Packing or Cracking Ethnic Groups? On the Colonial Design of Administrative Geographies in Sub-Sahara Africa

Unpublished Working Paper

Carl Müller-Crepon*

July 8, 2024

Abstract

The design of subnational administrative units is fundamental to the functioning of states and their political topography. Yet, we lack theoretical and empirical knowledge on the determinants of the partitioning of state territories into subnational governance units. I argue that aligning administrative borders with ethnic geography facilitates local institutional persistence and increases short term governance efficiency. Yet, because such "packing" of units stabilizes peripheral institutions and identities, some governments may instead design heterogeneous units and "crack" groups and their institutions to centralize power. I analyze these arguments by studying unit-design in Sub-Sahara Africa, contrasting indirect with direct forms of colonial rule using new data on administrative and early-colonial ethnic geographies derived from historical maps. Modeling administrative borders with a probabilistic spatial partition model, I find evidence that administrative borders are strongly and positively associated with ethnic boundaries. These effects are more pronounced under indirect British compared to more direct French rule and driven by their packing of uncracked groups into administrative units. With novel theory, data, and methods, this paper contributes to our understanding of the varying ethno-geographic roots of administrative geographies in Africa.

Keywords: Administrative geography; Ethnicity; Colonialism; African politics; Computational Methods; GIS

---

*Department of Government, London School of Economics and Political Science. c.a.muller-crepon@lse.ac.uk. I thank Sören Henn, Mateo Montenegro and Yannick Pengl, as well as seminar participants at LSE, IPERG at the Universitat de Barcelona, and at the 2024 Interdisciplinary Workshop on Identity and Politics at IAST/TSE and the 2024 EPSA Conference for very helpful comments and Elise Doumerge, Victoria Puglia, Jakob Schram, Jonathan Seel, and Daryna Sudachek for excellent research assistance. This project has received funding from the British Academy through grant SRG21\210081.

The design of administrative units is fundamental to states' political geography. Where administrative borders are drawn determines the number, shape, and demography of subnational governance units. The choice between few or many, ethnically homogeneous or diverse administrative units is far from innocuous. It affects, for example, local development (Alesina and Zhuravskaya 2011; Grossman, Pierskalla and Dean 2017), state capacity (Henn 2023; Müller-Crepon 2021), as well as violent conflict (Cunningham and Weidmann 2010; Juon 2024). Given these stakes, the design of regions and districts is inherently political. Yet, the question of why administrative units are shaped as they are is largely absent from the empirical political science literature, which mostly treats them as exogenously given (Soifer 2019). This paper addresses this gap with a focus on the ethnic underpinnings of colonial administrative geographies in Sub-Saharan Africa. Studying administrative unit design with new data and methods, I contribute to understanding the origins of political topographies, in particular in Africa.

A small literature examines administrative geographies as politically determined, focusing primarily on their proliferation. Following up on Green (2010), this literature examines the political drivers of unit splits along preexisting lower-level borders (see also Grossman and Lewis 2014; Hassan 2016; Resnick 2017). However, the literature's focus on relatively marginal changes of administrative geographies leaves unexplored the causes of the more fundamental *overall* partitioning of state territories into administrative units.[1] Two main factors can account for this oversight. First, we lack a political theory of administrative unit design. Second, any test of such theory must overcome the empirical difficulty of modeling the determinants of administrative partitionings as neither the number nor shapes of units are known ex ante. I address the first problem by theorizing administrative divisions as strategic choices made in response to the geography of pre-existing ethnic institutions. Newly collected data analyzed with a recently developed probabilistic spatial partition model allow me to empirically test that argument's observable implications.

[1]Note that the design of administrative borders in the territory of one state follows a different logic than that of borders between states or colonies, in particular when ruled by different colonizers. For recent evidence on the drawing of colonial state borders in Africa, see Paine, Qiu and Ricart-Huguet (2024).

This paper argues that a government can draw administrative borders to strategically "pack" ethnic groups and their pre-existing institutions into homogeneous units that are drawn to leave their settlement areas undivided or "uncracked". This facilitates their continuing function and indirect rule through them, but decreases the government's ability to centralize power and governance. This ability can be enhanced by "cracking" ethnic groups and their institutions into diverse, "non-packed" administrative units. Thus weakening local elites and undermining pre-existing institutions, doing so requires rebuilding effective local administrations but facilitates the consolidation of political power in the center. As a result, only governments with a preference and capacity for centralized governance will implement such disruptive designs. In turn, governments that seek or have to rely on decentralized governance draw administrative borders that leave groups packed and uncracked by administrative borders.

My empirical analysis tests this argument with a focus on colonial rule in Sub-Saharan Africa and with ethnic geography as a proxy for the political topography of pre-colonial socio-political institutions. In particular, I compare the effect of ethnic geography on administrative border designs under relatively decentralized, indirect British rule with those under more centralized, direct French rule. To do so, I combine three innovations in empirical measurement and methods.

First, I measure administrative partitionings and their change as the main outcome of interest. I therefore collect a temporal panel of administrative maps from across British and French colonies throughout the colonial period. These show the colonial evolution of territorial governance from few, imprecisely designed administrative units towards the late-colonial setup of districts, cercles, and regions which partially persisted until today. In additional analyses, I study the latter process using data on administrative regions over the decades since states' independence.

Second, I present new data on historical ethnic geography which remedy the low resolution and likely reverse causality affecting the classic data on ethnic settlement areas from George Murdock (1959). The new measure of ethnic geography is based on 49 newly digitized ethnic maps produced in the first half of the 20th century. While less comprehensive in coverage of the African continent than Murdock's (1959) map, these data feature much greater spatial detail, approximation of

local ethnic diversity, and reflection of uncertainty over relevant ethnic groupings. The fact that some maps stem from the early colonial period allows me to investigate subsequent administrative border *changes*, thus addressing the risk of reverse causality.

Third, I model the effect of ethnic geography on administrative borders using a Probabilistic Spatial Partition Model (PSPM) developed by Müller-Crepon, Schvitz and Cederman (2024). The PSPM models administrative borders as the partitioning of a spatial network of points into regions. Encoding covariates on the network's edges, the model allows for estimating the effect of ethnic boundaries on administrative borders conditional on spatial features (e.g., rivers or watersheds) that may cause ethnic and administrative geographies. Moving beyond the original setup of the model, I present a new set of estimators that allows for directly estimating the effects of macro-level strategies of ethnic *packing* and *cracking* on administrative partitionings. The extended model allows me to empirically separate these two correlated, yet theoretically and empirically distinct strategies.

The empirical results support the main argument. I find that ethnic boundaries are significantly associated with a substantively higher probability of district borders and border change in British colonies. Ethnic boundaries are weakly associated with the borders of French colonial cercles and do not explain their change over time. These results are robust to the inclusion of a lagged dependent variable which accounts for important sources of reverse causality bias and to a wide range of permutations of the data and model setup. Drawing on post-colonial data on administrative borders, I find that the more extensive ethno-geographic roots of administrative units in former British as compared to French colonies persist until today. Lastly, I distinguish between strategies of ethnic "packing" and "cracking" and find that British and French colonial governments ethnically packed their districts. Yet, British rulers drew internal borders that left ethnic groups significantly less cracked than French administrations.

Continuity or change in administrative unit designs

With the exception of the literature on administrative unit proliferation, political scientists for the most part treat administrative divisions as relatively stable political institutions.[2] Yet, once understood as the outcome of a set of political choices, questions relating to strategies of border drawing surface. In theorizing decisions over the partitioning of states' territories into administrative units of an ex ante unknown number and shape, I start by highlighting the dual character of administrative divisions which, like other instruments of infrastructural power, can foster state rule but can also empower local elites. As a result of this tension, the balance of power between the central government and local elites affects the drawing of administrative borders: Where local elites enjoy a power advantage, administrative divisions are aligned with their spheres of interests. Where states are stronger, units will be designed to break pre-existing institutions and pit them against each other. While empowering the state in the long-run, such misalignment is costly in the short-run as it disrupts local governance arrangements.

The main point of departure is the claim that the spatial decentralization of administrations is a "dual use" governance tool: it aids the state to bridge the spatial gap between its center and the population, yet can also foster collective action of populations and elites in the periphery. The extent to which decentralization empowers the center or periphery depends on many factors, such as selection and appointment mechanisms of local governments and administrations, their powers over budgets and policy making, and control over local public services (e.g. Oates 1999, and related literature). In addition, decentralized governance is *structurally* predisposed to serve either central or peripheral interests depending on the congruence of its border with the geography of local political topographies. I focus on this last factor.

The creation of administrative outposts is a crucial instrument by which states'

[2]While there is a (small) literature on the size of shape of *states* (e.g. Alesina and Spolaore 2005, 1997; Friedman 1977; Müller-Crepon, Schvitz and Cederman 2024), the theoretical starting point of this literature differs as international borders emerge between states rather than as an outcome of intra-state, center-periphery interactions. However, the reader may note that the origins of subnational borders will indirectly affect the locations of new interstate borders as these often follow pre-existing administrative ones (Carter and Goemans 2011).

increase their "infrastructural power" (Mann 1984). Decreasing the physical distance between state agents and the population, administrative decentralization facilitates greater and more spread-out control of, extraction of taxes from, and provision of services to citizens. As a result, most states' administrations are spatially organized with hierarchically nested administrative tiers reaching down from the center to local populations.[3] While there are decreasing and at some point, negative returns to fragmenting administrative geographies (Grossman, Pierskalla and Dean 2017), larger states tend to have more administrative units arranged in more extensive hierarchies.

At the same time, a large literature highlights the promise of administrative decentralization to empower local populations, by increasing citizens control over and knowledge of local policies thus better aligning demand- and supply-sides of governance (Tommasi and Weinschelbaum 2007; Grossman, Pierskalla and Dean 2017). Such benefits materialize in particular where collective action problems are more easily solved: in units that are spatially congruent with social institutions in general (Wilfahrt 2022) and ethnically homogeneous ones in particular (Alesina, Baqir and Easterly 1999; Miguel and Gugerty 2005). Beyond mobilizing for better public services, the capacity of peripheral collective action can be used to counter the central government more broadly. Some have, for example, linked the design of ethnically delineated "proto-states" to successful secessions (Roeder 2012; Griffiths 2016), lower political stability and generalized social trust, as well as stronger ethno-political identities (Alesina, Easterly and Matuszeski 2011; Robinson 2020; Müller-Crepon 2024).

As a result of the ambiguous effect of administrative borders, we can delineate two distinct administrative designs of administrative units: one enabling a continuing exertion of power by local institutions and elites, and the other breaking them by disruptive spatial misalignment.

The first strategy aligns administrative units with the geography of pre-existing informal and formal institutions of political control. In particular when coupled with the cooptation of local elites, designs that "pack" relatively homogeneous administrative divisions with unaltered pre-existing institutions and associated so-

[3] City-states are exceptional for their non-territorial character.

cial networks allow states to quickly and effectively build their reach on the cheap through cooptation. This comes at the cost of the need to share power and revenue with local institutions and their elites whose local entrenchment fragments states' institutional landscape and hampers state-wide governance.

More disruptive, but expensive, are conscious design of administrative geographies contra pre-existing institutions and social networks. In an attempt to break the local powers that are, governments can impose administrative borders that "crack" preexisting institutions such that they are spread across multiple units. This disrupts their internal organization and mobilization capacity. This effect can be increased by additionally creating non-packed administrative units that include parts of several pre-existing institutions. The resulting units will be internally divided, which further reduces their suitability for the organization of collective action and increases central governments' ability to rewire local society in the long-run.[4] Yet, the promise of breaking pre-existing institutions comes at the potentially steep short-term cost of local resistance and inefficiencies in local governance which has to make up for social fragmentation.

When do governments, then, pack units with pre-existing institutions and when do they make them compete in diverse units? And when do they crack institutions instead of trying to preserve them? As already hinted at, the design of administrative borders has to be understood in the broader context of institutional choices underlying local governance arrangements. Administrative "packing" of un-"cracked" predecessor institutions is congruent and indeed complementary to strategies of indirect rule – the cooptation of preexisting political institutions to build state reach works best if their territorial expanse remains unaltered. Any imposed territorial discontinuities lead to a change in local governance arrangements, contravening the indirect ruler's credo of "if it ain't broke, don't fix it" (Gerring et al. 2011, p. 385). In turn, direct rule involves breaking and replacing local governance arrangements. Disruptive administrative designs of diverse, non-"packed" divisions and those that "crack" pre-existing institutions are again complementary. Akin to dynamics of "divide and rule", they produce political tensions or even

[4]Writing on interstate borders, Englebert, Tarango and Carter (2002) call these two dimensions "dismemberment" and "suffocation".

conflict between and within preexisting political institutions that facilitate the imposition of rule by the center.[5]

Ethnic roots of colonial administrative unit designs

I turn to British and French rule over large parts of Sub-Saharan Africa to investigate my argument that administrative unit designs under indirect rule pack pre-existing institutions into comparatively homogeneous units, while direct rule comes with administrative borders designed to crack them into more heterogeneous units. The following contextualizes my theoretical argument in the relevant historical literature. In particular, I discuss the spatial character of pre-colonial political institutions and related ethnic identities, their encounter with rigid European conceptions of bounded territoriality, and differences between more indirect British and direct French colonial rule.

Understanding the historical roots of ethnic administrative unit designs in Africa is not only substantively important for contributing to understanding colonial strategies of rule and the enduring yet varying political importance of ethnic cleavages (Ali et al. 2019). The continent also promises three advantages over studying administrative border design elsewhere. First, territorial statehood was imposed by the colonizers on the continent comparatively late at the turn of the 19th century with the median colonial state border not settled until 1906 (Paine, Qiu and Ricart-Huguet 2024). This implies that we can study the drawing of administrative borders as a governance revolution occurring over a few decades rather than as the longer process it was elsewhere. Second, while local actors clearly constrained and shaped colonial policies, the general tendencies of British colonialists towards more indirect and the French ones towards more direct rule was likely not due to any particular set of interactions within individual colonies but originated in the empires' ideology and overall capabilities. This is, again, different to many other historical examples of the shift from indirect to direct rule which resulted

[5]What does, then, account for the choice between direct and indirect rule? While classical account focus on external warfare as driver of centralization (Tilly 1990, ch. 4), factors internal to governments, such as their capacity and ideology, as well as local factors likely play a role too (e.g. Müller-Crepon 2020).

from strategic center-periphery interactions (Tilly 1990; Hechter 1975). Third, the comparative overall haphazardness with which Africa was partitioned into empires (Herbst 2000)[6] facilitates the study of administrative borders within colonial empires which may face complications where state borders are endogenous to historical subnational divisions (Carter and Goemans 2011).

While studying unit designs in Sub-Sahara Africa comes with important empirical advantages,[7] I expect the theoretical argument to capture crucial patterns of the drawing of subnational units elsewhere. For example, post-unification Germany in 1871 remained administratively divided along the borders of the high-capacity former kingdoms in the South. This contrasts with the redrawing of administrative borders in post-unification unitary Italy which did not follow pre-existing political borders (cf. Ziblatt 2004, 2006). A similar contrast consists in the drawing of new borders in post-revolution France and Russia. The former case featured an almost grid-like homogeneous design with few compromises. In contrast, substate borders in Russia under Lenin cut across pre-existing units but were roughly aligned with the prevailing ethnic geography (e.g., Hirsch 2000).

Political topographies, territoriality, and the introduction of borders

Defining the state and its subordinate administrative divisions via its territory demarcated by borders is integral to our contemporary idea of statehood (e.g., Mann 1984; Weber 1919). Yet, the use of widthless lines to territorially bound political entities is a fundamentally modern phenomenon that "was virtually unknown in most places in Africa during the period before the European partition" (Asiwaju 1983, p. 45). Instead, frontier zones characterized the spaces between the political cores of states where they existed, mirroring the absence of interstate competition over abundant and sparsely populated territory (Herbst 2000).

Not dissimilar to premodern Europe (Ruggie 1983), precolonial rule oftentimes featured spatially overlapping and non-aligned jurisdiction over territory and people, with control over people being more valuable than control over ter-

[6] But see recent evidence by Paine, Qiu and Ricart-Huguet (2024) that the ultimate location of colonies' borders was often shaped by local interests and geographic conditions.

[7] Note that there are also disadvantages, in particular in data availability.

ritory (Herbst 2000; Wilks 1975). In practice, this implied strong political control over the core areas of states and only irregular and weak reach into their peripheries. The power of rulers was thus radiating outwards, gradually ranging from tight control in the center, to tributary rule in an intermediate range, to the mere ability to raid in peripheries. The rarity of precolonial maps only underscores the absence of demarcated boundaries (Herbst 2000). Evidently, the concept of territorially bounded political entities was even more foreign to acephalous societies where political power was not centralized. After all, the political boundary comes only to life as a separating line between political entities (Kristof 1959).

The prevailing diversity of political topographies changed rapidly with the colonial conquest in the late 19th and early 20th century. Not only did the European conquerors carve up the continent into empires and colonies, but partitioned their colonies into administrative units to create the basic infrastructure needed to establish "effective control." Thus shifting from rule over people to territorial rule, the elements of the hierarchical governance chains – the "thin white line" (Kirk-Greene 1980) – presided over whoever happened to reside in their territorial unit. The creation of non-overlapping and neatly bounded administrative division – regions, districts, and subdistricts – was as much of a governance revolution as the drawing of international borders. Both replaced the precolonial variety of territorial governance arrangements with inflexible, sharp lines that extended and delimited the territorial scope of states and their subnational entities across the entire African landmass (see Asiwaju 1983).

To make matters worse, the introduction of the concept of bounded, non-overlapping units was not limited to political entities but extended to the European conceptualization of ethnic geography. Historiographies of colonial rule describe the predominant mindset as expecting individuals to be nested in tribes, "discrete, bounded groups, whose distribution could be captured on an ethnic map" (Young 1985, p. 74). Incidentally, this understanding of ethnic geography – and the production of ethnic maps based on it – was intimately linked to the dominant European ethno-nationalist discourse that colonizers brought with them (Berman 1998).

Yet, the idea of geographically fixed, bounded, and non-overlapping groups met a reality where multiple groups often settled in the same environmental niche,

as for example pastoralists and sedentary agriculturalists inhabiting the same savanna (Cohen and Middleton 1970, p. 11). Even the identification of groups themselves sometimes proved difficult, in particular where identities and associated loyalties were "complex, flexible and amorphous, sometimes overlapping, sometimes complementary, and did not add up to clearly demarcated tribes" (Lentz 1995, p. 317, Southall 1970). Consequently, ethnic groups did not correspond neatly to political entities, which in turn often included ethnic minorities among their populations (e.g. Colson 1960; Wilfahrt 2022).

Yet even though ethno-linguistic groups, understood fuzzily, did not correspond perfectly with political institutions, I will in the following treat the former as theoretical and empirical proxies of the latter. This is for three reasons. First, the anthropological literature shows that cultural groups oftentimes come with distinct political institutions, in particular in cases without centralized political authorities (e.g. Fortes and Evans-Pritchard 1940). Even in cases where pre-colonial states included ethnic minorities that differed from the group group of the ruling elites, minorities oftentimes enjoyed special political rights (e.g. Colson 1960; Wilfahrt 2022), which indicates institutional differentiation. Second, the use of ethnic identities as proxies for institutions follows closely the dominant mindset of colonial governments which perceived them just as that, thus constituting a valid way of approximating *their* decision making. The third, more pragmatic argument is that good-resolution data on the spatial extent of (types of) political institutions does not exist independent of data on ethnic geography beyond relatively coarse data on the small set of pre-colonial states (Paine 2019; Paine, Qiu and Ricart-Huguet 2024; Wilfahrt 2022).

Strategies of local rule: British versus French styles

In particular for British colonies, there is ample case study evidence that colonial administrators often aimed at drawing district boundaries along the 'tribal' boundaries they perceived. Because the dominant strategy of indirect rule declared 'tribes' as 'natural' social units of local governance, tribal homelands were destined to become administrative units (e.g., Asiwaju 1970; Crowder 1968; Miles

1994; Spear 2003).

While the British application of indirect rule is well known, its impact on administrative unit design is not precisely documented. Writing on the internal borders of the British Gold Coast, today's Ghana, Lentz (2006, p. 53) notes that the colonial government was able to make administrative borders roughly but not fully consistent with the prevailing, complex, and fuzzy ethnic geography. Instead, pragmatism coupled with administrative and geographic exigencies ultimately determined the precise location of borders (see also Howard 2005). Yet, Sharpe (1986) describes how local chronicles were used in Northern Nigeria to determine ruling elites, their groups, and delimit administrative areas. Evidencing similar bottom-up processes, local populations and elites in Ghana at times successfully mobilized for border change that would increase the alignment between social and administrative geographies (Bening 1999). While Müller-Crepon (2020) notes that British rule was less direct in pre-colonially acephalous groups without centralized institutions due to creation of new subnational governments, the resulting governance arrangements where nevertheless ethnicized.

The historical case of French colonies is different for their more centralized governance. Although French administrations relied on local intermediaries as well, they tended to crush pre-existing institutions, replace them with more uniform institutions of their own making, and hand less power to local rulers (e.g., Cohen 1971; Conklin 1997; Crowder 1964; Müller-Crepon 2020). This strategy of comparatively direct rule likely originated in the higher level of administrative capacity available to the French and the French preference for more centralized governance that followed a Republican blueprint that despised hereditary and other forms of traditional rule (see, e.g. Müller-Crepon 2020).

Strategies of administrative unit designs among French governments are again less studied. The available evidence at least partially points towards a more fundamental disregard of pre-colonial ethno-political geographies. Pourtier (1989, p. 288) notes for the case of Gabon the French goal of establishing a tight administrative hierarchy modelled after Republican France that directly opposed tribalism and did not regard pre-existing institutions or ethnic geography. Bernier (1976) similarly concludes that French colonial *cercles* did not have roots in precolonial political or

ethno-linguistic topography, at least not at the end of the colonial period. This pattern has also been noted by Crowder (1968, p. 175) across French West Africa (see also Suret-Canale 1966; Guillaume 1999). In contrast, Lefebvre (2019) for the case of Niger notes that the colonial administration did indeed, presumably because of their weakness, aim at harnessing the social power of local elites and therefore modeled administrative units after their reach.

Yet, the evidence cited above is based on relatively few cases that may not generalize and does not account for potentially biasing influence of geographic features that simultaneously affect ethno-political geographies and colonial administrative borders. Even more importantly, given the impossibility of a perfect alignment between fuzzy precolonial ethnic and political topographies and sharp administrative borders, it remains unclear which yardsticks were applied to evaluate administrative unit designs in either colonial empire and to what extent "cracking" and "packing" strategies were used.

Observable implications

Clear expectations arise from the theoretical argument, its application to British and French (in)direct rule, and the suggestive historical evidence. In particular, I expect that fuzzy ethnic boundaries caused the drawing of aligned administrative borders in British colonies such that units are ethnically packed and groups remain comparatively uncracked. This effect of ethnic boundaries should be smaller if not absent in French colonies. It is, ex ante, unclear whether such an effect would be driven by the creation of non-packed, ethnically diverse units, the explicit cracking of ethnic groups, or a mix of both.

Data and empirical model

I test whether colonial governments aligned administrative geographies with prevailing ethno-linguistic geographies using newly digitized historical data on early 20^{th} century ethno-linguistic geography as well as spatial data on colonial administrative units. I estimate the effect of ethnic geography on administrative borders using a Probabilistic Spatial Partition Model (PSPM) recently developed by

Müller-Crepon, Schvitz and Cederman (2024), which allows for conditioning on observable covariates. The following presents the spatial data structure, the main variables of interest, and then introduces the empirical model and its specification.

Colonial state territories as planar graphs

The approach to modelling administrative borders follows Müller-Crepon, Schvitz and Cederman (2024) in understanding geographic space as a planar network of points. As discussed below, the PSPM takes such a graph as input and allows for estimating the effects of covariates encoded on its edges on the likelihood that two vertices belong to the same or different administrative units, also called partitions. As discussed by Müller-Crepon, Schvitz and Cederman (2024), the network approach discretizes the otherwise infinite number of possible outcomes, accounts for spatial interdependencies and observed covariates, and improves upon previous approaches to inferring the effects of spatial features such as ethnic geography on spatial partitionings.

The structure of the main graph used in the analysis takes into account both geographic area and the heavily skewed population distribution of the African continent, which comes with a strong correlation between population density and administrative units' size. This is captured by spatially sampling the graphs vertices from a population raster with a probability proportional to the natural logarithm of the raster cells' population estimated for the year 1880 by Goldewijk, Beusen and Janssen (2010).[8] The edges of the network are derived through a Delaunay triangulation of the points sampled. I sample 400 vertices per million square kilometers, which yields an average edge-length of approximately 50 kilometers. Figure 1 plots the full graph and Figure 4 shows more detail for the Gold Coast (Ghana). Appendix A.6 shows robustness of the results to varying spatial graph structures.

Administrative geographies

To measure the administrative borders and intersect them with the spatial network in Figure 1, I extend the spatial and temporal coverage of existing data on adminis-

[8] I add a constant of one.

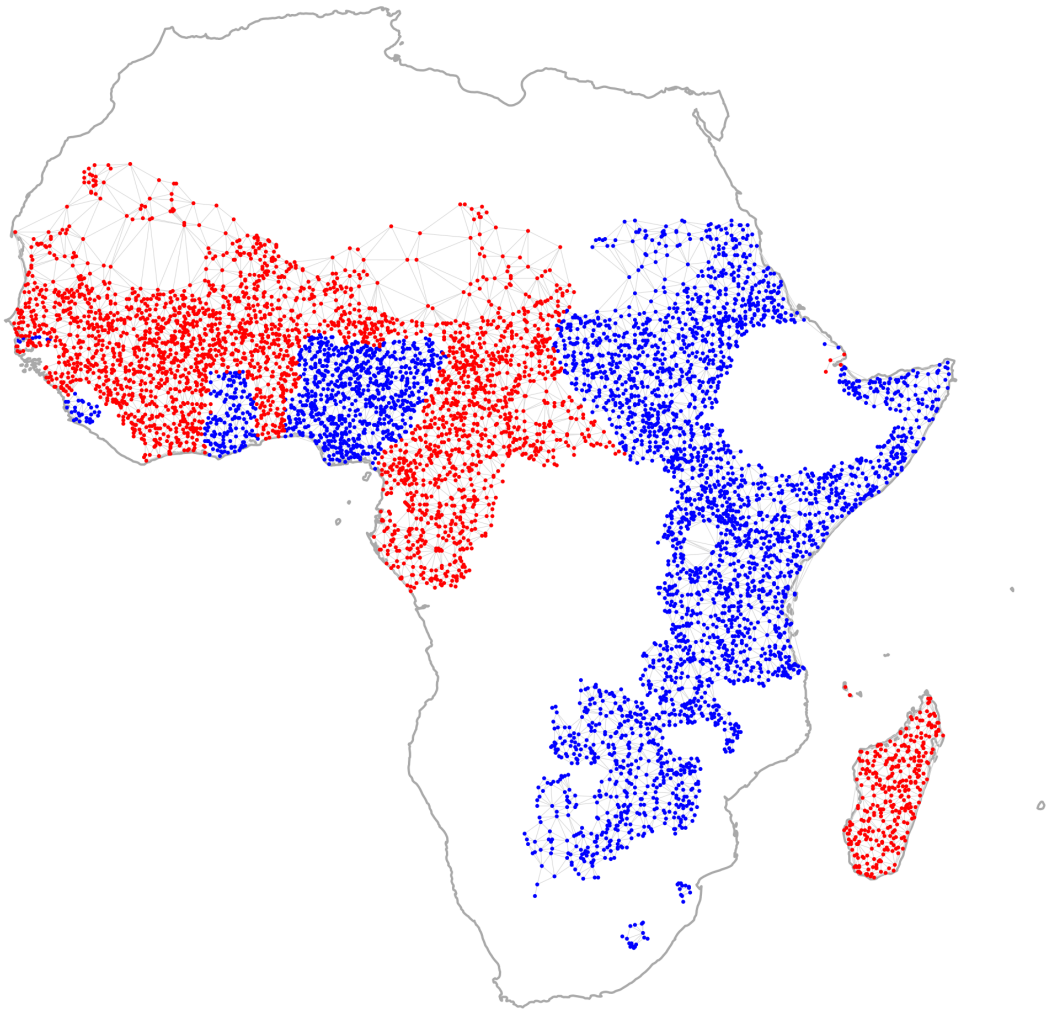


Figure 1: Former British and French colonies as planar graphs

trative units from British colonial sources (Müller-Crepon 2020) and in French West Africa (Huillery 2010). For the main analysis, I focus on British districts and French *cercles*.[9] As many authors highlight administrative borders have been changed with relative frequency (e.g. Bening 1999; Wilfahrt 2022). Similar to inter-colonial border change (Paine, Qiu and Ricart-Huguet 2024), intra-colonial borders were refined over time. The panel character of the newly collected administrative border data is particularly useful for estimating a lagged dependent variable model which addresses concerns over causal identification, particularly reverse causality.

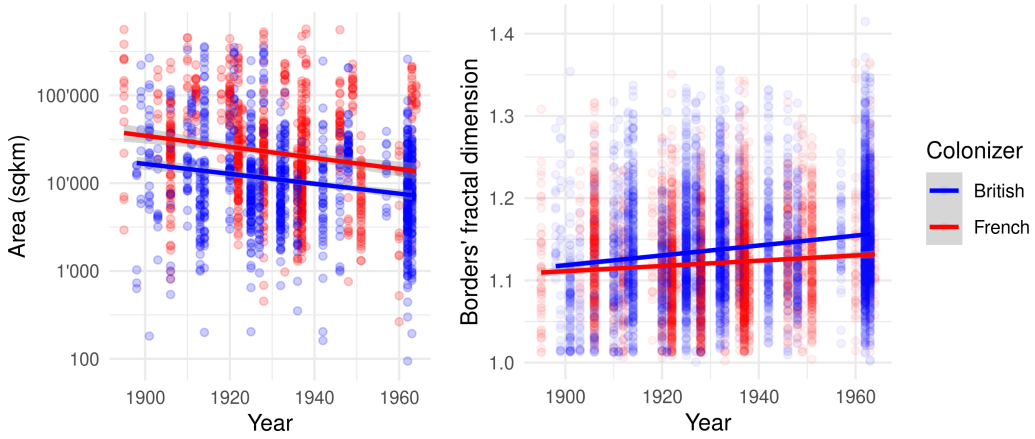
The data collection aimed at mapping British and French administrative geographies in the early (ca. 1900-1920), mid (ca. 1920-1940), and late (1940-independence) colonial period. In total, I have with the help of my research assistants digitized 53 new administrative maps. Joined with the existing data, the final database spans 82 unique colony-years in 26 colonies with a total of 1'858 administrative units.

The data clearly shows an iterative process of refining administrative units in the British and French colonies which is not unlike that of the borders of colonies themselves (Paine, Qiu and Ricart-Huguet 2024). As shown in Figure 2, units' average size roughly halves over time due to a parallel doubling of the number of administrative units. At the same time, the 'squiggliness' of units' borders increases steadily as they are drawn less often with a straight ruler but rather with better knowledge of local social and natural geography.[10] Because French colonies cover more desert areas that are barely populated, their administrative borders are, on average, straighter. The empirical analysis will test in how far border change over time aligned colonies administrative borders with ethnic geography.

Because the full data covers some colony-periods multiple times and others not at all, the analysis draws on a trimmed version of the dataset. It uses only one map per period, dropping colony-periods without coverage (see Figure 3). I define periods' start and end years flexible ($\pm$3 years) to maximize coverage. In additional analyses that explore the persistent effects of ethnic geography post-independence, I employ the time-series of post-colonial regional borders from Müller-Crepon

[9] French cercles are on average larger than British districts but smaller than British regions, for which no equivalent exists in French colonial Africa. British regions are the subject of an additional robustness check (see Appendix Table A2).

[10] Following Alesina, Easterly and Matuszeski (2011), I compute borders' fractal dimension which is 1 for perfectly straight lines and approaches 2 as a border is more squiggly and fills the plane.



(a) Size of administrative units over time

(b) Fractal dimension of administrative borders over time

Note: The fractal dimension of straight borders is 1. Its value increases towards two for very squiggly lines.

Figure 2: The refinement of colonial administrative partitionings over time

(2021).[11]

I intersect the data from every colony-period with the main graph introduced above. As the main outcome of the analysis shown in Figure 4b, I code the administrative unit within which every vertex fall. At the level of edges, I code whether an edge, in a given period, crosses an administrative boundary or not.

Historical ethnicity

While colonial administrative borders are comparatively well-documented on maps that were crucial for the functioning of the colonial state, data on ethnic geographies is scarcer and of worse quality. Even more importantly, ethnographic maps are affected by problems that directly relate to the concept of ethnicity itself.

First, as noted above, the boundaries of groups themselves were oftentimes fuzzy, with mutually unintelligible languages separated by dialect continua, further complicating the "ethnic grouping problem" which highlights the importance of the choice of granularity at which groups are conceptualized (Posner 2004). Second, groups often settled and continue to live in a spatially interspersed manner (Lentz 1995). Third, individuals' interethnic heritage and frequent multilingualism further complicates measurement (Buzasi 2016). A fourth problem consists in his-

[11] Unfortunately, post-colonial data does not exist at the level of districts before 1990.

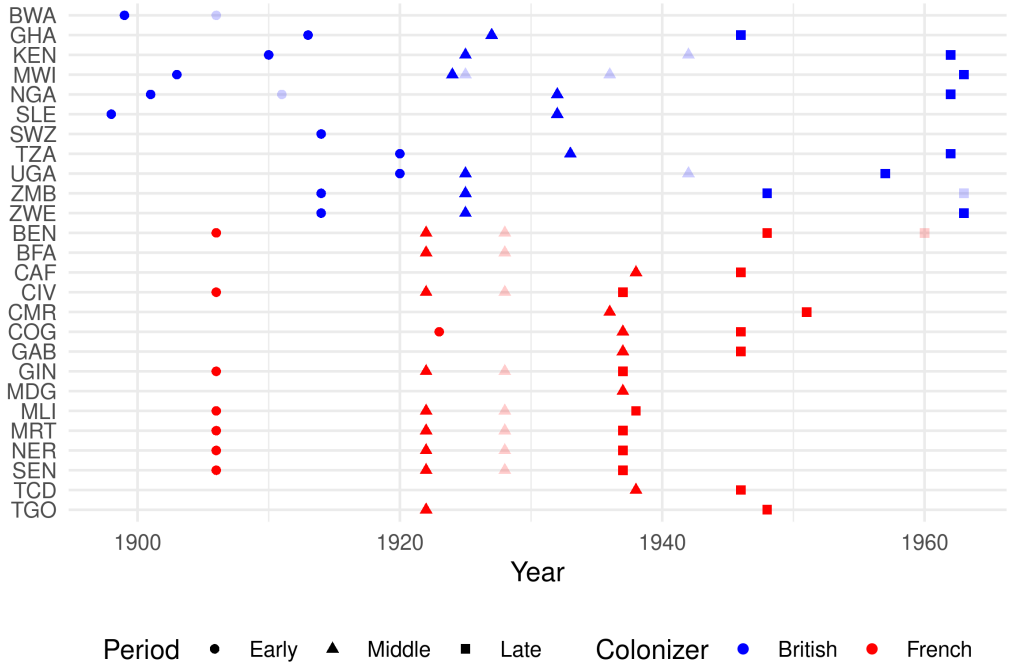
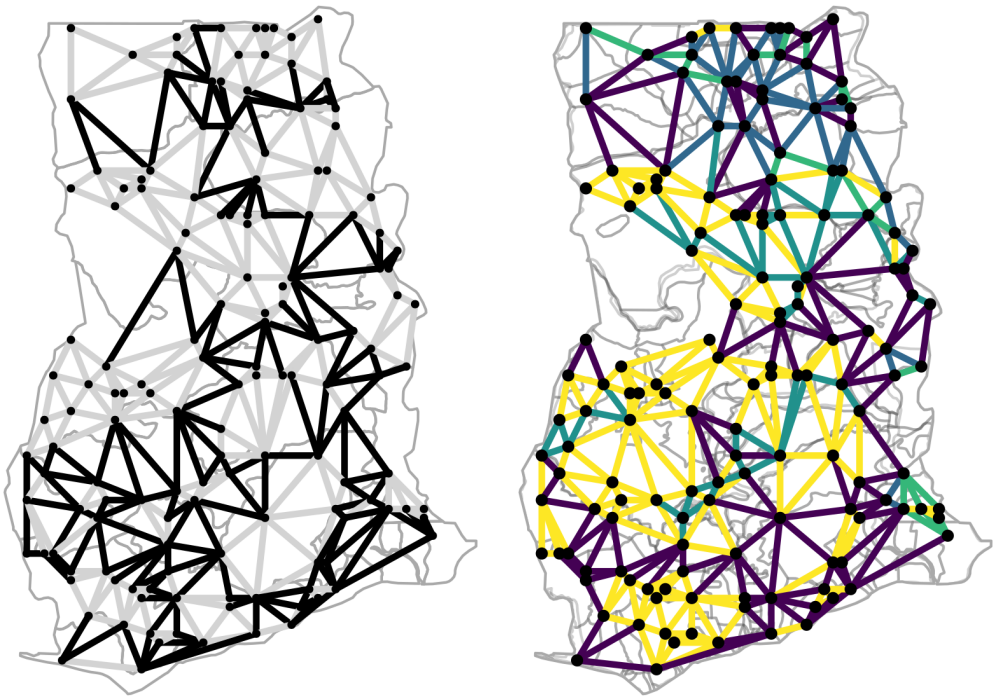


Figure 3: Administrative data used in the main analysis by colony and year. Transparent dots denote data that is dropped to avoid duplicate coverage of colony-periods.

torical reverse causality biases by which administrative borders shaped (perceptions of) ethnic groups and their geography (e.g. Müller-Crepon 2024; Singh and Vom Hau 2016).

While clearly important, the above concerns should not deflect from the fact that individuals across Sub-Sahara Africa spoke and still speak a vast diversity of languages, which differ to varying degrees and cluster geographically. In other words, categorical classifications of ethnic groups tend to be more informative in the center of groups' main settlement regions and less so in the linguistic and geographical space between them. The approach to measuring ethnic geography employed in this paper builds on this understanding while aiming to address the problems identified above. In particular, a collection of 49 newly digitized historical ethnographic maps drawn throughout the colonial period overcomes some of the problems associated with existing data on ethnic geography.

In the following, I first discuss problems relating to the quality and timing of the widely-used ethnographic map by George Murdock (1959). These severely limit



(a) Outcome: District borders

(b) Ethnic boundaries from colonial ethnographic maps

Note: Ethnic settlement polygons from six and points from one map in grey.

Figure 4: Data illustration: British Gold Coast (today's Ghana). Darker edges denote edges that cross district borders in (a) and greater ethnic differences in (b).

the map's usefulness for explaining colonial administrative borders. I then present a new collection of earlier ethnographic maps produced during the colonial period. They are of higher quality, together capture local ethnic mixing and uncertainty, and allow for analyzing administrative border changes occurring *after* their date of production thus avoiding an important aspect of reverse causality.[12]

Murdock's classic map of ethnic groups and its shortcomings: The ethnic map produced by Murdock (1959) is based on earlier secondary sources, was digitized by Nunn and Wantchekon (2011), and has been very influential in the burgeoning quantitative historical literature on the continent. The map shows the approximate settlement areas of 842 ethnic groups across the whole continent. Yet, while its coverage is extensive, it features three important drawbacks. First, the map is of low resolution (1:10 million or 1cm = 100km) which leads to low spatial detail.[13] Second, Murdock depicted groups' settlement areas as non-overlapping, neatly bounded, and shaped smoothly and regularly. This suggests either significant noise or, worse, bias in measurement of ethnic boundaries. Lastly, it is unclear how Murdock triaged between potentially diverging information on relevant ethnic groupings and their geography from secondary sources. The latter two caveats point to the substantive danger of reverse causality by which administrative regions might have influenced Murdock's mapping of ethnic geographies.

Murdock's map has the advantage of full and relatively uniform coverage of all of Africa, which makes it suitable for many research designs that do not rely on high spatial precision in the mapping of ethnic groups. Yet, for the present purpose of explaining the *location* of administrative borders, the map's low resolution, lack of detail, and risk of reverse causality motivate a new data collection which draws on ethnographic maps that predate Murdock's summary map.

[12] A previous version of this paper included an experimental measure of ethnic geography derived from today's placenames. While this measure correlates with colonial administrative boundaries in British (but not French) colonies, further analyses have shown that administrative boundaries have shaped the placenames we observe today along the lines of findings in Müller-Crepon (2024), raising the specter of reverse causality bias. Lacking a comprehensive list of (pre-)colonial toponyms, I am unable to address this bias and therefore dropped this analysis from the paper.

[13] The low resolution is compounded by the absence of coordinate lines or any other geographic features such as lakes or rivers on the map. This complicates georeferencing the map.

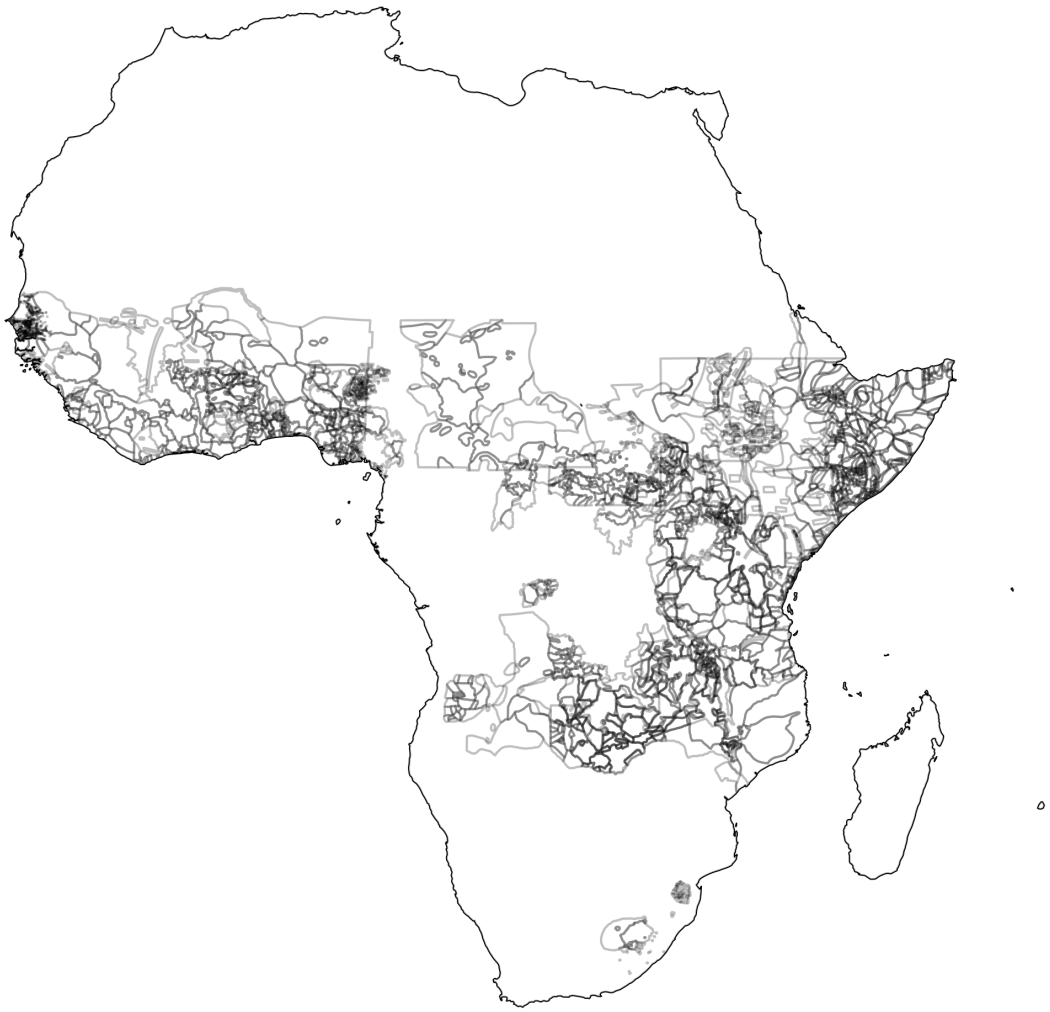


Figure 5: Ethnic settlement areas from 49 historical ethnic maps

Colonial ethnographic maps: The new collection of ethnographic maps from across Sub-Sahara Africa combines data from 49 newly digitized, historical maps. The resulting measure of ethnic boundaries captures the information on ethnic geographies and the uncertainties and ambiguities associated with it at the time during which administrative borders were drawn.

The maps were found through a systematic search in online map repositories and major library catalogues.[14] All maps were produced prior to that of Murdock (1959). Most maps follow the typical "polygon-style" also used by Murdock, though many depict ethnic settlement areas as overlapping. A few maps depict uncertainty directly by showing the names of some ethnic groups without any spatial delimiters, thus denoting fuzzy and unbounded ethnic settlement areas. As visible in Figure 4b, the maps have a much higher resolution than the one by Murdock, since they focus on regions, and more often even on (parts of) single colonies.

In sum then, the digitization of this ethnographic data allows for capturing local ethnic diversity and colonial ethnographers' uncertainty about ethnic groupings and their geographies. I encode this information on the graph by computing for each edge the fraction of maps on which an edge crosses an ethnic boundary (see Figure 4b). Where groups are depicted only as "limitless" labels, I associate vertices only with a group if it falls on the label.

Probabilistic Spatial Partition Model

I use a recently developed Probabilistic Spatial Partition Model (PSPM) to estimate the effect of ethnic geography encoded on the edges of the spatial graph on the partitioning of its vertices into administrative units. The model, presented in all detail by Müller-Crepon, Schvitz and Cederman (2024), captures the distribution of possible partitionings P of a graph as a Boltzman distribution

$$Pr(P = p_i) = \frac{e^{-\epsilon_i}}{\sum_{i=1}^{|\mathbb{P}|} e^{-\epsilon_i}}, \tag{1}$$

where the likelihood of a given partitioning p_i decreases with the "energy" ϵ_i

[14] Search terms include ethnic*, language, ethnographic, and similar. Libraries include the Bodleian Library at the University of Oxford, British Library, Library of Congress, and the Bibliothèque Nationale de France.

associated with it. This energy is in turn dependent on the realization of attractive and repulsive forces $\epsilon_{j,k}$ on the graph's edges between nodes j and k:

$$\epsilon_i = \sum_{j,k \in L} \mathbb{1}_{j,k} \epsilon_{j,k}, \tag{2}$$

$$\epsilon_{j,k} = \beta_0 + \beta_1 \text{ethnic boundary}_{j,k} + \gamma \mathbf{x}_{j,k}, \tag{3}$$

Most importantly and as denoted by the indicator $\mathbb{1}_{u_j = u_k}$ in Eq. 2, edges energy $\epsilon_{j,k}$ is only realized if its vertices j and k are part of the same administrative unit u. If realized, $\epsilon_{j,k}$ is determined by the weighted sum of a constant, the ethnic boundary measure, as well as observable covariates $\mathbf{x}_{j,k}$. The effects of these covariates consist in the β coefficients, which denote repulsive forces if positive and attractive forces if negative. These are estimated from the observed data in a conceptually similar way as typical regression coefficients. I therefore expect a positive coefficient for ethnic boundaries in British colonies which is larger than that obtained from the French sample. For ease of estimation and implementation, I estimate the models separately for the French and British samples throughout.

In the baseline specification, the covariate vector $\mathbf{x}_{j,k}$ captures potential joint determinants of administrative borders and ethnic geography. In particular, I include edges' length (logged), the size of the largest river and watershed they cross, their average elevation, as well as the average population density (Goldewijk, Beusen and Janssen 2010) and distance to the coast of the two vertices they connect (both logged). The baseline specification pools across all three periods.

A lagged dependent variable (LDV) specification addresses the risk of reverse causality and omitted variable bias. That specification adds to the vector of covariates $\mathbf{x}_{j,k}$ a lagged dependent variable that captures the presence of an administrative boundary in $t-1$. In addition, for the LDV model, I adjust the ethnic boundary measure such that it is only based on ethnic maps drawn in years prior to the observation of administrative borders in $t-1$.[15] Because of the scarcity of early ethnographic maps, the LDV model can only be estimated for the late colonial period,

[15] For Kenya and Uganda, the ethnic map (1943) was drawn the year *after* administrative borders are observed (1942). No administrative border changes are known to the author for that year and dropping the two cases from the analysis does not change the results, see Appendix Table A1.

with administrative borders mapped in the mid-colonial period as the lagged dependent variable. Ethnographic maps were drawn for West Africa in 1924, French Equatorial Africa in 1914,[16] and British East and Central Africa in 1943. While more limited in its empirical coverage this setup prevents direct reverse causality as well as unobserved confounders that affect the geography of ethnic boundaries as well as administrative borders.

The PSPM estimates the contribution of each variable to the overall potential energy of edges, captured as β coefficients such that the likelihood of the realized partitioning of colonies into administrative units is maximized. Estimation therefore follows a maximum composite likelihood approach with standard errors derived from a parametric bootstrap.[17] I split the graph into its British and French part to differentiate between the two colonial empires. Throughout, I assess all intra-empire borders of districts and cercles, including those that coincide with borders between colonies, since these were often changed and likely follow similar strategic concerns as administrative borders within colonies.

The varying effects of ethnic geography on administrative borders

The results from estimating the PSPM indicate that district borders where designed in relative congruence with ethnic geography in the British colonies. In turn, the borders of French *cercles* do not correlate systematically with ethnic boundaries. The difference between the two empires is consistent across the three colonial measurement periods and various permutations to the model and data structures.

Table 1 shows the main results using the baseline and LDV specification. Looking first at the British colonies, we see a consistent and precisely estimated effect of ethnic boundaries on the probability that two vertices are separated by district borders. For fully independent, so-called "bridge" edges which can change whether they cross a district border or not irrespective of all other edges in the network, the

[16] The results maybe biased by the inclusion of this very early and likely imprecise map. Dropping French Equatorial Africa does, however, not change the results of the LDV model, see Appendix Table A1.

[17] With 160 iterations and a burnin period of 10.

Table 1: Ethnic boundaries and administrative borders in British and French colonies

	British		French	
	(1)	(2)	(3)	(4)
Constant	-9.84^*	-9.12^*	-9.94^*	-6.48^*
	$[-10.58; -8.79]$	$[-10.49; -6.99]$	$[-10.71; -8.56]$	$[-8.05; -2.87]$
Ethnic boundary	0.47^*	0.38^*	0.14^*	0.02
	$[0.38; 0.57]$	$[0.25; 0.54]$	$[0.05; 0.24]$	$[-0.22; 0.25]$
Lagged dep. var.		0.81^*		1.06^*
		$[0.73; 0.97]$		$[0.89; 1.24]$
Controls	yes	yes	yes	yes
No. of vertices	5832	1662	5042	1010
No. of edges	15788	4493	13880	2584
No. of units	733	247	438	117

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

effects estimated associated with the boundaries drawn on colonial ethnographic maps amount to a hazard ratio of 1.60 [1.47, 1.77] in the baseline model and 1.46 [1.29, 1.71] in the LDV specification. Setting all covariates to their median values, ethnic boundaries in British colonies are associated with an increase of the chance of a bridge egde being crossed by a border from 24.6 [23.6, 26.5] to 34.3 [33.0, 36.7] percent in the baseline spcification. In the LDV model with a median lagged dependent variable of 0, the effect of ethnic boundaries amounts to an increase from 19.1 [16.7, 21.6] to 25.6 [22.7, 29.0]. Note that interdependence between non-bridge edges tends to vastly increase effect sizes as ethnic boundaries cross strings of connected edges. Most of the effect of ethnic boundaries in the LDV model is driven by the emergence of new administrative borders along ethnic boundaries and less so by greater stability of already existing ones (see Appendix Table A3).

Patterns in the set of French colonies look different with a much smaller and unstable effect associated with ethnic boundaries. While the ethnic boundaries show a association with administrative *cercle* boundaries in the baseline specification, this finding is not robust to the inclusion of the lagged dependent variable. In the former, the hazard ratio for bridge-edges associated with ethnic boundaries amount to 1.15 [1.05, 1.28] but decreases to a statistically and substantively insignificant 1.02 [0.80, 1.28] in the LDV specification.

Estimated effects of the control variables in models that exclude measures of ethnic geography conform with qualitative evidence on the influence of geographic

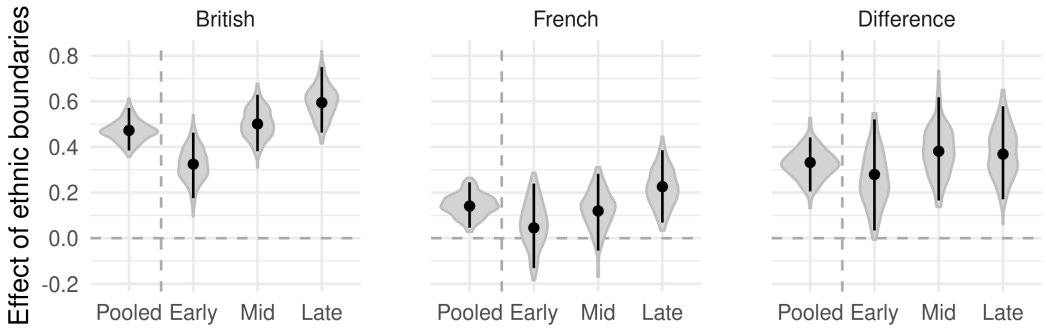


Figure 6: Baseline effect of ethnic boundaries on colonial administrative borders by measurement period

Note: Results from the baseline specification for the pooled and period-wise samples. 95% Confidence intervals and estimate distributions result from a parametric bootstrap with 160 iterations.

features on the design of administrative units in the colonial period (see Appendix A.5). In particular, rivers (but not watersheds) are frequent causes of district borders which are additionally more often drawn in densely populated areas. Average population density along edges is positively affected with the likelihood of them being crossed by administrative borders, indicating smaller units in more densely populated areas.

Figure 6 disaggregates the results from baseline Models 1 and 3 in Table 1 into the three measurement periods of colonial administrative borders.[18] The results show that the association between ethnic boundaries and administrative borders becomes stronger for the British and French colonies over time, yet at differing levels. While the association is substantive and statistically significant for British colonies throughout, it only becomes statistically significant in the French parts of Africa in the late colonial period. The small and insignificant results from the LDV model (4) in Table 1 suggests that this increase might be due to bias from reverse causality by which administrative borders have affected the drawing of later ethnic maps.

[18]The LDV results can not be disaggregated, since they are based on only on changes between the mid and late colonial period.

Robustness checks

The following discusses a series of robustness checks which are presented in more detail in Appendix A.

Linguistic distance: The risk of reverse causality is an important reason to carefully consider the interpretation of the main results. One channel through which administrative borders can reversely affect observed ethnic geographies is by affecting which ethno-linguistic groupings among all potential groupings become socially relevant and thus worthy of drawing on a map. Yet, such social construction is likely constrained by the structure of the ethnic "raw-material" – the emergence of ethnic groupings endogenous to administrative units such as the Luhya in Kenya (MacArthur 2013) is more likely among individuals speaking closely related languages than among unrelated linguistic spaces. It follows that reverse causality should bias effects of the boundaries between linguistically closely related groups more than the effects of linguistically distinct groupings. I therefore investigate whether effects of ethnic boundaries increase or decrease with linguistic differences, finding similarly sized effects on edges that cross small and large linguistic distances (Appendix Figure A1).[19] The consistent effect among edges that cross large linguistic distances suggests that the results are not exclusively due to the arbitrary invention of ethnic groups along or their disappearance within administrative borders.

Control variables: To avoid spurious results due to the potentially arbitrary choice of control variables, Appendix A.5 shows robustness to dropping all as well as extending the set of control variables. The latter variation adds additional measures that support previous arguments about the propensity of political units to be East-West oriented, and find that the French (but not British) aligned administrative borders with rugged terrain. The estimates for the effects of ethnic boundaries increase without any control variables and reach statistical significance in the French sample as well. This is likely due to the exclusion of spatial features – for

[19] Effects of large distances are smaller in the British LDV model, likely because such borders were drawn already by the mid-colonial period, leaving less variation to be explained when modeling modeling borders in the late colonial period.

example rivers or elevation – that affect administrative and ethnic geographies. In turn suggesting that the baseline specification includes most important covariates, estimated effects do not differ when more covariates are added to the baseline specification.

Colonial regions (British colonies): One potential caveat consists in the fact that the spatial organization of British and French colonies differed. While the British ruled through provinces and nested districts, the French *cercle* tend to be slightly larger than districts but smaller than provinces. Testing whether this difference in overall spatial organization drives the results by using British regions instead of districts as the outcome supports the baseline results as the estimated effects of ethnic boundaries are large (Table A2).[20]

Spatial structure: Additionally, Appendix A.6 report on a series of checks to investigate the sensitivity of the main results to changing (1) the spatial resolution of the planar graph, (2) its connectivity structure, and (3) the precise location of its edges. The results are robust to the use of spatially coarser and more disaggregated graphs but effect sizes generally increase with lower resolutions. This is not surprising, given that the encoding of ethnic boundaries on very short edges becomes more noisy and less likely to precisely coincide with the location of district borders (see also below). The results are robust to different connectivity structures, ranging from regular hexagonal, via quadratic graph structures, to graphs based on points sampled using normal population weights and completely random networks.[21] Lastly, the results are not due to the precise locations of nodes sampled for the baseline network.

Using standard regression models (with local fixed effects): Lastly, I employ a set of much simpler linear and logistic regression models which estimate the *edge-level* effect of ethnic boundaries on administrative borders. These models abstract

[20] Given the logistic structure of the model, estimated effects always relative to the size of units, captured by the baseline attraction on edges (β_0).

[21] Though note that standard errors increase for the LDV models when sampling points based on untransformed population distributions or randomly. While estimate sizes remain consistent, the effects of ethnic geography in the British colonies remain statistically significant but not their difference to the effects in the French colonial empire.

away from edges interdependence which is only taken into account for the clustering of standard errors[22]. Yet, they do allow for a much thorougher accounting of unobserved factors through fixed effects at the vertex level and are interpretable in a straightforward manner.

Shown in Appendix A.7, ethnic boundaries are associated with a positive, substantively large, and statistically significant effect of 22 (16) percentage points on the probability of British district borders in the baseline (LDV) model or 62 (40) percent of the mean outcome. In the French sample, this effect amounts to a statistically significant but much smaller 4 percentage points in the baseline model or 12 percent of the mean outcome. The LDV specification yields a non-significant effects of close to 0 percentage points. Adding vertex fixed effects does not change effects substantively in the linear model. These effects do not change when including fixed effects which account for significant amounts of local variation[23] and are consistent when estimating logistic regressions. There, including vertex fixed effects increases effect estimates. This is due multicollinearity which leads to the exclusion of all edges connected to vertices without any cross-border edges.

Post-colonial effect persistence and change

An additional analysis that employs the full panel of post-colonial *regional* borders (Müller-Crepon 2021) across former British and French colonies shows long-lasting persistence of the patterns of colonial administrative designs described above. Figure 7 shows the results of the baseline and LDV specifications[24] estimated for bi-decadal time periods as well as the pooled post-colonial data.

The results from the baseline analysis clearly show the differing levels of ethnic alignment former British and French colonies gain their independence with. With coefficients similar in size to those in the main analysis, administrative borders are significantly more in line with ethnic boundaries in countries gaining indepen-

[22] Using the cluster-robust variance estimator for dyadic data by Aronow, Samii and Assenova (2015).

[23] Including them raises the R^2 from below .18 (.10) to .61 (.55) in the British (French) baseline model.

[24] The post-colonial LDV specification uses all colonial ethno-graphic maps to measure ethnic boundaries. The lagged dependent variable consists in regional borders observed in the previous period as well as a deeper temporal lag of the mean edge-wise occurence of colonial district borders, based on the data from the main analysis.

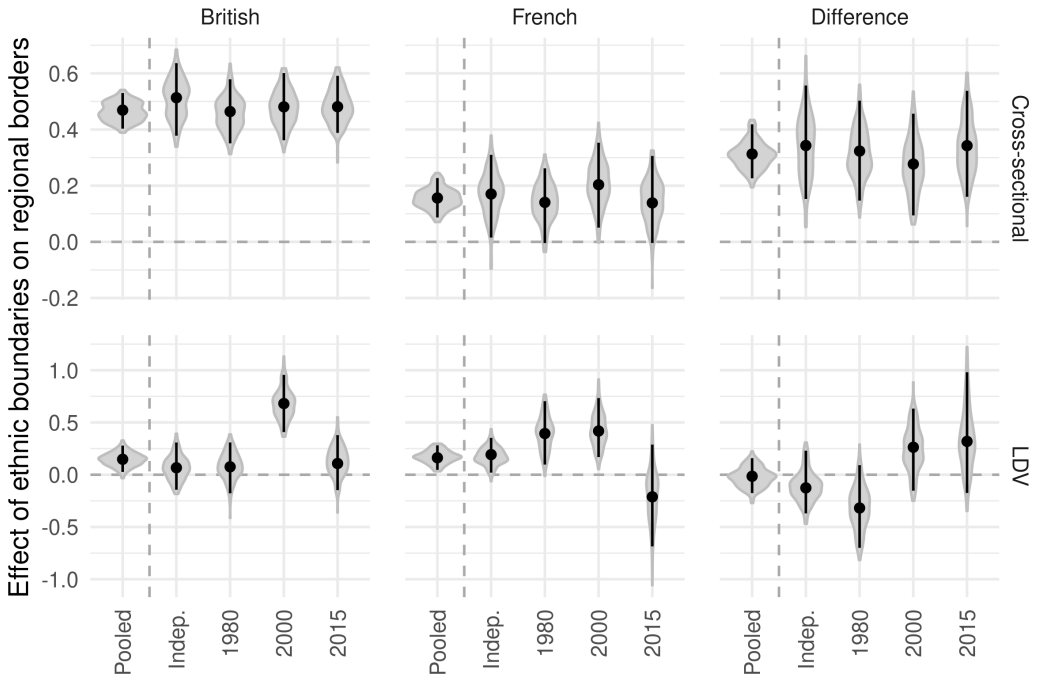


Figure 7: Effect of ethnic boundaries on the partitioning of post-colonial States into administrative regions

Note: Results from the baseline and LDV specification for pooled and bi-decadal samples. The LDV model includes an indicator for the presence of colonial borders in additional to the lagged dependent variable. 95% Confidence intervals and estimate distributions result from a parametric bootstrap with 160 iterations.

dence from the British as compared to the French empire. Over the years however, this difference in colonial legacies does not change much in size. This is because administrative border changes in former British *and* French colonies tend to, on average, follow ethnic boundaries. This is evidenced by the positive and statistically significant effect of ethnic boundaries in the LDV specification for the pooled samples which does not differ between countries with a French and British colonial past. Yet, when breaking down the effect into the different time periods, we observe that such ethnic alignments occur with significant temporal variation.

Packing or cracking ethnic groups?

The results so far show that, conditional on covariates, ethnic boundaries over-proportionally align with administrative borders in British but not French colonies. Yet, the extant focus on *edge-level* characteristics has prevented the analysis of the

ethnic packing of districts and administrative cracking of ethnic groups. Such an analysis requires the inclusion of predictors situated at higher levels of analysis. This section implements and estimates such predictors.

Beyond increasing the fit between theoretical argument and empirical analysis, macro-level predictors of ethnic packing and cracking address two additional potential caveats of the main analysis. First and as briefly mentioned above, the choice of spatial resolution is clearly important – a more detailed spatial setup comes with ever shorter edges on which the ethnic boundaries drawn on historical maps are ever less likely to systematically correlate with district borders. This is simply because of the fact that the precise location of ethnic boundaries drawn on the digitized maps is essentially random at the very local level. Second, even if – at the baseline level of spatial aggregation – ethnic boundaries do not correlate with district borders at all, the edge-level approach cannot distinguish between different extents of misalignment between ethnic and administrative geographies.

Packing and cracking as macro-level predictors

To implement macro-level predictors of ethnic packing and cracking above the micro edge-level, I return to the parametrization of the total "energy" that is realized by a given partitioning as formulated in Eq. 2. As noted by Müller-Crepon, Schvitz and Cederman (2024), its focus on energies associated with edges' characteristics can be extended by functions that capture energies produced by interactions between larger sets of vertices or edges. In the present case, this is the extent to which a partitioning p_i packs districts $u \in U_i$ so that they are ethnically homogeneous and cracks groups $g \in G$ so that they are fragmented by districts' borders:

$$\epsilon_i = \gamma_1 \,\mathsf{packing}(U_i, G) + \gamma_2 \,\mathsf{cracking}(U_i, G) + \sum_{j,k \in L} \mathbb{1}_{j=k} \epsilon_{j,k} \tag{4}$$

where partitionings p_i degree of packing and cracking is computed as Herfindahl-Hirschman indices aggregated across districts $u \in U$ and groups $g \in G$ weighted according to their absolute sizes (w_u and w_g).[25] Starting with districts,

[25] Using absolute sizes as weights mirrors the fact that the overall impact of edge-level predictors on partitionings' realization probability is also defined as the *sum* over all edges.

packing is defined as

$$\text{packing}(U_i, G) = \sum_{u \in U_i} w_u \sum_{g \in G} (\frac{u_e}{u})^2, \tag{5}$$

$\text{packing}(U_i, G)$ is 1 if every district u is fully ethnically homogenous it contains and decreases towards 0 as districts become more ethnically diverse. Accordingly, a *negative* γ_1 in Eq. 4 would signal that units' borders are drawn such that districts are "packed", since ethnically homogeneous districts then decrease the overall energy of a partitioning and increase its chance of realization.

Moving the the level of ethnic groups, cracking is similarly defined as

$$\text{cracking}(U_i, G) = \sum_{g \in G} w_g \, (1 - \sum_{u \in U_i} (\frac{g_u}{g})^2) \tag{6}$$

$\text{cracking}(U_i, G)$ is 0 if every group g is fully contained within one district (irrespective of its ethnic diversity) and moves towards 1 as a group is split into ever more parts by administrative borders. A *negative* γ_2 in Eq. 4 then suggests that borders are drawn to crack groups, whereas a positive one would suggest that borders leave them explicitly uncracked. Note that the implementation of the packing and cracking parameters has potential applications beyond this study. In particular, similar logics are often articulated in the gerrymandering literature as driving electoral district designs (e.g. Katz, King and Rosenblatt 2020) but not explicitly tested with account of potential covariates.

Such supra-edge predictors are very powerful in the extent to which they can be precisely tailored to capture theoretical arguments. Yet, the precision with which the model can be adjusted comes at the disadvantage of more difficult decisions regarding the appropriate modelling of potential omitted variables, all of which could also be used to construct an essentially limitless set of supra-edge predictors. Avoiding potentially arbitrary choices among essentially limitless possibilities, I here simply control for the edge-level covariates used in the baseline analysis which have most likely a local effect. The lagged-dependent variable for the LDV specification accounts for the effects of further omitted variables, including time-invariant effects of stable supra-edge predictors.

Since the packing(U_i, G) and cracking(U_i, G) predictors are based on a categorical measurement of ethnic geography,[26] I draw on the earliest available ethnographic maps for Central (1914), West (1924), and East (1943) Africa to estimate the γ parameters in Eq. 4.

Results on packing and cracking

The results in Table 2 provide insights that add substantive detail to the estimates of the overall effect of ethnic boundaries in Table 1.

Table 2: Packing or Cracking Ethnic Groups?

	British		French	
	(1)	(2)	(3)	(4)
Constant	-9.72^*	-8.94^*	-10.55^*	-6.54^*
	$[-10.80; -8.77]$	$[-9.77; -6.07]$	$[-11.40; -9.07]$	$[-7.57; -2.81]$
Packing	-0.44^*	-0.27^*	-0.31^*	-0.27
	$[-0.58; -0.37]$	$[-0.49; -0.12]$	$[-0.47; -0.20]$	$[-0.64; 0.04]$
Cracking	0.32^*	0.44^*	0.04	-0.15
	$[0.14; 0.40]$	$[0.17; 0.61]$	$[-0.17; 0.17]$	$[-0.51; 0.38]$
Lagged dep. var.		0.81^*		1.06^*
		$[0.69; 0.94]$		$[0.95; 1.30]$
Controls	yes	yes	yes	yes
No. of vertices	5209	1662	4030	1010
No. of edges	14042	4493	10566	2584
No. of units	681	247	428	117

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

First, the results for the extent of the packing of ethnic groups shows some, but relatively small and statistically insignificant differences between the British and French approaches to administrative unit designs. In both empires, the borders of districts and cercles were drawn such that the ethnic homogeneity of administrative units was increased as compared to district borders that are only affected by the edge-level covariates. The effect of the ethnic packing of district decreases slightly but remains statistically significant when adding the lagged dependent variable in the British sample. It is slightly smaller and looses significance in the French sample but retains substantive size. In substantive terms, the estimate in the British Model (1) suggests that increasing the ethnic homogeneity of districts that cover a

[26] I abstain from implementing a version that averages across maps which would add significant complexity to estimation and interpretation. Yet, in principle, the parameters could also be based on compositional data as retrieved, for example, from census data where available.

set of ten nodes by .25 increases a partitioning's chance of realization by a factor of three.[27]

Second, larger differences between the British and French empires appear to consist in the degree of cracking of ethnic settlement areas they engaged in. For the British sample, we observe a positive estimate for the impact of cracking, suggesting that administrative units left ethnic groups more undivided than one would expect only on the basis of the edge-level covariates. More specifically, increasing the average level of cracking of ethnic groups by .25 in the above toy example of 10 vertices would decrease the chance of realization of a given partitioning by a factor of .45.[28] This effect slightly increases when adding the lagged dependent variable in Model (2). In contrast, the estimated effect of cracking is small in the French baseline specification and even negative(but statistically insignificant) in the French LDV model.

In sum, these results suggest that the alignment of administrative borders with ethnic boundaries in British colonies is driven by the packing of ethnic groups which are left comparatively uncracked. In turn, French colonies feature packed cercles which do, however, crack groups into multiple units. This is consistent with previous findings that the French were substantively more hostile towards pre-colonial institutions and their rulers than the British colonial governments (Müller-Crepon 2020).

Discussion and conclusion

The partitioning of states' territories into administrative units is a crucial feature of territorial states with important outcomes on, among others, ethnic politics, economic development, and political stability. Going beyond previous literature that focuses on unit change, this paper has investigated the strategies underlying initial designs of administrative geographies. Administrative units can be drawn either in alignment with pre-existing social and political geographies, packing important social groups and their institutions into their own, homogeneous divisions without

[27] This hazard ratio is computed as $e^{-1*-.44*.25*10}$, see Eq. 1.
[28] This hazard ratio is computed as $e^{-1*.32*.25*10}$, see Eq. 1.

dividing them internally. Or, administrative borders can be aimed at transforming societies, cracking pre-existing groups and institutions into diverse units. While the first strategy is relatively easy to implement, it leaves political power decentralized. The second approach, in turn, is more costly in the short run but centralizes power at the expense of local institutions and elites.

I have tested this argument by analyzing administrative unit designs in British and French colonies in Sub-Sahara Africa. While the British had a comparatively decentralized style of indirect rule, the French, for ideological and material reasons, used more direct rule to centralize power. My results show that the design of administrative units differed accordingly. Using newly collected data on administrative borders and ethnic geography and a recently developed probabilistic spatial partition model, the analyses show that British colonial governments tended to design administrative units along ethnic boundaries, packing groups into comparatively homogeneous districts that left them uncracked. In turn, administrative borders in French colonies do not systematically correlate with ethnic boundaries, with the main difference to British designs consisting in comparatively higher levels of administrative cracking of ethnic groups.

Taken together with recent results of the effects of administrative units on ethnic groups' geography (Müller-Crepon 2024),[29] my current findings point towards the possibility of a dynamic co-development of administrative units and ethnic groups. The mindset of particularly British colonial rulers distributed power roughly along initially fuzzy and geographically ill-defined ethnic lines. This led to a crystallization of identities along borders drawn, as well as demands to change unit designs along sharpened ethnic lines (Grossman and Lewis 2014), which in turn likely further strengthened ethnic identities. This dynamic alignment of administrative geographies and ethnic identities is fundamentally driven by the introduction of the concept of territorial governance through neatly bounded and non-overlapping administrative divisions.

[29] Note that the present results does not invalidate the research design or findings in Müller-Crepon (2024), which rely on local variation at administrative borders in a regression discontinuity design, including at straight borders drawn in an as-if-random manner *at the local level*.

References

Alesina, Alberto and Ekaterina Zhuravskaya. 2011. "Segregation and the quality of government in a cross-section of countries." *American Economic Review* 101(5):1872–1911.

Alesina, Alberto and Enrico Spolaore. 1997. "On the number and size of nations." *The Quarterly Journal of Economics* 112(4):1027–1056.

Alesina, Alberto and Enrico Spolaore. 2005. *The size of nations*. Cambridge, MA: MIT Press.

Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public goods and ethnic divisions." *Quarterly Journal of Economics* 114(4):1243–1284.

Alesina, Alberto, William Easterly and Janina Matuszeski. 2011. "Artificial states." *Journal of the European Economic Association* 9(2):246–277.

Ali, Merima, Odd-Helge Fjeldstad, Boqian Jiang and Abdulaziz B. Shifa. 2019. "Colonial Legacy, State-building and the Salience of Ethnicity in Sub-saharan Africa." *Economic Journal* 129(619):1048–1081.

Aronow, Peter M, Cyrus Samii and Valentina A Assenova. 2015. "Cluster–robust variance estimation for dyadic data." *Political analysis* 23(4):564–577.

Asiwaju, Anthony I. 1970. The Aleketu of Ketu and the Onimek of Meko: The Changing Status of two Yoruba Rulers under French and British rule. In *West African Chiefs. Their Changing Status under Colonial Rule and Independence*, ed. Michael Crowder and Obaro Ikime. Ile-Ife: University of Ife Press pp. 134–161.

Asiwaju, Anthony I. 1983. "The Concept of Frontier in the Setting of States in Precolonial Africa." *Présence Africaine* 127:43.

Bening, R Bagulo. 1999. *Ghana: regional boundaries and national integration*. Ghana University Press.

Berman, Bruce J. 1998. "Ethnicitiy, Patronage, and the African State: The Politics of Uncivil Nationalism." *African Affairs* 388(97):305–341.

Bernier, Jacques. 1976. "La formation territoriale du Sénégal." *Cahiers de géographie du Québec* 20(51):447–477.

Buzasi, Katalin. 2016. "Linguistic situation in twenty sub-Saharan African countries: A survey-based approach." *African Studies* 75(3):358–380.

Carter, David B and Hein E Goemans. 2011. "The making of the territorial order: New borders and the emergence of interstate conflict." *International Organization* 65(2):275–309.

Cohen, Ronald and John Middleton. 1970. *From Tribe to Nation in Africa: studies in incorporation processes*. Chandler Pub. Co.

Cohen, William B. 1971. *Rulers of Empire: the French Colonial Service in Africa*. Stanford: Hoover Institution Press.

Colson, Elizabeth. 1960. African Society at the Time of the Scramble.

Conklin, Alice L. 1997. *A Mission to Civilize. The Republican Idea of Empire in France and West Africa, 1895–1930.* Stanford: Stanford University Press.

Crowder, M. 1964. "Indirect rule: French and British style." *Africa: Journal of the International African Institute* 34(3):197–205.

Crowder, Michael. 1968. *West Africa under Colonial Rule*. London: Hutchinson.

Cunningham, Kathleen Gallagher and Nils B. Weidmann. 2010. "Shared Space: Ethnic Groups, State Accommodation, and Localized Conflict." *International Studies Quarterly* 54(4):1035–1054.

Englebert, Pierre, Stacy Tarango and Matthew Carter. 2002. "Dismemberment and Suffocation. A Contribution to the Debate on African Boundaries." *Comparative Political Studies* 35(10):1093–1118.

Fortes, Meyer and Edward E. Evans-Pritchard, eds. 1940. *African Political Systems*. 1958 ed. Oxford: Oxford University Press.

Friedman, David. 1977. "A Theory of the Size and Shape of Nations." *Journal of political economy* 85(1):59–77.

Gerring, John, Daniel Ziblatt, Johan van Gorp and Julián Arévalo. 2011. "An Institutional Theory of Direct and Indirect Rule." *World Politics* 63(3):377–433.

Goldewijk, Kees Klein, Arthur Beusen and Peter Janssen. 2010. "Long-term dynamic modeling of global population and built-up area in a spatially explicit way: HYDE 3.1." *The Holocene* 2010(1):1–9.

Green, Elliott D. 2010. "Patronage, district creation, and reform in Uganda." *Studies in Comparative International Development* 45(1):83–103.

Griffiths, Ryan D. 2016. *The Age of Secession: The International and Domestic Determinants of State Birth*. Cambridge: Cambridge University Press.

Grossman, Guy, Jan H. Pierskalla and Emma Boswell Dean. 2017. "Government Fragmentation and Public Goods Provision." *Journal Of Politics* 79(3):823–839.

Grossman, Guy and Janet I Lewis. 2014. "Administrative unit proliferation." *American Political Science Review* 108(1):196–217.

Guillaume, P. 1999. *Le Monde colonial: XIXe-XXe siècle*. Coll. U Armand Colin.
URL: *https://books.google.fr/books?id=QoAdHAAACAAJ*

Hassan, Mai. 2016. "A State of Change: District Creation in Kenya after the Beginning of Multi-party Elections." *Political Research Quarterly* 69(3):510–521.

Hechter, Michael. 1975. *Internal Colonialism. The Celtic Fringe in British National Development, 1536-1966*. London: Routledge & Kegan Paul.

Henn, Soeren J. 2023. "Complements or substitutes? How institutional arrangements bind traditional authorities and the state in Africa." *American Political Science Review* 117(3):871–890.

Herbst, Jeffrey. 2000. *States and Power in Africa*. Princeton: Princeton University Press.

Hirsch, Francine. 2000. "Toward an empire of nations: border-making and the formation of Soviet national identities." *The Russian Review* 59(2):201–226.

Howard, Allen M. 2005. Nodes, Networks, Landscapes, and Regions: Reading the Social History of Tropical Africa 1700s–1920. In *The Spatial Factor in African History*. Brill pp. 21–140.

Huillery, Elise. 2010. "The impact of European settlement within French West Africa: Did precolonial prosperous areas fall behind?" *Journal of African Economies* 20(2):263–311.

Juon, Andreas. 2024. "Territorial Autonomy and the Trade-Off between Civil and Communal Violence." *American Political Science Review* pp. 1–17.

Katz, Jonathan N, Gary King and Elizabeth Rosenblatt. 2020. "Theoretical foundations and empirical evaluations of partisan fairness in district-based democracies." *American Political Science Review* 114(1):164–178.

Kirk-Greene, Anthony Hamilton Millard. 1980. "The Thin White Line: The Size of the British Colonial Service in Africa." *African Affairs* 79(314):25–44.

Kristof, Ladis K D. 1959. "The Nature of Frontiers and Boundaries." *Annals of the Association of American Geographers* 49(3):269–282.

Lefebvre, C. 2019. *Frontières de sable, frontières de papier: Histoire de territoires et de frontières, du jihad de Sokoto à la colonisation française du Niger, XIXe-XXe siècles*. Bibliothèque historique des pays d'Islam Éditions de la Sorbonne.
URL: *https://books.google.fr/books?id=J8C9DwAAQBAJ*

Lentz, Carola. 1995. "'Tribalism'and ethnicity in Africa." *Cah. Sci. hum* 31(2):303–328.

Lentz, Carola. 2006. *Ethnicity and the Making of History in Northern Ghana*. Edinburgh: Edinburgh University Press.

MacArthur, Julie. 2013. "When did the Luyia (or any other group) become a tribe?" *Canadian Journal of African Studies/La Revue canadienne des études africaines* 47(3):351–363.

Mann, Michael. 1984. "The autonomous power of the state: its origins, mechanisms and results." *European Journal of Sociology* 25(2):185–213.

Miguel, Edward and Mary Kay Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of public Economics* 89(11-12):2325–2368.

Miles, William F. S. 1994. *Hausaland Divided. Colonialism and Independence in Nigeria and Niger*. Ithaca: Cornell University Press.

Müller-Crepon, Carl. 2020. "Continuity or Change: (In)direct Rule in British and French Colonial Africa." *International Organization* 74(4):707–741.

Müller-Crepon, Carl. 2021. "State reach and development in Africa since the 1960s: new data and analysis." *Political Science Research and Methods* pp. 1–10.

Müller-Crepon, Carl. 2024. "Building Tribes: How Administrative Units Shaped Ethnic Groups in Africa." *American Journal of Political Science, Early View* .

Müller-Crepon, Carl, Guy Schvitz and Lars-Erik Cederman. 2024. "Shaping States into Nations: The Effects of Ethnic Geography on State Borders." *American Journal of Political Science, Early View* .

Murdock, George Peter. 1959. *Africa. Its Peoples and Their Culture History.* New York: McGraw-Hill Book Company.

Nunn, Nathan and Leonard Wantchekon. 2011. "The slave trade and the origins of Mistrust in Africa." *American Economic Review* 101(7):3221–3252.

Oates, Wallace E. 1999. "An Essay on Fiscal Federalism." *Journal of Economic Literature* 37(3):1120–1149.

Paine, Jack. 2019. "Ethnic Violence in Africa: Destructive Legacies of Pre-Colonial States." *International Organization* 73(3):645–683.

Paine, Jack, Xiaoyan Qiu and Joan Ricart-Huguet. 2024. "Endogenous Colonial Borders: Precolonial States and Geography in the Partition of Africa." *American Political Science Review, FirstView* .

Posner, Daniel N. 2004. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi." *The American Political Science Review* 98(4):529–545.

Pourtier, Roland. 1989. "Les États et le contrôle territorial en Afrique centrale : principes et pratiques." *Annales de géographie* 98(547):286–301.
URL: *https://www.persee.fr/doc/geo*$_0$003 − 4010$_1$989$_n$*um*$_9$8$_5$47$_2$0907

Resnick, Danielle. 2017. "Democracy, decentralization, and district proliferation: The case of Ghana." *Political Geography* 59:47–60.

Robinson, Amanda Lea. 2020. "Ethnic diversity, segregation and ethnocentric trust in Africa." *British Journal of Political Science* 50(1):217–239.

Roeder, Philip G. 2012. *Where nation-states come from*. Princeton University Press.

Ruggie, John Gerard. 1983. "Continuity and transformation in the world polity: Toward a neorealist synthesis." *World Politics* 35(2):261–285.

Sharpe, Barrie. 1986. "Ethnography and a regional system: Mental maps and the myth of states and tribes in north-central Nigeria." *Critique of Anthropology* 6(3):33–65.

Singh, Prerna and Matthias Vom Hau. 2016. "Ethnicity in time: Politics, history, and the relationship between ethnic diversity and public goods provision." *Comparative Political Studies* 49(10):1303–1340.

Soifer, Hillel David. 2019. Units of analysis in subnational research. In *Inside countries: Subnational research in comparative politics*, ed. Agustina Giraudy, Eduardo Moncada and Richard Snyder. Cambridge University Press pp. 92–112.

Southall, Aidan W. 1970. "The illusion of tribe." *Journal of Asian and African Studies* 5(1-2):28–50.

Spear, Thomas. 2003. "Neo-Traditionalism and the Limits of Invention in British Colonial Africa." *Journal of African History* 44(1):3–27.

Suret-Canale, Jean. 1966. "La Fin de la Chefferie En Guinee." *The Journal of African History* 7(3):459–493.

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990-1992*. Cambridge, MA: Blackwell.

Tommasi, Mariano and Federico Weinschelbaum. 2007. "Centralization vs. decentralization: A principal-agent analysis." *Journal of public economic theory* 9(2):369–389.

Weber, Max. 1919. *Politik als Beruf*. Berlin: Duncker & Humblot.

Wilfahrt, Martha. 2022. *Precolonial Legacies in Postcolonial Politics: Representation and Redistribution in Decentralized West Africa*. Cambridge University Press.

Wilks, Ivor. 1975. *Asante in the ninetheenth century. The structure and evolution of political order.* Cambridge: Cambridge University Press.

Young, Crawford. 1985. Ethnicity and the Colonial and Postcolonial State in Africa. In *Ethnic Groups and the State*, ed. Paul R. Brass. London: Croom Helm pp. 57–93.

Ziblatt, Daniel. 2004. "Rethinking The Origins of Federalism: Puzzle, Theory, and Evidence from Nineteenth-Century Europe." *World Politics* 57(01):70–98.

Ziblatt, Daniel. 2006. *Structuring the State. The Formation of Italy and Germany and the Puzzle of Federalism*. Princeton: Princeton University Press.

Supplementary Material

Table of Contents

A Robustness checks

A.1 Sample adjustments

Table A1: Varying the sample definition

	Baseline for LDV Sample		LDV: Dropping colonies	
	1: Brit.	2: French	3: Brit.	4: French
Constant	-9.66^*	-10.06^*	-8.51^*	-6.33^*
	$[-10.42; -8.41]$	$[-11.16; -8.10]$	$[-9.54; -6.13]$	$[-7.82; -2.49]$
Ethnic boundary	0.48^*	0.23^*	0.37^*	0.03
	$[0.41; 0.57]$	$[0.10; 0.34]$	$[0.24; 0.56]$	$[-0.21; 0.27]$
Lagged dep. var.			0.75^*	1.09^*
			$[0.57; 0.88]$	$[0.95; 1.31]$
Dropped	No LDV	No LDV	KEN & UGA	AEF
Controls	yes	yes	yes	yes
No. of vertices	4988	2984	1399	970
No. of edges	13489	7679	3834	2490
No. of units	620	344	193	113

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

A.2 Regional borders in British colonies

Table A2: Ethnic boundaries and regional borders in British colonies

	Baseline	LDV
	(1)	(2)
Constant	-9.63^*	-9.21^*
	$[-10.71; -7.53]$	$[-9.85; -6.01]$
Ethnic boundary	0.60^*	0.56^*
	$[0.47; 0.71]$	$[0.35; 0.79]$
Lagged dep. var.		0.73^*
		$[0.52; 1.00]$
Controls	yes	yes
No. of vertices	2318	880
No. of edges	6183	2354
No. of units	254	108

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

A.3 Persistence or change

Table A3: Ethnic boundaries and subnational borders: Persistence vs change

	British			French		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-9.63^*	-9.06^*	-9.47^*	-6.41^*	-6.48^*	-6.42^*
	$[-10.43; -6.56]$	$[-10.33; -6.58]$	$[-10.50; -6.46]$	$[-7.92; -3.18]$	$[-8.13; -3.33]$	$[-7.83; -3.15]$
Lagged dep. var. (LDV)	2.89	1.00^*	2.57	-0.35	1.03^*	-0.33
	$[-1.49; 3.91]$	$[0.84; 1.21]$	$[-2.27; 3.72]$	$[-3.89; 2.20]$	$[0.79; 1.29]$	$[-4.60; 2.05]$
Ethnic boundary	0.38^*	0.56^*	0.54^*	0.01	-0.03	-0.00
	$[0.27; 0.53]$	$[0.39; 0.79]$	$[0.38; 0.77]$	$[-0.25; 0.20]$	$[-0.36; 0.31]$	$[-0.32; 0.28]$
Ethnic boundary$\times$ LDV		-0.42^*	-0.36^*		0.10	0.02
		$[-0.75; -0.15]$	$[-0.69; -0.11]$		$[-0.33; 0.62]$	$[-0.41; 0.49]$
Controls	yes	yes	yes	yes	yes	yes
Controls $imes$ LDV	yes	no	yes	yes	no	yes
No. of vertices	1662	1662	1662	1010	1010	1010
No. of edges	4493	4493	4493	2584	2584	2584
No. of units	247	247	247	117	117	117

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

A.4 Linguistic distance

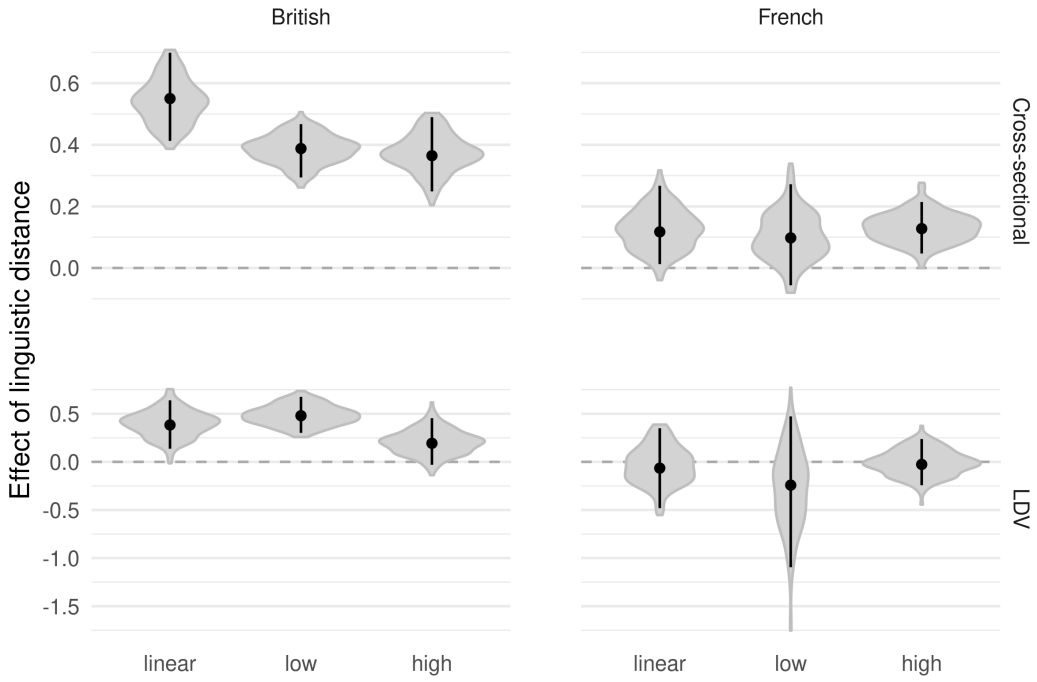


Figure A1: Effect of linguistic distances modeled linearly and binned (low/high)

Note: 95% CIs and estimate distributions result from a parametric bootstrap with 160 iterations.

A.5 Control variables

Table A4: British colonies: Sensitivity towards vector of controls

	Baseline			LDV		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-1.27^*	-10.92^*	-11.99^*	-1.49^*	-9.73^*	-9.01^*
	$[-1.29; -1.18]$	$[-11.73; -9.93]$	$[-12.70; -9.15]$	$[-1.58; -1.38]$	$[-10.79; -7.84]$	$[-10.09; -5.31]$
Lagged dep. var.				0.92^*	0.81^*	0.81^*
				$[0.81; 1.05]$	$[0.69; 0.91]$	$[0.70; 0.97]$
Ethnic boundary	0.75^*		0.47^*	0.53^*		0.37^*
	$[0.65; 0.81]$		$[0.38; 0.57]$	$[0.41; 0.69]$		$[0.24; 0.53]$
Edge length		0.87^*	0.98^*		0.75^*	0.71^*
		$[0.79; 0.94]$	$[0.72; 1.02]$		$[0.58; 0.82]$	$[0.33; 0.81]$
River		0.98^*	0.81^*		1.03^*	0.97^*
		$[0.79; 1.15]$	$[0.65; 1.02]$		$[0.75; 1.36]$	$[0.54; 1.50]$
Watershed		0.02	0.10		-0.00	-0.03
		$[-0.12; 0.15]$	$[-0.03; 0.22]$		$[-0.20; 0.27]$	$[-0.34; 0.17]$
Elevation mean		0.03	-0.61^*		0.60	0.43
		$[-0.28; 0.35]$	$[-1.10; -0.20]$		$[-0.06; 1.26]$	$[-0.63; 1.31]$
Population 1880		0.12^*	0.11^*		0.09^*	0.08^*
		$[0.09; 0.13]$	$[0.07; 0.13]$		$[0.03; 0.13]$	$[0.02; 0.13]$
Dist. coast		-0.01	0.03		-0.03	-0.08
		$[-0.05; 0.03]$	$[-0.02; 0.08]$		$[-0.12; 0.04]$	$[-0.22; 0.02]$
Δ Long.			-3.31^*			-0.87
			$[-3.98; -0.24]$			$[-1.80; 3.20]$
Δ Lat.			-3.16			0.59
			$[-3.73; 0.08]$			$[-0.89; 5.05]$
Δ Elevation			1.22			1.08
			$[-1.12; 2.97]$			$[-1.99; 3.68]$
Elevation Std. Dev.			0.15			0.14
			$[-0.52; 0.94]$			$[-1.01; 1.53]$
No. of vertices	5832	7113	5832	1662	2046	1662
No. of edges	15788	20045	15788	4493	5773	4493
No. of units	733	855	733	247	261	247

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

Table A5: French colonies: Sensitivity towards vector of controls

	Baseline			LDV		
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-1.22^*	-9.71^*	-11.75^*	-1.48^*	-6.05^*	-5.41^*
	$[-1.23; -1.10]$	$[-10.41; -8.88]$	$[-12.31; -8.49]$	$[-1.61; -1.32]$	$[-7.66; -4.16]$	$[-7.21; -2.57]$
Lagged dep. var.				1.17^*	0.93^*	1.07^*
				$[1.02; 1.35]$	$[0.81; 1.04]$	$[0.93; 1.28]$
Ethnic boundary	0.39^*		0.14^*	0.14		0.02
	$[0.28; 0.46]$		$[0.05; 0.24]$	$[-0.09; 0.37]$		$[-0.20; 0.28]$
Edge length		0.76^*	0.99^*		0.51^*	0.46^*
		$[0.69; 0.83]$	$[0.68; 1.07]$		$[0.34; 0.63]$	$[0.18; 0.62]$
River		0.78^*	0.68^*		0.68^*	0.65^*
		$[0.61; 0.91]$	$[0.48; 0.84]$		$[0.30; 1.05]$	$[0.15; 1.18]$
Watershed		0.06	-0.02		-0.23	0.06
		$[-0.05; 0.22]$	$[-0.19; 0.15]$		$[-0.53; 0.13]$	$[-0.29; 0.62]$
Elevation mean		0.84^*	-0.00		-0.60	-2.41
		$[0.20; 1.36]$	$[-1.22; 0.96]$		$[-2.28; 1.35]$	$[-4.46; 1.00]$
Population 1880		0.14^*	0.10^*		-0.01	-0.06
		$[0.11; 0.16]$	$[0.06; 0.14]$		$[-0.08; 0.06]$	$[-0.21; 0.03]$
Dist. coast		-0.08^*	-0.04		-0.10	-0.06
		$[-0.10; -0.04]$	$[-0.09; 0.02]$		$[-0.19; 0.00]$	$[-0.20; 0.06]$
Δ Long.			-3.67^*			2.83^*
			$[-3.89; -0.02]$			$[0.13; 5.12]$
Δ Lat.			-2.03			-0.19
			$[-2.64; 2.16]$			$[-1.74; 3.30]$
Δ Elevation			2.84			1.73
			$[-1.20; 5.12]$			$[-1.52; 3.12]$
Elevation Std. Dev.			-0.62			1.40
			$[-1.72; 1.04]$			$[-1.64; 3.16]$
No. of vertices	5042	7433	5042	1010	1825	1010
No. of edges	13880	21135	13880	2584	5125	2584
No. of units	438	590	438	117	132	117

Notes: 95% confidence intervals from parametric bootstrap in parenthesis. * Statistically significant at the 95% level.

A.6 Spatial structure

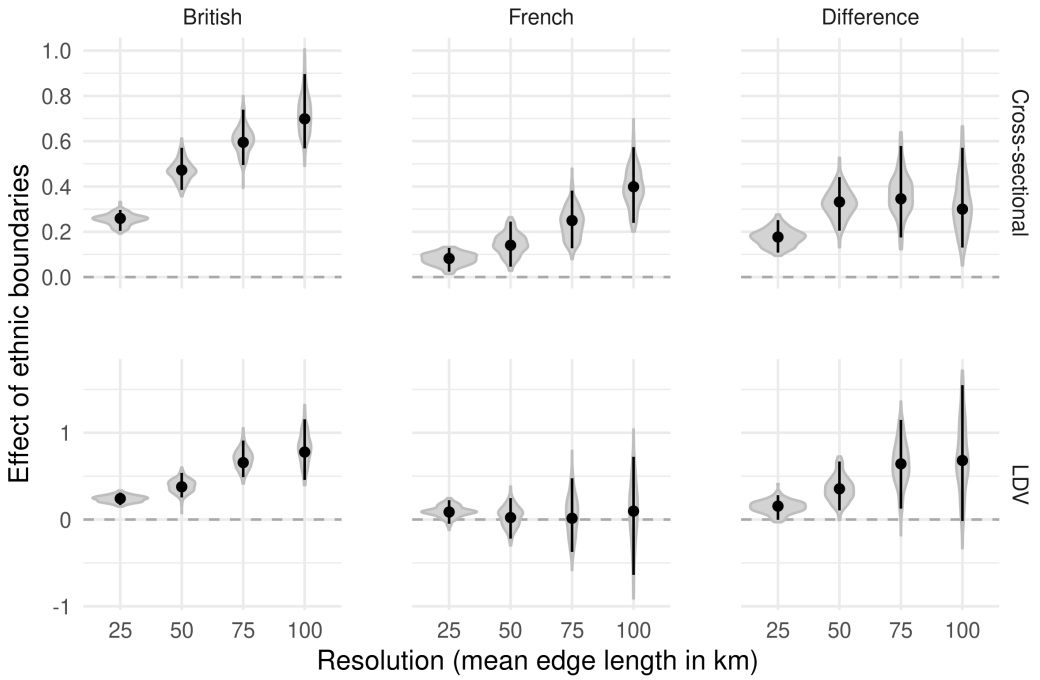


Figure A2: Effect of ethnic boundaries at varying resolutions of the spatial lattice

Note: 95% CIs and estimate distributions result from a parametric bootstrap with 160 iterations.

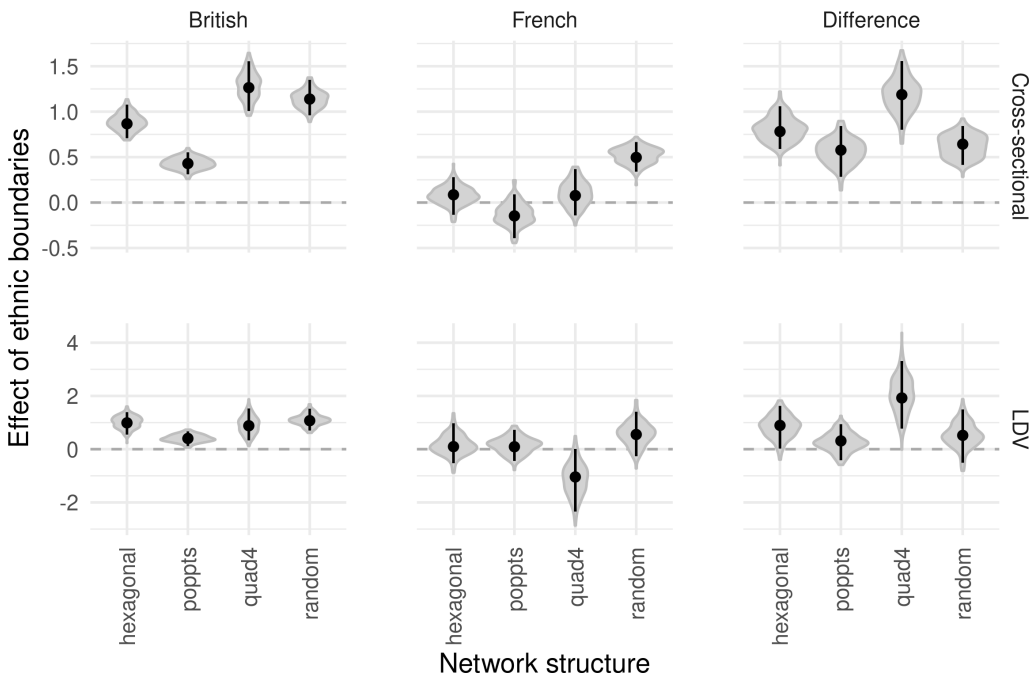


Figure A3: Effect of ethnic boundaries using hexagonal, quadratic, triangular, and random planar graph structures

Note: 95% CIs and estimate distributions result from a parametric bootstrap with 160 iterations.

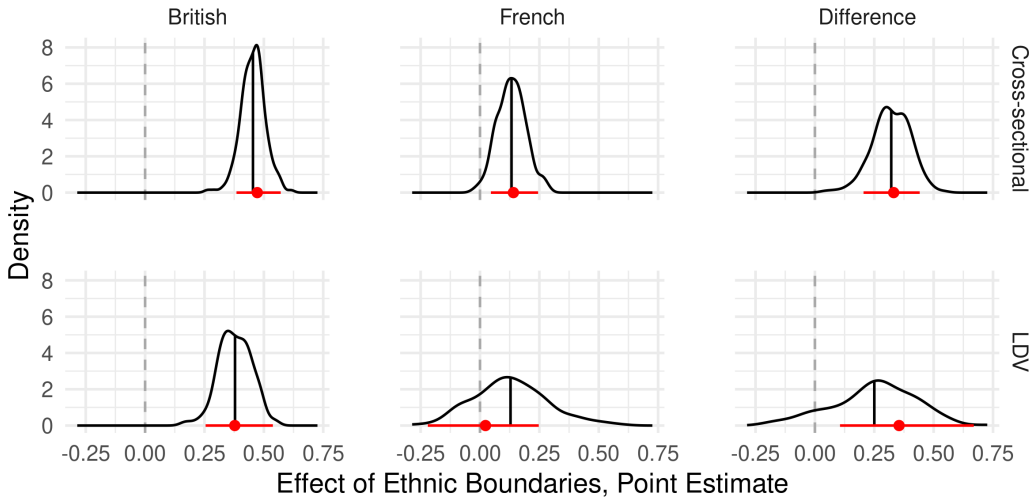


Figure A4: Point estimates of the effect of ethnic boundaries: Shifting the spatial graph

Note: Distributions result from re-estimating the main models 100 times, with data from 100 resampled planar graphs. Red estimates show the results of the main results in Table 1.

A.7 Standard regression models (with fixed effects)

Table A6: Linear Probability Model (OLS, edge-level)

Dependent Variable:	District border							
Colonizer	British				French			
Specification	Baseline		LDV		Baseline		LDV	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	-1.76**		-1.44**		-1.90**		-0.67**	
	(0.07)		(0.13)		(0.08)		(0.15)	
Ethnic boundary	0.22**	0.22**	0.16**	0.16**	0.04**	0.04**	0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Lagged dep. var.			0.40**	0.38**			0.67**	0.65**
			(0.02)	(0.02)			(0.02)	(0.02)
Fixed-effects								
Vertex 1 x Period		Yes		Yes		Yes		Yes
Vertex 2 x Period		Yes		Yes		Yes		Yes
Fit statistics								
Outcome mean	0.35	0.35	0.41	0.41	0.32	0.32	0.33	0.33
Observations	15,788	15,788	4,493	4,493	13,880	13,880	2,584	2,584
R^2	0.18	0.61	0.34	0.65	0.10	0.55	0.52	0.73
Within R^2		0.15		0.29		0.09		0.48

Custom standard-errors in parentheses
*Signif. Codes: **: 0.01, *: 0.05, +: 0.1*

Table A7: Logistic Regression Model (edge-level)

Dependent Variable:	District border							
Colonizer	British				French			
Specification	Baseline		LDV		Baseline		LDV	
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Constant	-13.14**		-12.88**		-13.48**		-10.97**	
	(0.48)		(0.98)		(0.52)		(1.70)	
Ethnic boundary	1.03**	1.98**	0.87**	1.46**	0.18**	0.32**	0.01	0.09
	(0.04)	(0.11)	(0.08)	(0.20)	(0.04)	(0.10)	(0.13)	(0.36)
Lagged dep. var.			1.95**	2.86**			3.47**	4.73**
			(0.08)	(0.20)			(0.12)	(0.34)
Fixed-effects								
Vertex 1 x Period		Yes		Yes		Yes		Yes
Vertex 2 x Period		Yes		Yes		Yes		Yes
Fit statistics								
Outcome mean	0.35	0.49	0.41	0.49	0.32	0.46	0.33	0.45
Observations	15,788	7,727	4,493	2,503	13,880	7,000	2,584	1,317
Squared Correlation	0.18	0.38	0.34	0.51	0.10	0.29	0.52	0.66
Pseudo R^2	0.15	0.31	0.28	0.42	0.09	0.23	0.44	0.55

Custom standard-errors in parentheses
*Signif. Codes: **: 0.01, *: 0.05, +: 0.1*