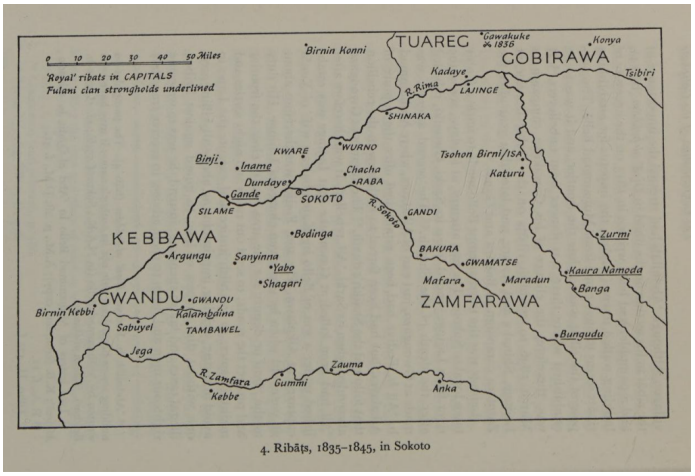


Figure E4: Royal Ribats, 1839–1845

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

to high- and very-high-confidence legitimacy classifications; Table 6 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 6.4.

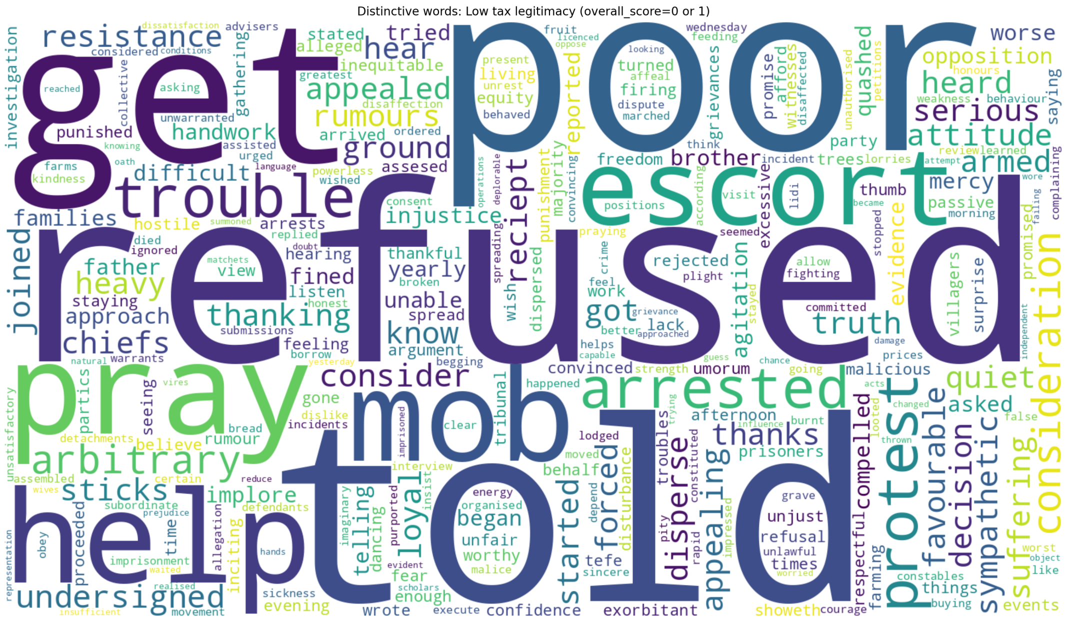
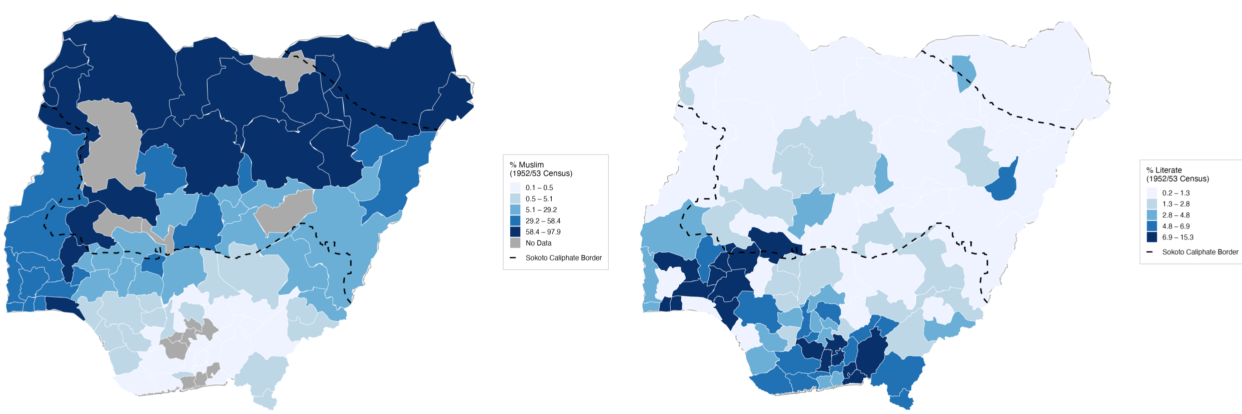


Figure F1: Word Cloud: Contested Legitimacy

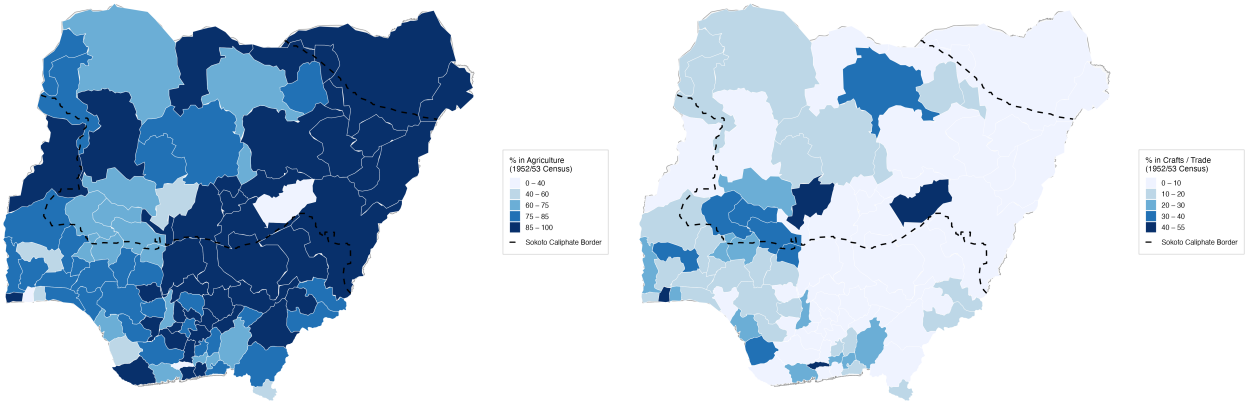
G.1 Demography

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)

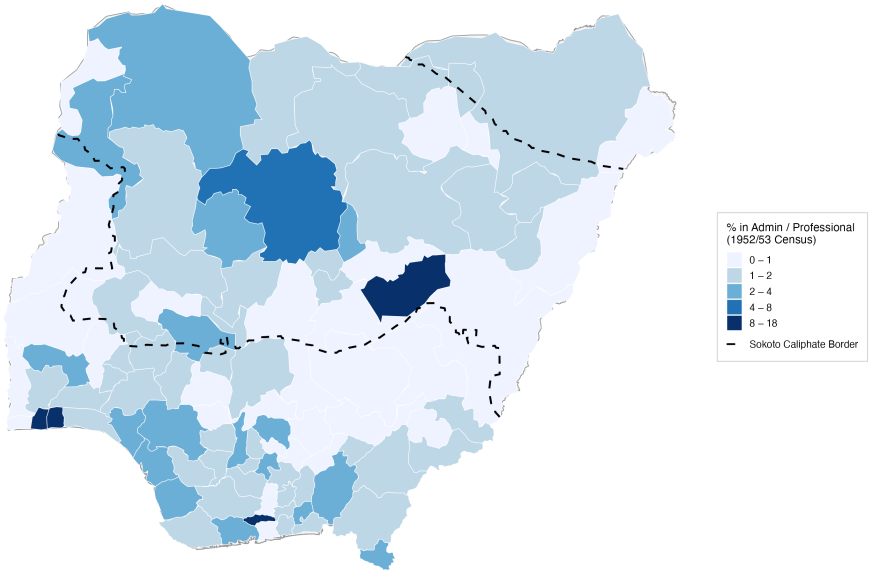
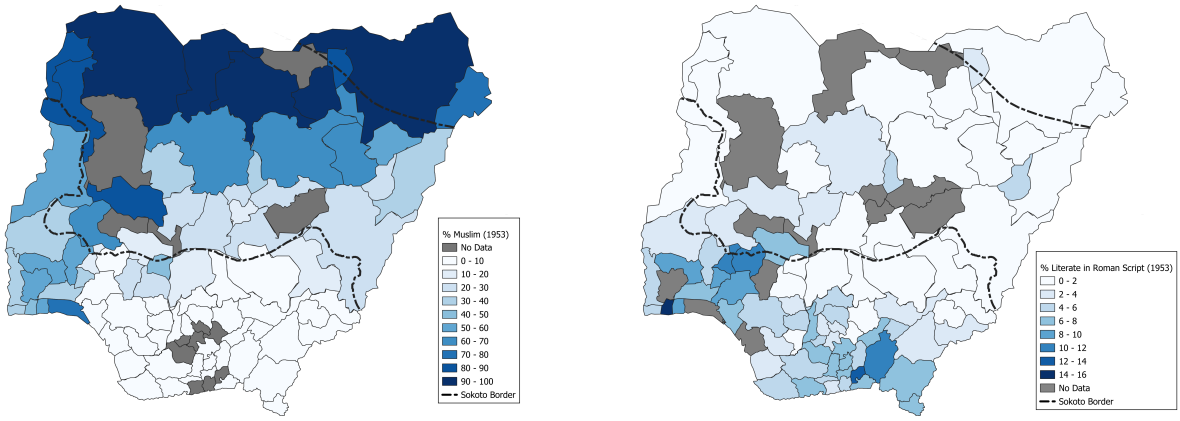


Figure G3: Occupation Shares (1952/53 Census), Administration



(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 6.4. Appendix Figure G6 maps civil servants per 1,000 population from

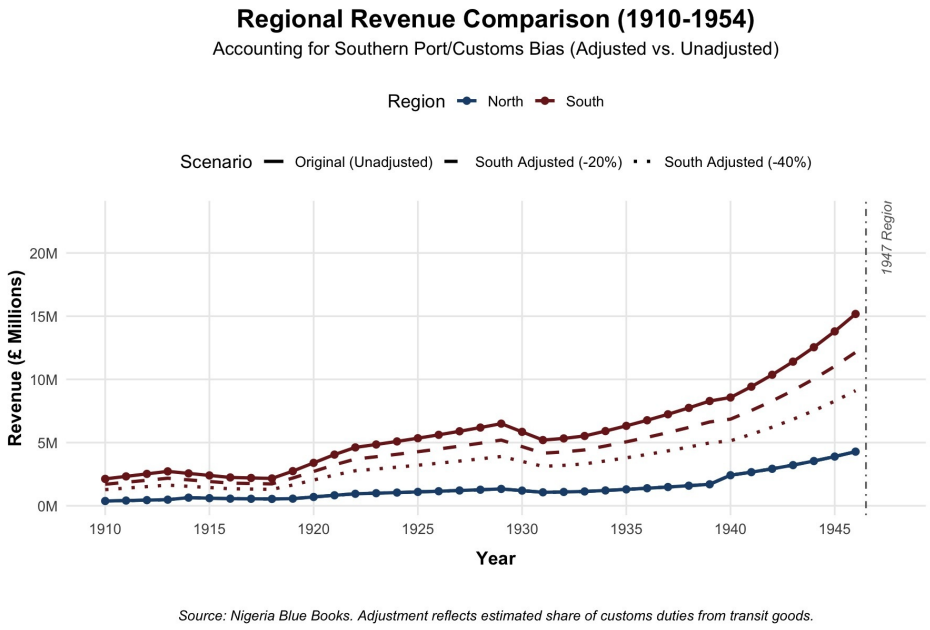


Figure G5: Tax Revenue: Northern vs. Southern Protectorate

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 6.4.

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

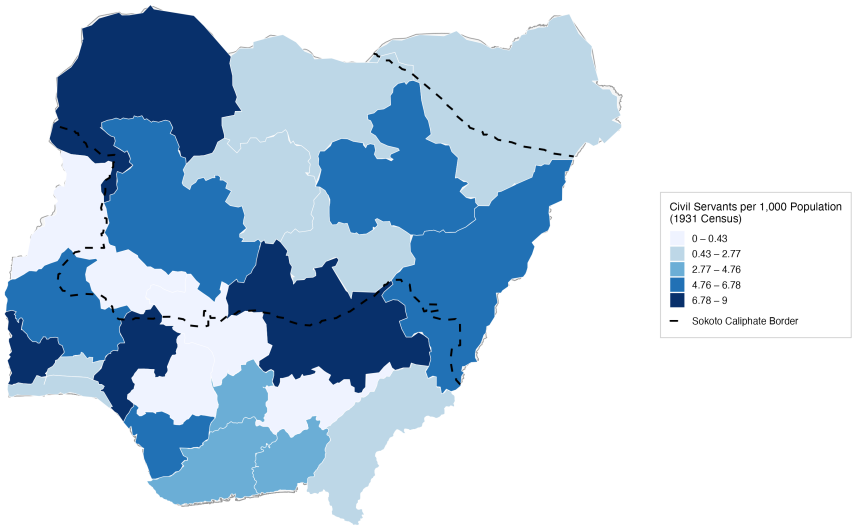


Figure G6: Civil Servants per 1,000 Population (1931 Census)

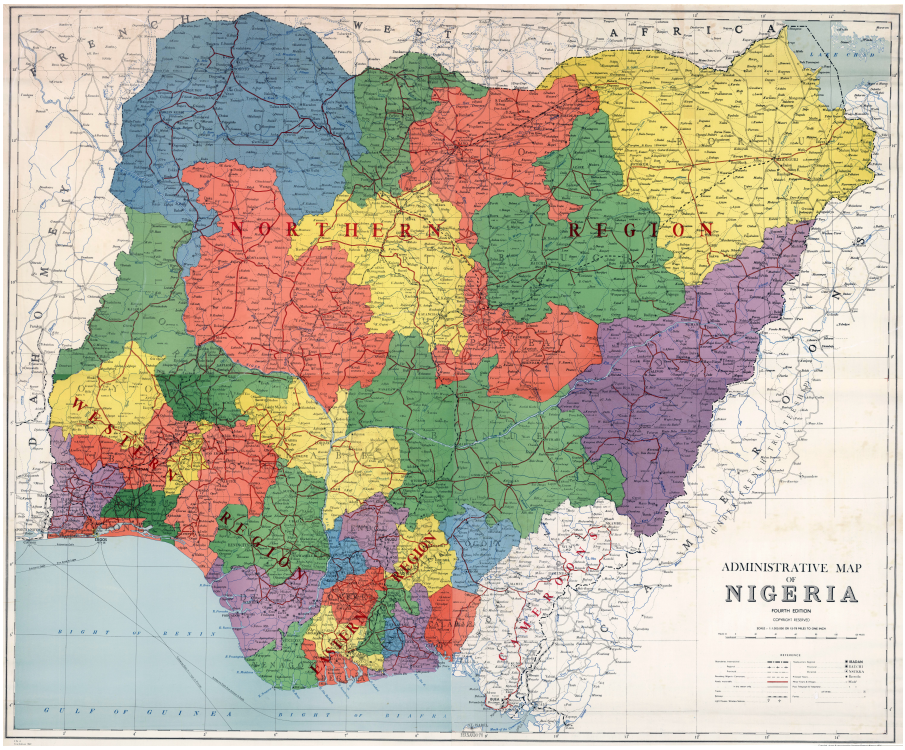


Figure G7: Roads and Railroads, 1924

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*; Hopkins (1973), *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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.