DISCUSSION PUBLISHERS
Research Center · Open Access Edition

Pre-colonial Ethnic Institutions and Contemporary African Development

Stelios Michalopoulos, Elias Papaioannou

Discussion Research ID	DPR-2026-1409
Year	2012
Field	Social Sciences
DOI	10.3386/w18224
Author ORCID	0009-0009-3058-5189
Licence	CC-BY
Original source	National Bureau of Economic Research
Permanent link	https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9
Publisher	Discussion Publishers discussionpublishers.com

Abstract

Acemoglu for detailed comments and useful

Republished by Discussion Publishers under the original CC-BY licence. Copyright remains with the original authors; please cite the original work.

Cite as: Stelios Michalopoulos, Elias Papaioannou (2012). Pre-colonial Ethnic Institutions and Contemporary . Discussion Publisher

NBER WORKING PAPER SERIES

PRE-COLONIAL ETHNIC INSTITUTIONS AND CONTEMPORARY AFRICAN DEVELOPMENT

Stelios Michalopoulos
Elias Papaioannou

Working Paper 18224
http://www.nber.org/papers/w18224

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2012

We thank 4 anonymous referees and the Editor for many insightful comments and useful suggestions. We thank seminar participants at Dartmouth, Tufts, Oxford, Vienna, Brown, Harvard, Stanford, UC-Berkeley, UC-Davis, NYU, AUEB, the CEPR Development Economics Workshop, the World Bank, the IMF, the NBER Political Economy Meetings, the NBER Summer Institute Meetings in Economic Growth and Income Distribution and the Macroeconomy for valuable comments. We also benefited from discussions with Yannis Ioannides, Rafael La Porta, Antonio Ciccone, Rob Johnson, Raphael Frank, Jim Feyrer, Ross Levine, Avner Greif, Jeremiah Dittmar, David Weil, Sandip Sukhtankar, Quamrul Ashraf, Oded Galor, Ed Kutsoati, Pauline Grosjean, Enrico Perotti, Pedro Dal Bó, Nathan Nunn, Raquel Fernandez, Jim Robinson, and Enrico Spolaore. We are particularly thankful to Andy Zeitlin, Melissa Dell, Andei Shleifer, Nico Voigtländer, and Daron Acemoglu for detailed comments and useful suggestions. We also thank Nathan Nunn for providing the digitized version of Murdock's Tribal Map of Africa. This paper draws on material from our previous paper titled "Divide and Rule or the Rule of the Divided? Evidence from Africa". All errors are our sole responsibility. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Stelios Michalopoulos and Elias Papaioannou. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Pre-colonial Ethnic Institutions and Contemporary African Development
Stelios Michalopoulos and Elias Papaioannou
NBER Working Paper No. 18224
July 2012
JEL No. N17,O4,O43,Z1

ABSTRACT

We investigate the role of deeply-rooted pre-colonial ethnic institutions in shaping comparative regional development within African countries. We combine information on the spatial distribution of ethnicities before colonization with regional variation in contemporary economic performance, as proxied by satellite images of light density at night. We document a strong association between pre-colonial ethnic political centralization and regional development. This pattern is not driven by differences in local geographic features or by other observable ethnic-specific cultural and economic variables. The strong positive association between pre-colonial political complexity and contemporary development obtains also within pairs of adjacent ethnic homelands with different legacies of pre-colonial political institutions.

Stelios Michalopoulos
Brown University
Department of Economics
64 Waterman Street
Providence, RI 02912
and NBER
smichalo@brown.edu

Elias Papaioannou
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
papaioannou.elias@gmail.com

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

1 Introduction

There has been ample research on the institutional origins of African (under)development both in economics and the broader literature in social sciences; yet the two strands have followed somewhat different paths. On the one hand, influenced by the studies of Acemoglu *et al.* (2001, 2002) and La Porta *et al.* (1997, 1998), the empirical literature in economics has mainly focused on the impact of colonization in comparative development primarily via its effect on contractual institutions and property rights protection at the national level (see Acemoglu and Johnson (2005)). On the other hand, the African historiography has invariably stressed the role of deeply-rooted, ethnic institutional characteristics (see Herbst (2000) for a summary). Motivated by the richness of anecdotal evidence and case studies documenting the importance of ethnic-specific institutional traits, in this study we explore in a systematic way the relationship between pre-colonial ethnic institutions, political centralization in particular, and regional development.

We utilize data from the pioneering work of Murdock (1959, 1967), who has mapped the spatial distribution of African ethnicities and compiled various quantitative indicators reflecting political institutions, cultural, and economic traits of several ethnic groups around colonization. To overcome the paucity of economic indicators across African ethnic homelands, we combine the anthropological data with satellite images of light density at night.

Our analysis shows that the complexity and hierarchical structure of pre-colonial ethnic institutions correlate significantly with contemporary regional development, as reflected in light density at night. This correlation does not necessarily imply a causal relationship, because one cannot rule out the possibility that other ethnic characteristics and hard-to-account-for factors drive the association. Nevertheless this correlation obtains across numerous permutations. First, it is robust to an array of controls related to the disease environment, land endowments, and natural resources at the local level. Accounting properly for geography is important as there is a fierce debate in the literature on whether the correlation between institutional and economic development is driven by hard-to-account-for geographical features. Second, the strong positive association between pre-colonial political centralization and regional development retains its economic and statistical significance, when we solely examine within-country variation. Including country fixed effects is crucial since we are able to account for all country-specific, time-invariant features. Third, regressing luminosity on a variety of alternative pre-colonial ethnic characteristics, such as occupational specialization, economic organization, the presence of polygyny, slavery, and proxies of early development, we find that political centralization is the only robust correlate of contemporary economic performance. This reassures that the uncovered positive association does not reflect differences in observable cultural and economic

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

attributes across African ethnicities. Fourth, the positive correlation between ethnic political complexity and regional development prevails when we limit our analysis within pairs of neighboring homelands falling in the same country where ethnicities with different pre-colonial institutions reside.

These patterns obtain both when the unit of analysis is the ethnic homeland and when we exploit the finer structure of the luminosity data to obtain multiple observations (pixels) for each homeland. Hence, although we do not have random assignment in ethnic institutions, the results clearly point out that traits manifested in differences in the pre-colonial institutional legacy matter crucially for contemporary African development.

Ethnic Institutions: Past and Present There was significant heterogeneity in political centralization across African ethnicities before colonization (Murdock (1967)). At the one extreme, there were states with centralized administration and hierarchical organization such as the Shongai Empire in Western Africa, the Luba kingdom in Central Africa, and the kingdoms of Buganda and Ankole in Eastern Africa. At the other extreme, there were acephalous societies without political organization beyond the village level, such as the Nuer in Sudan or the Konkomba in Ghana and Togo. The middle of the spectrum occupied societies organized in large chiefdoms and loose alliances, such as the Ewe and the Wolof in Western Africa. While these societies lacked statehood, they tended to have conflict resolution mechanisms and a somewhat centralized decision making process (Diamond (1997)).

The advent of the Europeans in Africa had limited impact on these pre-existing local political structures. This was because colonization was (with some exceptions) quite limited both regarding timing and location (Herbst (2000)). Mamdani (1996) argues that, in fact, the European colonizers in several occasions strengthened tribal chiefs and kings via their doctrine of indirect rule. In the eve of African independence some countries attempted to limit the role of ethnic institutions; however, the inability of African states to provide public goods and broadcast power beyond the capitals, led African citizens to continue relying on the local ethnic-specific structures rather than the national government (Englebert (2009)). Herbst (2000), for example, notes that in Mauritania, Mozambique, Niger, Nigeria and Chad while new states initially marginalized local chiefs, when they realized the extraordinary difficulties in governing rural areas, they quickly invited them back.

There is ample evidence pointing to the ongoing importance of ethnic-specific institutions. First, ethnic leaders and chiefs enjoy considerable support and popularity across local communities (e.g. Baldwin (2010)). Second, both survey data and case studies show that local chiefs have significant power in allocating land rights. Analyzing data from the Afrobarometer

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Surveys, Logan (2011) documents that ethnic institutions are instrumental in assigning property rights and resolving disputes.[1] Along the same lines, Goldstein and Udry (2008) show that informal ethnic institutions and local chiefs today exert significant de facto power in assigning land in rural Ghana (see also Bubb (2012)). Third, in many countries local leaders collect taxes and provide some basic public goods (e.g. Glennerster, Miguel, and Rothenberg (2010) and Acemoglu, Reed, and Robinson (2012)). Fourth, since the early 1990s many countries (15 out of 39 according to Herbst (2000)) have passed legislation or even constitutional amendments (in the case of Uganda and Ghana) formally recognizing the role of ethnic institutional structures in settling property rights disputes and enforcing customary law (see Baldwin (2011)).

The African historiography has proposed various channels via which ethnic institutions shape contemporary economic activity. First, Herbst (2000) and Boone (2003) argue that in centralized societies there is a high degree of accountability of local chiefs. For example, in ethnic groups that had a state structure, poorly-performing local rulers could be replaced by the king or superior administrators. Even nowadays some ethnic groups have assemblies and supreme officials that make local chiefs accountable. Second, Diamond (1997) and Acemoglu and Robinson (2012) describe how ethnic groups that formed large states, had organized bureaucracies providing policing and other public goods. Third, in centralized ethnicities there was access to some formal legal resolution mechanism and some form or property rights steadily emerged (Herbst (2000). Fourth, others have argued that centralized societies were quicker in adopting Western technologies, because the colonizers collaborated more strongly with politically complex ethnicities (Schapera (1967, 1970)). Fifth, tribal societies with strong political institutions have been more successful in getting concessions both from colonial powers and from national governments after independence. For example, Acemoglu and Robinson (2012) describe how the Tswana leaders travelled from Bechuanaland (current Botswana) to England and convinced the British government to allow for a greater degree of autonomy.[2]

Related Literature Our study contributes to the literature on the role of pre-colonial, institutional and cultural features in African development reviewed above (Fortes and Evans-Pritchard (1940), Schapera (1967), Stevenson (1968), Goody (1971), Bates (1983), Robinson (2002), Boone (2003), Englebert (2009); Besley and Reynal-Querol (2012)). The most closely related line of research is that of Gennaioli and Rainer (2006, 2007) who present cross-

[1] Logan (2011) shows that ethnic leaders are equally important as the local and central governments is assigning property rights. Respondents tend to rely more on local chiefs and ethnic institutional structures for the resolution of disputes as compared to national and local government. Ethnic-specific political actors and institutions play also some role in the provision of education and health.

[2] Mamdani (1996), nevertheless, differs in his assessment on the beneficial contemporary role of hierarchical pre-colonial structures arguing that the legacy of indirect rule in Africa through traditional chiefs was a basis for post-independence poor institutional and economic performance.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

country evidence showing that pre-colonial political centralization correlates positively with public goods provision and contemporary institutions. We advance this literature by establishing that, unlike other observable ethnicity-level variables, pre-colonial ethnic institutions captured by the degree of political complexity are systematically linked to contemporary regional development within countries as well as within pairs of contiguous ethnic homelands.

Our study also belongs to a growing body of work on the historical origins and the political economy of African development. Nunn (2008) stresses the importance of the slave trades, while Huillery (2009), Berger (2009) and Arbesu (2011) quantify the long-run effects of colonial investments and tax collection systems. Englebert, Tarango, and Carter (2002), Alesina, Easterly, and Matuszeski (2011) and Michalopoulos and Papaioannou (2011b) examine the negative effects of the improper colonial border design during the Scramble for Africa.

On a broader scale our work relates to the literature on the institutional origins of contemporary development (see Acemoglu, Johnson, and Robinson (2005) for a review). Our micro approach enables us to overcome problems inherent to the cross-country framework adding to a vibrant body of research that examines the within-country impact of various historical institutional arrangements (e.g. Banerjee and Iyer (2005); Iyer (2010); Dell (2010)). Moreover, our within-country results linking pre-colonial political centralization to contemporary regional development complement the cross-country findings of Bockstette, Chanda, and Putterman (2002) on the beneficial long-run consequences of statehood.

Finally, our work has implications for the literature on state capacity that examines the origins and consequences of weak states' inability to monopolize violence, collect taxes, and protect private property (e.g. Besley and Persson (2011)). Our work shows that in presence of weak states, local institutions (ethnic in Africa) may fill in the void created by the limited penetration of national institutions (Michalopoulos and Papaioannou (2012)).

Paper Structure In the next section we present the pre-colonial ethnic institutional measures and discuss the luminosity data. We cross-validate our main data and report descriptive statistics illustrating the relationship between ethnic-level political organization and development. Section 3 presents our results at the ethnic homeland level. First, we lay down the general econometric framework and report our benchmark within-country estimates. Then, we examine the role of other ethnic-specific features. In Section 4 we first report our baseline results at the pixel level and then examine whether regional development differs systematically across contiguous territories where ethnicities with a different degree of political centralization reside. In section 5 we conclude discussing directions for future research.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

2 Data

2.1 Data on the location of historical ethnic homelands

The starting point in compiling our dataset is George Peter Murdock's (1959) Ethnolinguistic Map that portrays the spatial distribution of ethnicities across Africa in the eve of European colonization in the mid/late 19^{th} century. Murdock's Map (Figure 1a) includes 843 tribal areas (the mapped groups correspond roughly to levels $7-8$ of the Ethnologue's language family tree); 8 areas are classified as uninhabited upon colonization and are therefore excluded. We also drop the Guanche, a group in the Madeira islands that is currently part of Portugal. One may wonder how much the spatial distribution of ethnicities across the continent has changed over the past 150 years. Reassuringly, using individual data from the Afrobarometer Nunn and Wantchekon (2011) show a 0.55 correlation between the location of the respondents in 2005 and the historical homeland of their ethnicity as identified in Murdock's map. Similarly, Glennerster, Miguel, and Rothenberg (2010) document that in Sierra Leone after the massive displacement of the $1991-2002$ civil war there has been a systematic movement of individuals towards their ethnic group's historical homeland. We intersect Murdock's ethnolinguistic map with the 2000 Digital Chart of the World that portrays contemporary national boundaries to identify partitioned ethnicities and assign each area to the respective country.

2.2 Ethnic Institutional Traits

In work following the mapping of African ethnicities, Murdock (1967) produced an Ethnographic Atlas (published in twenty nine installments in the anthropological journal *Ethnology*) that coded around 60 variables, capturing cultural, geographical, and economic characteristics of $1,270$ ethnicities around the world. We assigned the 834 African ethnicities of Murdock's Map of 1959 to the ethnic groups in his Ethnographic Atlas of 1967. The two sources do not always use the same name to identify an ethnic group. Utilizing several sources and the updated version of Murdock's Atlas produced by Gray (1999), we match 534 ethnicities from the Ethnographic Atlas to 490 ethnic homelands in Murdock's Map (Figure 1a).[3]

We measure pre-colonial political institutions using Murdock's (1967) "Jurisdictional Hierarchy beyond the Local Community Level" index (see also Gennaioli and Rainer (2006, 2007)). This is an ordered variable, ranging from 0 to 4, that describes the number of political jurisdictions above the local (usually village) level for each ethnicity. A zero score indicates stateless societies "*lacking any form of centralized political organization*". A score of 1 indicates

[3]In 34 instances an ethnic homeland from Murdock's Map is assigned to more than one groups in the Ethnographic Atlas; in these cases we assigned to the ethnic homeland the median value of the ethnic institutions index.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

petty chiefdoms; a score of 2 designates paramount chiefdoms; and 3 and 4 indicate groups that were part of large states. Murdock (1967) explicitly excludes colonial regimes and attempts to capture political complexity before European colonization. This classification is similar to Diamond (1997) who distinguishes between four main types of societal arrangements: bands, tribes, chiefdoms, and centralized states. Figure 1*b* illustrates the significant heterogeneity in pre-colonial political organization across African groups. Examples of ethnicities without any level of political organization above the local level include the Bura and the Lango in Uganda. Examples of tribes belonging to petty chiefdoms are the Mende in Sierra Leone and the Ibo of Nigeria. The Mbundu in Angola and the Zerma in Niger were part of paramount chiefdoms, while the Ndebele in Zimbabwe and the Mossi in Burkina Faso are societies that were parts of states before colonization. The Bubi in Equatorial Guinea and the Beduin Arabs are the only groups classified as having been part of large complex states (score of 4).

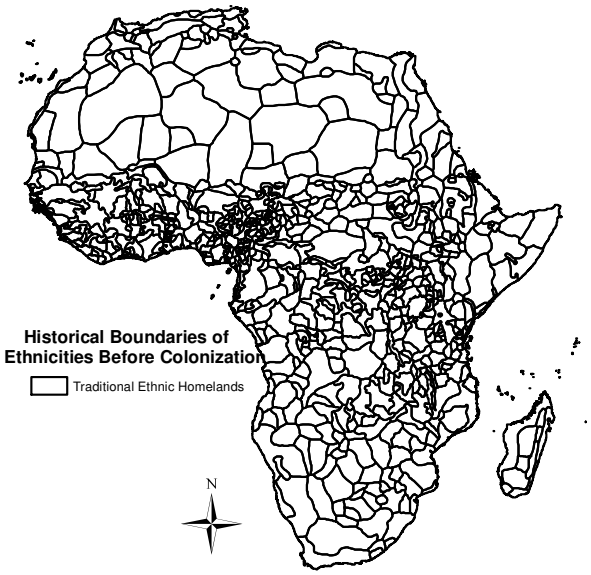


Figure 1*a*: Ethnic Boundaries

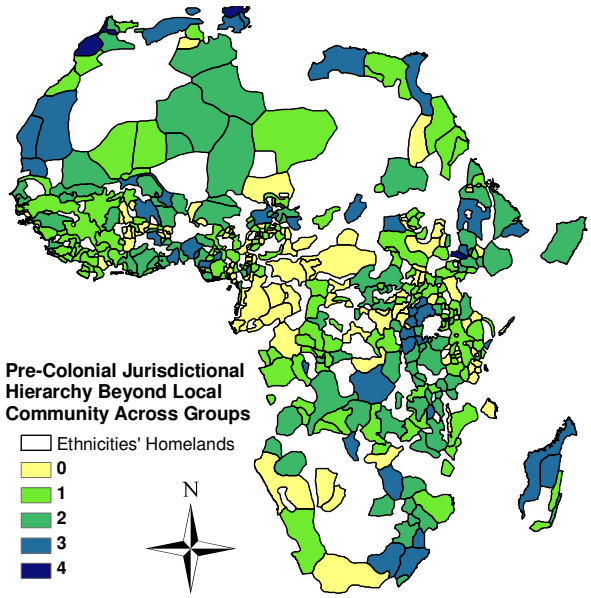


Figure 1*b*: Ethnic Pre-Colonial Institutions

Cross-validation of Murdock's Jurisdictional Hierarchy Index We cross-validated Murdock's data going over the African historiography. Our reading of the literature suggests that the jurisdictional hierarchy index -while not perfect- is in accordance with works describing the degree of political complexity in pre-colonial Africa. Murdock (1967) classifies as centralized the dominant ethnic groups of most major pre-colonial states. For example, the Ankale and the Buganda, which were the central ethnic groups in the strong kingdoms of Eastern Africa, get a score of 3. The same applies to other ethnic groups that were part of large states, such as the Zulu and the Swazi in South Africa, the Ife and the Igala in Nigeria, and the Shongai in Mali (Fortes and Evans-Pritchard (1940); Goody (1971)).

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Murdock also seems to correctly identify stateless ethnicities. The jurisdictional hierarchy index equals zero or one for the Amba, the Konkomba, the Tiv, the Dinka, and the Lugbara, in line with the analysis of Middleton and Tait (1958) who describe them as acephalous societies. Regarding the Amba, for example, Winter (1958) writes that "*the village is the largest unilateral unit of power*", whereas Tait (1958) characterizes the Konkomba as an ethnic group that is "*organized locally, without formal laws, and central authority.*" Likewise, the Lobbi is classified as stateless in line with Goody (1971) who characterizes them as "*people with no state organization at all*". The classification also identifies properly societies with intermediate levels of political centralization (paramount chiefdoms). The Nupe in Nigeria, the Bemba in Zambia, and the Ngwato in Botswana which formed small states get a score of 2 (Fortes and Evans-Pritchard (1940)). Clearly there is some subjectivity in Murdock's characterization of ethnic-specific institutional structures. Yet to the extent that these biases are not systematic, this should lead to attenuation and as such our estimates will be on the conservative side.

2.3 Satellite Light Density at Night

The nature of our study requires detailed spatial data on economic development. To the best of our knowledge, geocoded high-resolution measures of economic development spanning all Africa are not available. To overcome this limitation we use satellite light density at night to proxy for local economic activity. The luminosity data come from the Defense Meteorological Satellite Program's Operational Linescan System that reports images of the earth at night captured from $20:30$ to $22:00$ local time. The satellite detects lights from human settlements, fires, gas flares, lightning, and the aurora. The measure is a six-bit number (ranging from 0 to 63) calculated for every 30-second area (approximately 1 square kilometer). The resulting annual composite images of time-stable lights are created by overlaying all images captured during a calendar year, dropping images where lights are shrouded by cloud or overpowered by the aurora or solar glare (near the poles), and removing ephemeral lights like fires and lightning. We construct average light density per square kilometer for 2007 and 2008 averaging across pixels at the desired level of aggregation.

The use of luminosity data as a proxy for development builds on the recent contribution of Henderson, Storeygard, and Weil (2011) and previous works (e.g. Elvidge, Baugh, Kihn, Kroehl, and Davis (1997); Doll, Muller, and Morley (2006)) showing that light density at night is a robust proxy of economic activity (see also Pinkovskiy (2011)). These studies establish a strong within-country correlation between light density at night and GDP levels and growth rates. There is also a strong association between luminosity and access to electricity and public-goods provision, especially across low income countries (see Min (2008)). Even Chen

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

and Nordhaus (2011), who emphasize some problems of the satellite image data, argue that luminosity can be quite useful for regional analysis in war-prone countries with poor quality income data. Luminosity data are subject to saturation and blooming. Saturation occurs at a level of light similar to that in the urban centers of rich countries and results in top-coded values. Blooming occurs because lights tend to appear larger than they actually are, especially for bright lights over water and snow. These issues, however, are less pressing within Africa. First, there are very few instances of top-coding (out of the $30,457,572$ pixels of light density only 0.00017% are top-coded). Second, since luminosity is quite low across African regions, blooming (bleeding) is not a major problem. Moreover, in the within-contiguous-ethnic-homelands analysis blooming would work against our hypothesis.

Cross-Validation: Satellite Light Density and Regional Development In our empirical analysis we primarily explore within-country variation. Thus we examined the relationship between luminosity and economic performance using micro-level data from the Demographic and Health Surveys (DHS) (see the Supplementary Appendix for additional cross-validation checks). The DHS team in each country produces a composite wealth index, based on individual responses on whether households have access to basic public goods, such as electrification, clean water, etc. We examine the correlation between log light density and the wealth index within four large countries from different parts of Africa; Nigeria from Western Africa, Tanzania from Eastern Africa, the Democratic Republic of Congo from Central Africa, and Zimbabwe from Southern Africa. We derive the average wealth index across households for each enumeration area (usually a village or a small town) and associate it with light density of each DHS area using a radius of $10km$. Figures $2a - 2d$ offer a visual representation of the significant correlation (around 0.75) between luminosity and the composite wealth index. To mitigate concerns that the correlation is driven by outliers we drop the top 1% of lit areas effectively excluding the capital city and few other major urban hubs (the dashed line representing the median regression slope is almost identical to the LS line).

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

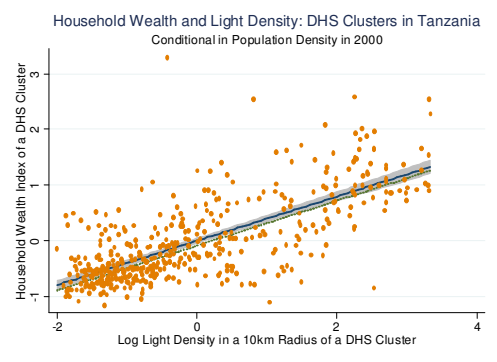


Figure 2*a*

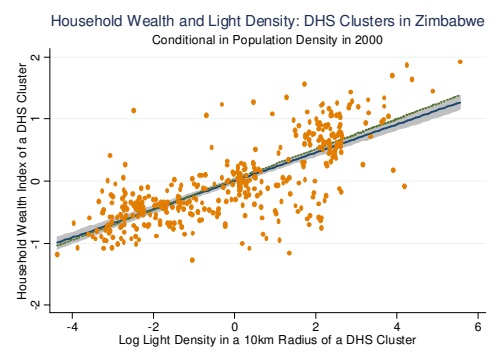


Figure 2*b*

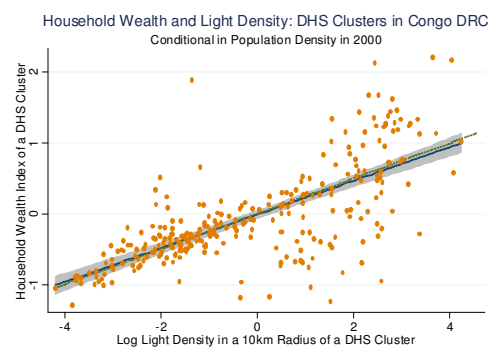


Figure 2*c*

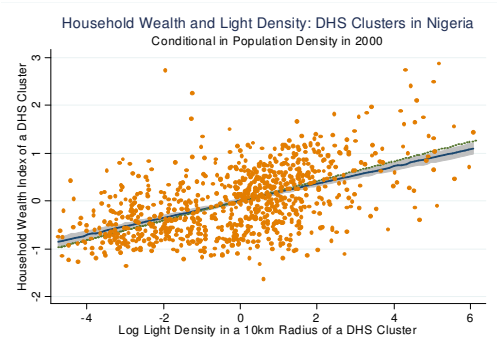


Figure 2*d*

2.4 Summary Statistics

Figures 3*a* and 3*b* portray the distribution of luminosity across African ethnic homelands. In Figure 3*a* we aggregate the luminosity data at the country-ethnic homeland level, which serves as our unit of analysis in Section 3. In Figure 3*b* we divide the continent into pixels of $12.5 * 12.5$ decimal degrees (approximately $12.5km * 12.5km$) and map lit and unlit pixels. Table 1 reports descriptive statistics of the luminosity data both at the ethnic-country homeland level and at the pixel level. The mean value of luminosity at the ethnic homeland level is 0.368. The median is significantly lower, 0.022, because of few areas where light density is extremely high. There are 7 observations where luminosity exceeds 7.4 and 14 observations where light density exceeds 4.06. On average 16.7% of all populated pixels are lit, while in the remaining pixels satellite sensors do not detect the presence of light.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

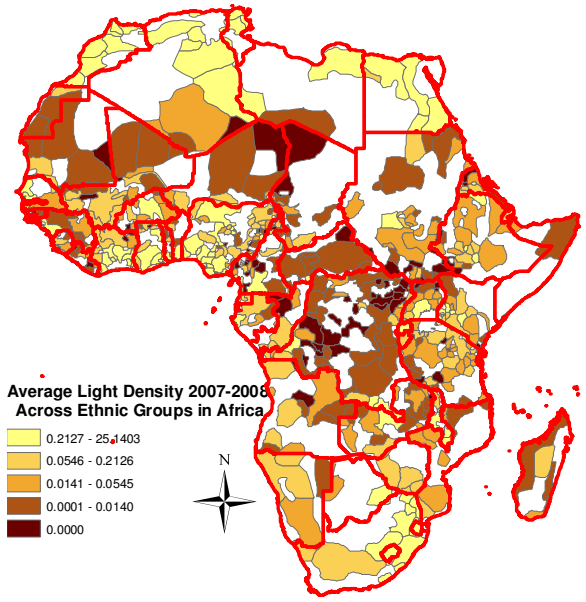


Figure 3*a*: Luminosity at the Ethnic Homeland

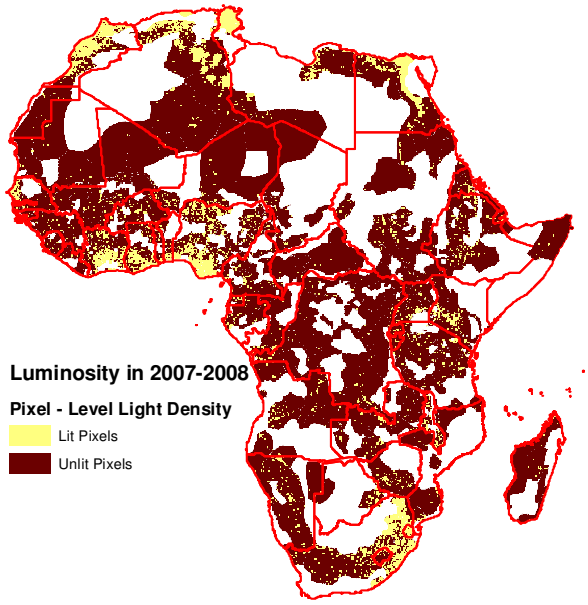


Figure 3*b*: Pixel-Level Luminosity

The summary statistics reveal large differences in luminosity across homelands where ethnicities with different pre-colonial political institutions reside. The mean (median) luminosity in the homelands of stateless societies is 0.248 (0.017) and for petty chiefdoms the respective values are 0.269 (0.013); and only 10% and 12.9% of populated pixels are lit, respectively. Focusing on groups that formed paramount chiefdoms, average (median) luminosity is 0.311 (0.037), while the likelihood that a pixel is lit is 16.9%. Average (median) luminosity in the homelands of ethnicities that were part of centralized states before colonization is 0.993 (0.082). On average 30.2% of pixels falling in the homelands of highly centralized groups are lit, three times more than the respective likelihood for stateless societies. Light density in the homelands of pre-colonial states is significantly higher, even when compared to groups organized as paramount chiefdoms. The mean (median) difference is 0.68 (0.045); and simple test of means (medians) suggest that these differences are significant at the 99% confidence level. The descriptive statistics reveal that regional development across ethnic homelands correlates with the form of the pre-colonial political organization. Light density increases significantly when one moves from the homelands of stateless at the time of colonization societies and petty chiefdoms to the homelands of ethnicities organized as paramount chiefdoms; and luminosity is even higher in the homelands of ethnicities that were part of large states.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

3 Ethnic Homeland Analysis

3.1 Empirical Framework

To formally examine the relationship between pre-colonial ethnic institutions and development across ethnic homelands, we estimate variants of the following specification:

$$y_{i,c} = a_0 + \gamma IQL_i + X'_{i,c}\Phi + \lambda PD_{i,c} + a_c + \varepsilon_{i,c}. \quad (1)$$

The dependent variable, $y_{i,c}$, reflects the level of economic activity in the historical homeland of ethnic group i in country c, as proxied by light density at night. IQL_i denotes local ethnic institutions as reflected in the degree of jurisdictional hierarchy beyond the local level. For ethnicities that fall into more than one country each partition is assigned to the corresponding country c. For example, regional light density in the part of the Ewe in Ghana is assigned to Ghana, while the adjacent region of the Ewe in Togo is assigned to Togo.[4] In most specifications we include country fixed effects (a_c), so as to exploit within-country variation. While fixed effects estimation may magnify problems of measurement error (by absorbing a sizable portion of the variation), it accounts for differences in national policies, the quality of national institutions, the identity of the colonizer, the type of colonization, as well as other country-wide factors.

A merit of our regional focus is that we can account for local geography and other factors (captured in vector $X_{i,c}$). In many specifications we include a rich set of controls, reflecting land endowments (elevation and area under water), ecological features (a malaria stability index, land suitability for agriculture), and natural resources (diamond mines and petroleum fields). Several studies suggest the inclusion of these variables. First, Nunn and Puga (2012) show that elevation and terrain ruggedness have affected African development both via goods and via slave trades. Second, the inclusion of surface under water accounts for blooming in the lights data and for the potential positive effect of water streams on development via trade. Third, controlling for malaria prevalence is important as Gallup and Sachs (2001) and subsequent studies have shown a negative impact of malaria on development. Fourth, there is a vast literature linking natural resources like oil and diamonds to development (e.g. Ross (2006)). Fifth, Michalopoulos (2012) shows that differences in land suitability and elevation across regions lead to the formation of ethnic groups, whereas Ashraf and Galor (2011) show that land quality is strongly correlated with pre-colonial population densities. We also control for the location of each ethnic area inside a country augmenting the specification with the distance of the centroid of each ethnic

[4] After intersecting Murdock's ethnolinguistic map with the 2000 Digital Chart of the World we drop ethnic partitions of less than 100 km^2, as such tiny partitions are most likely due to the lack of precision in the underlying mapping.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

group i in country c from the respective capital, the national border, and the nearest sea coast. The coefficient on distance from the capital reflects the impact of colonization and the limited penetration of national institutions. Distance to the national border captures the potentially lower level of development in border areas whereas distance to the sea reflects the effect of trade as well as the penetration of colonization. In several specifications we control for log population density ($PD_{i,c}$) though the latter is likely endogenous to ethnic institutional development. Appendix Table 2 reports the summary statistics for all control variables.

The distribution of luminosity across ethnic homelands is not normal, as (i) a significant fraction (around 24%) of the observations takes on the value of zero and (ii) we have a few extreme observations in the right tail of the distribution (Appendix Figure 1a). To account for both issues we use as dependent variable the log of light density adding a small number (($y_{i,c} \equiv \ln(0.01 + LightDensity_{i,c})$, Appendix Figure 1$b$).[5] This transformation ensures that we use all observations and that we minimize the problem of outliers. We also estimate specifications ignoring unlit areas ($y_{i,c} \equiv \ln(LightDensity_{i,c})$), as in this case the dependent variable is normally distributed (Appendix Figure 1c). Moreover, in our pixel-level analysis, where we focus on regions of $0.125 * 0.125$ decimal degrees, we use as dependent variable a dummy that takes on the value one when the pixel is lit and zero otherwise.

In all specifications we employ the approach of Cameron, Gelbach, and Miller (2011) and cluster standard errors at the country level and at the ethnic-family level. Murdock assigns the 834 groups into 96 ethnolinguistic clusters/families. Double-clustering accounts for the fact that ethnicity-level characteristics are likely to be correlated within an ethnolinguistic family. Moreover, clustering at the ethnic-family level is appropriate because partitioned ethnicities appear more than once. Finally, the multi-way clustering method allows for arbitrary residual correlation within both dimensions and thus accounts for spatial correlation (Cameron, Gelbach, and Miller (2011) explicitly cite spatial correlation as an application of the multi-clustering approach). We also estimated standard errors accounting for spatial correlation of an unknown form using Conley's (1999) method. The two approaches yield similar standard errors; and if anything the two-way clustering produces somewhat larger standard errors.

3.2 Preliminary Evidence

Table 2 reports cross-sectional LS specifications that associate regional development with pre-colonial ethnic institutions. Below the estimates we report both double-clustered (in parentheses) and Conley's (in brackets) standard errors.[6] Column (1) reports the unconditional

[5] In the previous draft of the paper we added one to the luminosity data before taking the logarithm finding similar results.

[6] Conley's method requires a cutoff distance beyond which the spatial correlation is assumed to be zero; we experimented with values between $100km$ and $3000km$. We report errors with a cutoff of $2000km$ that delivers

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

estimate. In line with the pattern shown in Table 1, the coefficient on the jurisdictional hierarchy index is positive (0.411) and highly significant. The coefficient remains significant when we control for population density in column (2). In column (3) we control for distance to the capital city, distance to the border, and distance to the coast ("location controls") whereas in column (4) we augment the specification with a rich set of geographic features.[7] Adding these controls reduces the size of the coefficient on the jurisdictional hierarchy index; yet the estimate retains significance at the 99% confidence level. In columns (5) and (6) we examine whether the strong positive correlation between pre-colonial political institutions and regional development is driven by differences in national institutional quality or income per capita, respectively. This check is motivated by Gennaioli and Rainer (2006) who show that across African countries there is a positive association between the average level of pre-colonial political centralization and contemporary national institutions (in our sample the correlation between the rule of law in 2007 and the jurisdictional hierarchy index is 0.19). Conditioning on either (or both) of these country-level measures of institutional and economic development has little effect on our main result. The coefficient on jurisdictional hierarchy remains intact.[8]

3.3 Benchmark Fixed Effects Estimates

The positive correlation between local institutions and regional development may be driven by a myriad of nationwide features. In Table 3 we estimate country fixed effects specifications associating regional development with pre-colonial ethnic institutions. Table 3*A* reports estimates using all observations. In Table 3*B* we focus on the intensive margin of luminosity. By doing so we (i) account for nonlinearities in the dependent variable and (ii) focus on densely populated areas (since non-lit areas have a median population density of 11.06 people per square kilometer whereas lit regions have a median of 35.54).

Jurisdictional Hierarchy beyond the Local Community The coefficient on the jurisdictional hierarchy index in column (1) of table 3A is 0.326 and highly significant.[9] The estimate is moderately smaller than the analogous unconditional specification reported in Table 2 - column (1), suggesting that common-to-all-ethnicities country-level factors are not driving the

the largest in magnitude standard errors.

[7] Land suitability for agriculture, which reflects climatic and soil conditions, enters most models with a positive and significant estimate. The malaria stability index enters with a statistically negative estimate. The coefficient on land area under water is positive and in many specifications significant. Elevation enters with a negative estimate which is significant in some models. The petroleum dummy enters always with a positive and significant coefficient. The diamond dummy enters in most specifications with a negative estimate.

[8] We lose three observations when we condition on the rule of law or GDP, because we lack data on Western Sahara. The results are unaffected if we assign the Western Saharan ethnic homelands to Morocco.

[9] When we add country fixed effects we lose one observation. This is because in Swaziland we have only one group, the Swazi.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

positive cross-sectional correlation.[10] Standard errors drop also with the inclusion of country fixed effects and as such the statistical significance of the estimate is unaffected. In column (2) we augment the specification with distance to the coast, distance to the border, distance to the capital and the rich set of geographic controls. The coefficient on the jurisdictional hierarchy beyond the local community index retains its statistical and economic significance. In column (3) we control for population density, while in column (4) we control jointly for geography, location, and population density. Compared to column (1) the estimate on jurisdictional hierarchy falls by almost a half. This is not surprising as according to the African historiography (e.g. Stevenson (1968), Fenske (2009)) there is a strong interplay between geography, population density, and political complexity.[11] The size of the coefficient in column (3) - Table 3*A* implies that a one-standard-deviation increase in the jurisdictional hierarchy index (which corresponds to approximately one-unit increase; see Appendix Table 2) is associated with a 0.12 standard-deviation increase in luminosity. This magnitude is similar to the one documented by Nunn and Wantchekon (2011) in their within-country cross-regional examination of the effect of the African slave trades on trust (where they report "beta" coefficients in the range of $0.10 - 0.16$).

Political Centralization In columns (5) and (8) we use an alternative indicator of pre-colonial political institutions. Following Gennaioli and Rainer (2006, 2007) we define a dummy variable that takes the value of zero when the group lacks any political organization beyond the local level or is organized as a petty chiefdom; the index equals one if Murdock classifies the ethnicity as being a large chiefdom or part of a state. Experimenting with the re-scaled index is useful because the aggregation may account for measurement error in the jurisdictional hierarchy index. Moreover, the binary classification is in line with the distinction of African pre-colonial political systems into centralized ones and those lacking any form of centralized political authority.[12] The coefficient on political centralization is positive and highly significant. The estimate retains significance, when we control for geography (in (6)), current

[10] The Hausman-type test that compares the coefficient on the jurisdictional hierarchy index of the cross-sectional to the within-country model, suggests that one cannot reject the null hypothesis of coefficient equality.

[11] Since population density may be both a cause and an effect of ethnic institutions, the specifications where we also control for population density should be cautiously interpreted. Following Angrist and Pishcke's (2008) recommendation we also used lagged (at independence) population density. In these models (not reported) the estimates on the ethnic institutions measures are larger (and always significant at the 95% level).

[12] Fortes and Evans-Pritchard (1940) argue that "*the political systems fall into two main categories. One group consists of those societies which have centralized authority, administrative machinery, and judicial institutions-in short, a government-and in which cleavages of wealth, privilege, and status correspond to the distribution of power and authority. This group comprises the Zulu, the Ngwato, the Bemba, the Banyankole, and the Kede. The other group consists of those societies which lack centralized authority, administrative machinery, and judicial institutions-in short which lack government-and in which there are no sharp divisions of rank, status, or wealth. This group comprises the Logoli, the Tallensi, and the Nuer.*" Other African scholars make a trichotomous distinction between stateless societies, large chiefdoms, and centralized states.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

levels of population density (in (7)) or both (in (8)). The magnitude of political centralization in column (8) in Table $3B$ suggests that luminosity is 34 percent ($exp(0.295)-1 = 0.343$) higher in ethnic homelands where politically centralized societies reside (e.g. Yoruba in Nigeria), as compared to stateless societies or small chiefdoms (e.g. the Sokoto or the Tiv in Nigeria).

Flexibly Estimating the Role of Jurisdictional Hierarchy In columns (9)-(12) we flexibly estimate the relationship between pre-colonial political institutional structures and contemporary development. We define three dummy variables that take on the value one for petty chiefdoms, paramount chiefdoms, and pre-colonial states, respectively; the comparison group being stateless societies.[13] The difference in regional development between stateless societies and small chiefdoms is statistically indistinguishable from zero. This result is in accord with the African historiography that usually does not distinguish between these organizational structures (see also Gennaioli and Rainer (2006, 2007)). Sizable differences in regional development emerge for large paramount chiefdoms and particularly for groups that were part of pre-colonial states. This finding is consistent with Diamond (1997), Bockstette, Chanda, and Putterman (2002), and Acemoglu and Robinson (2012), who argue that centralization and statehood experience of preindustrial societies are the traits most conducive to development.

Nigeria offers an illustration of these results. Average (median) luminosity in the homelands of the five ethnic groups that were part of states in pre-colonial Africa, namely the Yoruba, the Fon, the Ife, the Igala, and the Edo is 1 (0.72). Mean (median) luminosity in the homelands of ethnic groups organized solely at the local level or in petty chiefdoms is 0.88 (0.075). Likewise, in the Democratic Republic of Congo, average luminosity in the homeland of stateless ethnicities and petty chiefdoms is 0.037; luminosity in paramount chiefdoms is only slightly higher, 0.042; yet mean luminosity across homelands of ethnicities belonging to pre-colonial centralized states is three times larger, 0.12.

3.4 Institutions or Other Ethnic Traits?

One concern with the previous estimates is that some other ethnicity feature related to the economy, culture, or societal structure, is driving the positive correlation between luminosity and pre-colonial institutions. To address this issue we examined whether some other ethnic trait, in lieu of political centralization, correlates with contemporary development. In Table 4 we report within-country specifications associating log light density with around twenty different variables from Murdock's Ethnographic Atlas (see the Data Appendix for detailed variable

[13] Since we have just two ethnic groups where the jurisdictional hierarchy index equals four, we assign these ethnicities into the groups where the jurisdictional hierarchy index equals 3.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

definitions).[14] These measures reflect the type of economic activity (dependence on gathering, hunting, fishing, animal husbandry, milking of domesticated animals, and agriculture), societal arrangements (polygyny, presence of clans at the village level, slavery), early development (size and complexity of pre-colonial settlements), and proxies of local institutional arrangements (an indicator for the presence of inheritance rule for property, elections for local headman, class stratification and jurisdictional hierarchy at the village level).

In Specification *A* we regress regional light density on the ethnic-level variables, simply conditioning on country fixed effects and on population density (the results are similar if we omit population density). Most of the additional variables are statistically insignificant. An indicator for societies where fishing contributes more than 5% in the pre-colonial subsistence economy enters with a positive coefficient as economic development is higher in regions close to the coast and other streams and potentially because of blooming in luminosity. An agricultural intensity index ranging from 0 to 9, where higher values indicate higher dependence, is negative and significant, but the correlation between pre-colonial agricultural intensity and regional development is not robust to an alternative index of agricultural dependence.

The results in columns (1)-(2) show that class stratification, a societal trait that has been linked to property rights protection and the emergence of centralized states with a bureaucratic structure, correlates significantly with luminosity.[15] Regional development is higher across regions populated by stratified, as compared to egalitarian, societies. The positive association between stratification and regional development, though surprising at first glance, is in line with recent works in Southern America (e.g. Acemoglu, Bautista, Querubin, and Robinson (2008), Dell (2010)). A potential explanation is that in weakly institutionalized societies inequality may lead to some form of legal institutions, property rights, and policing, as the elite has the incentive to establish constraints (Diamond (1997); Herbst (2000)).

In Specification *B* we add the jurisdictional hierarchy beyond the local community index to test whether it correlates with regional development conditional on the other ethnic traits. In all specifications the jurisdictional hierarchy index enters with a positive and stable coefficient (around 0.20), similar in magnitude to the (more efficient) estimate in Table 3*A* - column 3. The coefficient is always significant at standard confidence levels (usually at the 99% level). Clearly the positive correlation between pre-colonial political institutions and contemporary development may still be driven by some other unobserved or hard-to-account for factor, related for example to genetics or cultural similarities with some local frontier economy (see for example Spolaore and Wacziarg (2009) and Ashraf and Galor (2012)). However, the results in Table 4

[14] We are grateful to an anonymous referee for proposing this test.

[15] In line with these arguments in our sample the correlation of class stratification and the jurisdictional hierarchy index is 0.63.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

reassure that we are not capturing the effect of cultural traits, the type of economic activity, or early development, at least as reflected in Murdock's statistics.

3.5 Further Sensitivity Checks

In the Supplementary Appendix we further explore the sensitivity of our results: (1) dropping observations where luminosity exceeds the 99th percentile; (2) excluding capitals; (3) dropping each time a different part of the continent; (4) using log population density as an alternative proxy for development. Moreover, using data from the Afrobarometer Surveys on living conditions and schooling, we associate pre-colonial institutions with these alternative proxies of regional development. Across all specifications, we find a significantly positive correlation between a group's current economic performance and pre-colonial political centralization.

4 Pixel-Level Analysis

We now proceed to the pixel-level analysis. In this section the unit of analysis is a pixel of $0.125*$ 0.125 decimal degrees. As a result we now have multiple observations within each ethnic area in each country. Since there are several unpopulated pixels (in the Sahara and the rainforests) we exclude pixels with zero population (including unpopulated pixels if anything strengthens the results). Figure 4 illustrates the new unit of analysis showing pixel-level luminosity within two Bantu groups in Northern Zambia, the Lala and the Lamba.

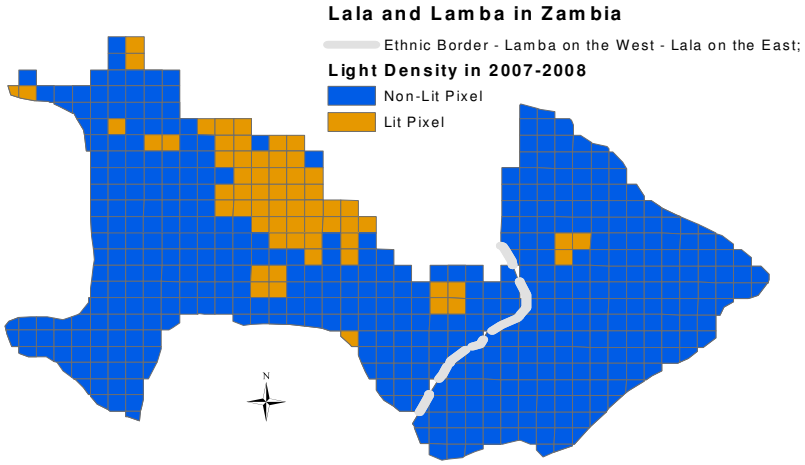


Figure 4: Example of the Pixel-Level Analysis

Moving to the pixel level offers some advantages. First, we can condition on geography, natural resources, and the disease environment at an even finer level. Second, since the dependent variable is an indicator for lit pixels, the non-linear nature of luminosity is no longer a concern. Third, we account for the possibility that average luminosity at the ethnic homeland

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

also reflects inequality; this may be the case when average light density at the ethnic homeland is driven by a few extremely lit pixels.

4.1 Benchmark Pixel-Level Estimates

Our specification for the pixel-level analysis reads:

$$y_{p,i,c} = a_0 + a_c + \gamma IQL_i + \lambda PD_{p,i,c} + Z'_{p,i,c}\Psi + X'_{i,c}\Phi + \zeta_{p,i,c}.$$

The dependent variable, $y_{p,i,c}$, reflects economic activity in pixel p that belongs to the historical homeland of ethnicity i in country c. $PD_{p,i,c}$ denotes log population density, while vector $Z'_{p,i,c}$ includes other controls at the pixel level; $X'_{i,c}$ is the set of conditioning variables at the ethnic-country level.

Table 5 - Panel A reports the results. In columns (1)-(5) we report linear probability models where the dependent variable equals one if the pixel is lit and zero otherwise. The coefficient on the jurisdictional hierarchy index in the unconditional specification in column (1) is positive and highly significant. The estimate retains significance when we add a vector of country constants (in (2)). In column (3) we control for log pixel population density. As in our analysis at the ethnic homeland level, the coefficient on the pre-colonial ethnic institutions index declines, though it remains significant. In column (4) we augment the specification with a rich set of pixel-specific controls. Namely, we control for land suitability for agriculture, elevation, malaria stability, surface area, distance from the centroid of each pixel to the sea coast, the capital city, and the national border and we add indicators capturing the presence of diamond mines, petroleum fields, and water bodies.[16] In spite of the inclusion of this rich set of controls, the jurisdictional hierarchy beyond the local community continues to enter with a positive and highly significant (at the 99% confidence level) coefficient. In column (5) we condition on the location and geographic controls at the country-ethnic level. The coefficient on the jurisdictional hierarchy index remains intact. The estimate (0.031) in column (4) implies that compared to stateless ethnicities, pixels in the homelands of ethnic groups that were part of paramount chiefdoms are on average 6% more likely to be lit. Similarly, the likelihood that a pixel is lit is 9 percentage points higher if it falls in the homeland of groups that had complex centralized institutions at the time of colonization. These magnitudes are not negligible, since only 17% of populated pixels are lit across Africa.[17]

In Panel B we estimate in a flexible manner the relationship between pre-colonial political

[16] Note that not all pixels have the same surface area since pixels by the coast, lakes, and ethnic boundaries are smaller.

[17] The results are similar using the Gennaioli and Rainer (2006, 2007) binary index of political centralization (see Appendix Table 6).

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

organizational forms and contemporary development. The estimates show that differences in development become economically and statistically significant when one compares paramount chiefdoms to stateless societies or groups organized as petty chiefdoms. Contemporary development is even higher in areas populated by societies that were part of pre-colonial states. The most conservative estimates imply that the likelihood that a pixel is lit is approximately 8 percentage points higher when one moves from the homeland of stateless ethnicities to regions with ethnic groups that pre-colonially were part of a centralized state. Examples from Botswana illustrate the point estimates. The Naron and the Kung are two stateless societies whereas the (Ba)Ngwato (a traditional Sotho-Tswana tribe) and the Ndebele (which originate to the Zulus, the dominant ethnic group of one of the largest pre-colonial states in Southern Africa) are centralized groups. On average 27.8% of the homeland of the Ndebele and the Ngwato is lit, while only 5.4% of the homeland of the Naron and the Kung is lit.

In columns (6)-(10) of Table 5 we report otherwise identical to columns (1)-(5) LS specifications using as the dependent variable the log of luminosity adding a small number (0.01). The coefficient on the jurisdictional hierarchy index is more than two standard errors larger than zero across all perturbations; this shows that our results at the ethnic homeland level were not driven by the transformation of luminosity.

4.2 Contiguous Ethnic Homeland Analysis

Approach and Empirical Specification In spite of employing a rich conditioning set, one may still be worried that some unobservable local geographic feature is driving the results. To mitigate such concerns we focus on contiguous ethnicities with a different degree of pre-colonial political centralization and exploit within-country, within-adjacent ethnicities variation in luminosity and ethnic institutions. In some sense this approach extends the pioneering case study of Douglas (1962), who attributed the large differences in well-being between the neighboring Bushong and the Lele in the Democratic Republic of Congo to their local institutions and the degree of political centralization in particular.[18]

We first identified contiguous ethnic homelands, where groups differ in the degree of political centralization, using the Gennaioli and Rainer (2007) binary classification. There are 252 unique adjacent ethnic pairs comprising a centralized and a non-centralized ethnicity. Figure 4 illustrates this using the Lala and the Lamba. The Lala were organized as a petty chiefdom at the time of colonization; as such the binary political centralization index equals zero. The Lamba are classified as a paramount chiefdom and therefore as politically centralized. When a group is adjacent to more than one ethnicities with different pre-colonial centralization

[18] We are thankful to Jim Robinson for providing us with this reference.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

in the same country, we include all pairs.[19] Then we examine whether there are systematic differences in development within contiguous ethnic homelands in the same country running specifications of the following form:

$$y_{p,i(j),c} = a_{i(j),c} + \delta IQL_i + \lambda PD_{p,i,c} + Z'_{p,i,c}\Psi + \zeta_{p,i(j),c}.$$

The dependent variable takes on the value of one if pixel p is lit and zero otherwise. Every pixel p falls into the historical homeland of ethnicity i in country c that is adjacent to the homeland of ethnicity j in the same country c (where ethnicities i and j differ in their degree of pre-colonial political centralization). Since we now include country-specific, ethnicity-pair fixed effects, $a_{i(j),c}$, the coefficient on the jurisdictional hierarchy beyond the local community index, δ, captures whether differences in pre-colonial ethnic institutions translate into differences in light density across pixels within pairs of contiguous ethnicities in the same country.

Validation Before we present the results, we examine whether there is a systematic correlation between pre-colonial institutions and various characteristics within adjacent ethnic pairs in the same country. To do so we run ethnic-pair-country fixed effects specifications (with $a_{i(j),c}$) associating the jurisdictional hierarchy index with natural resources (presence of diamond mines or petroleum fields), location (distance to capital, to the sea and the national border), and geography (elevation, presence of water bodies, soil quality for agriculture and the malaria stability index). These regressions, reported in Table 6, yield statistically and economically insignificant estimates suggesting that by focusing on neighboring ethnic areas we neutralize the role of local (observable) geographic and location factors.

Results Table 7 reports the results of the contiguous-ethnic-homeland analysis. The estimate in (1) shows that within country, within pairs of contiguous ethnic homelands luminosity is significantly higher in the historical homeland of ethnicities with more complex political institutions. In column (2) we condition on pixel population density. The coefficient on the jurisdictional hierarchy beyond the local community index falls, though it becomes more precisely estimated. In column (3) we control for the rich set of pixel-level geographic variables. While some of these variables enter with significant estimates, given their minimal correlation with the jurisdictional hierarchy index (shown in Table 6), this has a negligible effect on the estimate. The coefficient in column (3) implies that the likelihood that a pixel is lit is approximately 2.5 percentage points higher if one moves from the homeland of a stateless group

[19] For example, the Dagomba in Ghana, a centralized group (the jurisdictional hierarchy index equals 3) is adjacent to two non-centralized groups in Ghana, the Basari and the Konkomba. In such cases we include both pairs. The median (average) distance between the centroids of neighboring ethnicities is $179km$ ($215km$).

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

to the neighboring homeland in the same country of an ethnic group that was organized as a paramount chiefdom. In columns (4)-(6) we restrict our analysis to pairs of contiguous ethnic homelands with large differences (two levels or greater) in the jurisdictional hierarchy index; this is helpful not only because we now focus on sharper discontinuities, but also because we can account (to some degree) for measurement error in Murdock's classification of pre-colonial political organization. The estimate on the pre-colonial ethnic institutions index retains its statistical and economic significance. In columns (7)-(9) we require that one of the two adjacent ethnic groups was part of a pre-colonial state. Thus, in these models in each pair of adjacent ethnicities we compare a group which had been either stateless or part of a petty chiefdom (the Gennaioli and Rainer (2007) index equals zero) to an ethnicity that was organized as a state at the time of colonization. There is a strong positive correlation between differences in luminosity and differences in the degree of pre-colonial political institutions. The estimates suggest that the probability that a pixel is lit is $5.5\% - 7.5\%$ higher when one moves from the homeland of stateless societies to the areas of groups that formed large states before the colonial era.

A couple of examples are useful. In Uganda 2% of the pixels falling in the homeland of the Acholi, a non-centralized group (jurisdictional hierarchy index equals 1) are lit, while 4.2% of the pixels are lit in the adjacent homeland of the Nyoro, an ethnic group that was part the large Banyoro kingdom (jurisdictional hierarchy index equals 3). Similarly, 21.4% of the pixels are lit in the homeland of the Ganda, the central ethnic group of the powerful kingdom of Buganda that had a highly centralized bureaucracy under the kabaka/king, compared to only 6.7% lit pixels in the neighboring territory of the stateless Lango.

Further Evidence To further assuage concerns that some local unobserved geographic feature is driving the results, we narrowed our analysis to pixels close to the ethnic boundary. This approach is similar in spirit to regression discontinuity type analyses that are becoming increasingly popular in institutional economics.[20] In our context, implementing a standard regression discontinuity design across ethnic boundaries like the ones that are usually performed across the national border is not advisable for several reasons. First, while national borders are accurately delineated, drawing error in Murdock's map on the exact location of ethnic boundaries is likely to be non-trivial. Second, since Murdock's map, originally printed in the end of his book on African ethnicities, is available at a small scale, its digitization magnifies any noise inherent to the initial border drawing. Third, Murdock assigns each part of Africa to a single dominant group, while (some) ethnicities (may) overlap; and naturally population mixing is higher closer to ethnic boundaries. Fourth, due to bleeding in the luminosity data

[20] See, for example, Dell (2010), Bubb (2012), and Michalopoulos and Papaioannou (2012), among others.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

(occurring from the diffusion of light) and since electricity grids are crossing adjacent regions within the same country, we may not be able to detect significant differences in luminosity in areas very close to ethnic borders.

In spite of these limitations we took the (heroic) step to estimate the role of local institutions close to the ethnic boundaries. In an effort to counterbalance the potential merits of focusing very close to the ethnic border and accounting for the aforementioned problems, we perform estimation in areas close to the ethnic boundaries, but excluding pixels that fall within 25 kilometers or within 50 kilometers from each side of the border. Essentially, this boils down to assuming that the ethnic border is "thick" (by either $50km$ or $100km$). We perform the analysis within adjacent ethnic homelands with different pre-colonial political institutions in the same country. In case of ethnic homelands having multiple neighbors with different pre-colonial centralization we chose the largest in size bordering group.

Table $8A$ reports LS regression estimates using three different bandwidths ($100km$, $150km$, and $200km$) from the original ethnic border. In the most restrictive specification in column (1) of Panel A, when we limit our attention to areas within 100 kilometers from the ethnic border (while excluding pixels within a 25 kilometers range), the coefficient on the index of jurisdictional hierarchy beyond the local community is positive (0.019) and statistically significant at the 90% level. When we increase the bandwidth to 150 kilometers in column (2) the coefficient increases somewhat (0.023) and retains its statistical significance; further increasing the bandwidth to 200 kilometers (or more) has no impact on the coefficient, while due to the increase in the sample, the standard errors become tighter. Turning now to Panel B, when we exclude pixels 50 kilometers from the ethnic border, the coefficient is 0.023 when we use the narrow bandwidth of 100 kilometers and 0.028 when we increase the bandwidth to 150 or 200 kilometers. In columns (4)-(6) we focus on pairs of ethnicities in the same country with sharp discontinuities in the strength of pre-colonial ethnic institutions. The estimates show that regional development is significantly higher in the homeland of societies with advanced pre-colonial institutions. Finally, in columns (7)-(9) we perform the analysis requiring that the centralized ethnic group has been part of a state before colonization. The estimates are somewhat larger, while the standard errors fall. Across all specifications the coefficient on the jurisdictional hierarchy index is in the range of $0.020 - 0.035$, quite similar to the estimates in Tables 5 and 7. This reassures that our benchmark estimates were not driven by unobserved local features.

In Table $8B$ we estimate locally linear regressions including in the set of controls an RD-type fourth-order polynomial in distance to the "thick" ethnic border, allowing the coefficients on the distance terms to be different for the relatively high and the relatively low institutional

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

quality homelands, respectively. Compared to the analogous estimates in Table 8*A* this allows us to estimate the role of pre-colonial ethnic institutions exactly at the ethnic border. Across all specifications the coefficient on the jurisdictional hierarchy index is positive and if anything somewhat higher than the corresponding specifications in Table 8*A* where we did not include the RD-type polynomial in distance to the ethnic boundary. While standard errors are somewhat larger, the estimates are statistically significant at the standard confidence levels in most specifications.

Figures 5*a* and 5*b* illustrate graphically the relationship between pixel luminosity and distance to the ethnic border for adjacent groups with large differences (two levels or greater) in the jurisdictional hierarchy index. Figure 5*a*, which includes the boundary pixels, suggests that while light density is overall higher in the homelands of centralized ethnic groups, these differences become miniscule and are statistically indistinguishable from zero for pixels exactly at the ethnic border. Yet, as Figure 5*b* shows, when we just exclude 25 kilometers from each side of the ethnic border, then differences in pixel-level light density become both economically and statistically significant.

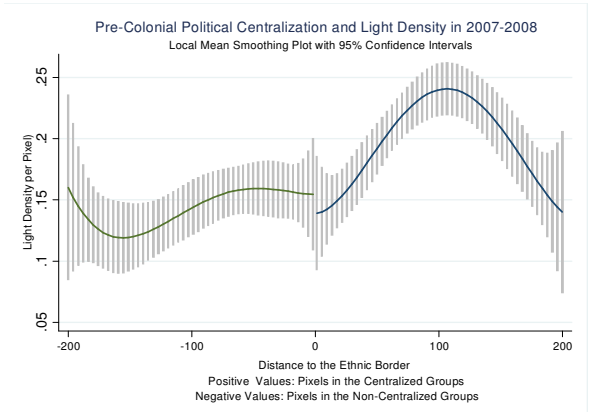


Figure 5*a* - Border Thickness: 0 km

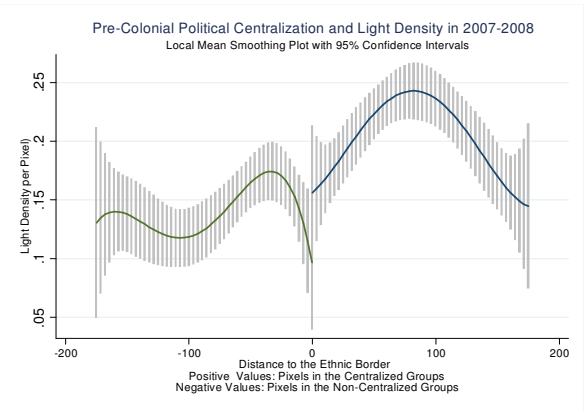


Figure 5*b* - Border Thickness: 25 km

5 Conclusion

In this study we combine anthropological data on the spatial distribution and local institutions of African ethnicities at the time of colonization with satellite images on light density at night to assess the role of deeply-rooted ethnic institutions in shaping contemporary comparative African development. Exploiting within-country variation, we show that regional development is significantly higher in the historical homelands of ethnicities with centralized, hierarchical, pre-colonial political institutions.

Since we do not have random assignment on ethnic institutions, this correlation does

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

not necessarily imply causation. Hard-to-account-for factors related to geography, culture, or early development may confound these results. Yet, the uncovered pattern is robust to a host of alternative explanations. First, we show that the strong correlation between pre-colonial institutional complexity and current development is not driven by observable differences in geographic, ecological, and natural resource endowments both at the ethnic homeland and at the pixel level. Second, the uncovered link between historical political centralization and contemporary development is not mediated by observable ethnic differences in culture, occupational specialization, and the structure of economic activity before colonization. Third, the positive association between pre-colonial ethnic political institutions and luminosity is present within pairs of adjacent ethnic homelands in the same country where groups with different pre-colonial institutions reside. Our analysis, therefore, provides large-scale formal econometric evidence in support of the African historiography that dates back to Fortes and Evans-Pritchard (1940) emphasizing the importance of ethnic institutions in shaping contemporary economic performance.

The uncovered empirical regularities call for future research. First, our results imply that the literature on the political economy of African development should move beyond country-level features and examine the role of ethnic-specific attributes. Second, future research should shed light on the mechanisms via which ethnic institutional and cultural traits shape economic performance. Third, empirical and theoretical work is needed to understand how local ethnicity-specific institutions and cultural norms emerge. Finally, our approach to combine high resolution proxies of development (such as satellite light density at night) with anthropological data on culture and institutions provides a platform for subsequent research, allowing, for example, one to investigate the interplay between ethnic traits and national policies.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

References

Acemoglu, D., M. A. Bautista, P. Querubin, and J. A. Robinson (2008): "Economic and Political Inequality in Development: The Case of Cundinamarca, Colombia," in *Institutions and Economic Performance*, ed. by E. Helpman. Harvard University Press.

Acemoglu, D., and S. Johnson (2005): "Unbundling Institutions," *Journal of Political Economy*, 113(5), 949–995.

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

——— (2002): "Reversal Of Fortune: Geography And Institutions In The Making Of The Modern World Income Distribution," *Quarterly Journal of Economics*, 107(4), 1231–1294.

——— (2005): "Institutions as a Fundamental Cause of Long-Run Growth," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. N. Durlauf, pp. 109–139. Elsevier North-Holland, Amsterdam, The Netherlands.

Acemoglu, D., T. Reed, and J. A. Robinson (2012): "Chiefs," mimeo, Harvard University and MIT.

Acemoglu, D., and J. A. Robinson (2012): *Why Nations Fail? The Origins of Power, Prosperity, and Poverty*. Crown Publishers, New York, NY.

Alesina, A., W. Easterly, and J. Matuszeski (2011): "Artificial States," *Journal of the European Economic Association*, 9(2), 246–277.

Angrist, J., and J.-S. Pischke (2008): *Mostly Harmless Econometrics*. Princeton University Press, Princeton, NJ.

Arbesu, H. A. L. (2011): "Colonialism and Economic Development: Evidence from a Natural Experiment in Colonial Nigeria," mimeo MIT.

Ashraf, Q., and O. Galor (2011): "Dynamics and Stagnation in the Malthusian Epoch: Theory and Evidence," American Economic Review, forthcoming.

——— (2012): "Human Genetic Diversity and Comparative Economic Development," American Economic Review, forthcoming.

Baldwin, K. (2010): "Why Vote with the Chief? Political Connections and Public Goods Provision in Zambia," Working Paper, Department of Political Science, University of Florida.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

——— (2011): "When Politicians Cede Control of Resources: Land, Chiefs, and Coalition Building in Africa," Working Paper, Department of Political Science, University of Florida.

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

Bates, R. H. (1983): "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa," in *State versus Ethnic Claims: African Policy Dilemmas*, ed. by D. Rothchild, and V. A. Olunsorola. Westview Press, Boulder, CO.

Berger, D. (2009): "Taxes, Institutions and Governance: Evidence from Colonial Nigeria," mimeo, New York University.

Besley, T., and T. Persson (2011): *Pillars of Prosperity. The Political Economics of Development Clusters*. Princeton University Press, Princeton, NJ.

Besley, T., and M. Reynal-Querol (2012): "The Legacy of Historical Conflicts. Evidence from Africa," mimeo, London School of Economics and UPF.

Bockstette, V., A. Chanda, and L. Putterman (2002): "States and Markets: The Advantage of an Early Start," *Journal of Economic Growth*, 7(4), 347–369.

Boone, C. (2003): *Political Topographies of the African State: Territorial Authority and Institutional Choice*. Cambridge University Press.

Bubb, R. (2012): "States, Law, and Property Rights in West Africa," Working Paper, New York University School of Law.

Cameron, C. A., J. Gelbach, and D. Miller (2011): "Robust Inference with Multi-Way Clustering," *Journal of Business and Economic Statistics*, 28(2), 238–249.

Chen, X., and W. D. Nordhaus (2011): "Using Luminosity Data as a Proxy for Economic Statistics," *Proceeedings of the National Academy of Sciences*, 108(21), 8589–8594.

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

Dell, M. (2010): "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 78(6), 839–857.

Diamond, J. (1997): *Guns, Germs, and Steel: The Fates of Human Societies*. W. W. Norton & Co., New York, NY.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Doll, C. N., J.-P. Muller, and J. G. Morley (2006): "Mapping Regional Economic Activity from Night-Time Light Satellite Imagery," *Ecological Economics*, 57, 75–92.

Douglas, M. (1962): "Lele Economy Compared with the Bushong: A Study of Economic Backwardness," in *Markets in Africa*, ed. by P. Bohannan, and G. Dalton, pp. 211–233. Northwestern University Press, Evanston, ILL.

Elvidge, C., K. Baugh, E. Kihn, H. Kroehl, and E. Davis (1997): "Mapping of City Lights Using DMSP Operational Linescan System data," *Photogrammetric Engineering and Remote Sensing*, 63, 727–734.

Englebert, P. (2009): *Africa, Unity, Sovereignty and Sorrow*. Lynne Rienner Publishers, Inc., Boulder, Colorado.

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

Ethnologue (2005): Languages of the World, SIL International; Fifteenth edition.

Fenske, J. (2009): "Does Land Abundance Explain African Institutions," mimeo Yale University.

Fortes, M., and E. Evans-Pritchard (1940): *African Political Systems*. Oxford University Press.

Gallup, J. L., and J. D. Sachs (2001): "The Economic Burden of Malaria," *The American Journal of Tropical Medicine and Hygiene*, 64(1), 85–96.

Gennaioli, N., and I. Rainer (2006): "Precolonial Centralization and Institutional Quality in Africa," in *Institutions and Norms in Economic Development*, ed. by M. Gradstein, and K. Konrad. MIT Press.

——— (2007): "The Modern Impact of Precolonial Centralization in Africa," *Journal of Economic Growth*, 12(3), 185–234.

Glennerster, R., E. Miguel, and A. Rothenberg (2010): "Collective Action in Diverse Sierra Leone Communities," Working Paper, NBER, No. 16196.

Goldstein, M., and C. Udry (2008): "The Profits of Power: Land Rights and Agricultural Investment in Ghana," *Journal of Political Economy*, 116(6), 981–1022.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Goody, J. (1971): *Technology, Tradition, and the State in Africa*. International African Institute by Oxford University Press.

Gray, P. J. (1999): "A Corrected Ethnographic Atlas," *Wolrd Cultures*, 10, 24–85.

Henderson, V. J., A. Storeygard, and D. N. Weil (2011): "Measuring Economic Growth from Outer Space," *American Economic Review, forthcoming*.

Herbst, J. (2000): *States and Power in Africa*. Princeton University Press, Princeton, NJ.

Huillery, E. (2009): "History Matters: The Long Term Impact of Colonial Public Investments in French West Africa," *American Economic Journal - Applied Economics*, 1(2), 176–215.

Iyer, L. (2010): "Direct versus Indirect Colonial Rule in India: Long-term Consequences," *Review of Economics and Statistics*, 92(4), 693–713.

Logan, C. (2011): "The Roots of Resilience: Exploring Popular Support for African Traditional Authorities," Afrobarometer Working Paper No. 128.

Mamdani, M. (1996): *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton University Press, Princeton, NJ.

Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4), 1508–1539.

Michalopoulos, S., and E. Papaioannou (2011a): "Divide and Rule or the Rule of the Divided? Evidence from Africa," NBER Working Paper 17184.

——— (2011b): "The Long-Run Effects of the Scramble for Africa," CEPR Discussion Paper No. 8676.

——— (2012): "National Institutions and African Development: Evidence from Partitioned Ethnicities," mimeo, Brown University and Dartmouth College.

Middleton, J., and D. Tait (1958): *Tribes without Rulers*. Routlege and Kegan Paul, London.

Min, B. (2008): "Democracy and Light: Electoral Accountability and the Provision of Public Goods," Working Paper, mimeo UCLA.

Murdock, G. P. (1959): *Africa: Its Peoples and Their Culture History*. McGraw-Hill Book Company, New York.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

——— (1967): *Ethnographic Atlas*. University of Pittsburgh Press, Pittsburgh, PA.

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

Nunn, N., and D. Puga (2012): "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics*, 94(1), 20–36.

Nunn, N., and L. Wantchekon (2011): "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 101(7), 3221–3252.

Pinkovskiy, M. L. (2011): "Economic Discontinuities across Borders: Evidence from Satellite Data on Lights at Night," mimeo, MIT, Department of Economics.

Porta, R. L., F. L. de Silanes, A. Shleifer, and R. Vishny (1997): "Legal Determinants of External Finance," *Journal of Finance*, 52(3), 1131–1150.

——— (1998): "Law and Finance," *Journal of Political Economy*, 106(6), 1113–1155.

Robinson, J. A. (2002): "States and Power in Africa: Comparative Lessons in Authority and Control by Jeffrey I. Herbst: A Review Essay," *Journal of Economic Literature*, 40, 510–519.

Ross, M. L. (2006): "A Closer Look on Oil, Diamonds, and Civil War," *Anuual Review of Political Science*, 9(1), 265–300.

Schapera, I. (1967): *Government and Politics in Tribal Societies*. Humanities Press, London: Watts.

——— (1970): "Tribal Innovators: Tswana Chiefs and Social Change, 1795-1940," *London School of Economics Monographs on Social Anthropology*, (43).

Spolaore, E., and R. Wacziarg (2009): "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2), 469–529.

Stevenson, R. F. (1968): *Population and Political Systems in Tropical Africa*. Columbia University Press.

Tait, D. (1958): "The Territorial Pattern and Lineage System of Konkomba," in *Tribes without Rulers*, ed. by J. Middleton, and D. Tait, pp. 167–202. Routledge and Kegan Paul, London.

Winter, E. (1958): "The Aboriginal Political Structure of Bwamba," in *Tribes without Rulers*, ed. by J. Middleton, and D. Tait, pp. 136–166. Routledge and Kegan Paul, London.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 1: Summary Statistics

variable	Obs.	mean	st. dev.	p25	median	p75	min	max
Panel A: All Observations								
Light Density	683	0.368	1.528	0.000	0.022	0.150	0.000	25.140
Ln (0.01 + Light Density)	683	-2.946	1.701	-4.575	-3.429	-1.835	-4.605	3.225
Pixel-Level Light Density	66570	0.560	3.422	0.000	0.000	0.000	0.000	62.978
Lit Pixel	66570	0.167	0.373	0.000	0.000	0.000	0.000	1.000
Panel B: Stateless Ethnicities								
Light Density	183	0.248	1.878	0.000	0.017	0.080	0.000	25.140
Ln (0.01 + Light Density)	183	-3.273	1.427	-4.605	-3.621	-2.408	-4.605	3.225
Pixel-Level Light Density	13174	0.172	1.556	0.000	0.000	0.000	0.000	55.634
Lit Pixel	13174	0.100	0.301	0.000	0.000	0.000	0.000	1.000
Panel C: Petty Chiefdoms								
Light Density	276	0.269	1.155	0.000	0.013	0.084	0.000	13.086
Ln (0.01 + Light Density)	276	-3.238	1.584	-4.605	-3.753	-2.370	-4.605	2.572
Pixel-Level Light Density	20259	0.283	2.084	0.000	0.000	0.000	0.000	60.022
Lit Pixel	20259	0.129	0.335	0.000	0.000	0.000	0.000	1.000
Panel D: Paramount Chiefdoms								
Light Density	175	0.311	0.940	0.001	0.037	0.191	0.000	9.976
Ln (0.01 + Light Density)	175	-2.788	1.711	-4.545	-3.058	-1.604	-4.605	2.301
Pixel-Level Light Density	20972	0.388	2.201	0.000	0.000	0.000	0.000	58.546
Lit Pixel	20972	0.169	0.375	0.000	0.000	0.000	0.000	1.000
Panel E: Pre-Colonial States								
Light Density	80	0.993	2.246	0.007	0.082	0.803	0.000	14.142
Ln (0.01 + Light Density)	80	-2.022	2.183	-4.106	-2.391	-0.207	-4.605	2.650
Pixel-Level Light Density	12165	1.739	6.644	0.000	0.000	0.160	0.000	62.978
Lit Pixel	12165	0.302	0.459	0.000	0.000	1.000	0.000	1.000

The table reports descriptive statistics for the luminosity data that we use to proxy economic development at the country-ethnic homeland level and at the pixel level. Panel A gives summary statistics for the full sample. Panel B reports summary statistics for ethnicities that lacked any form of political organization beyond the local level at the time of colonization. Panel C reports summary statistics for ethnicities organized in petty chiefdoms. Panel D reports summary statistics for ethnicities organized in large paramount chiefdoms. Panel E reports summary statistics for ethnicities organized in large centralized states. The classification follows Murdock (1967). The Data Appendix gives detailed variable definitions and data sources.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 2: Pre-colonial Ethnic Institutions and Regional Development
Cross-Sectional Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Jurisdictional Hierarchy	0.4106***	0.3483**	0.3213***	0.1852***	0.1599***	0.1966***
Double-clustered s.e.	(0.1246)	(0.1397)	(0.1026)	(0.0676)	(0.0605)	(0.0539)
Conley's s.e.	[0.1294]	[0.1288]	[0.1014]	[0.0646]	[0.0599]	[0.0545]
Rule of Law (in 2007)					0.4809**	
Double-clustered s.e.					(0.2213)	
Conley's s.e.					[0.1747]	
Log GDP p.c. (in 2007)						0.5522***
Double-clustered s.e.						(0.1232)
Conley's s.e.						[0.1021]
Adjusted R-squared	0.056	0.246	0.361	0.47	0.488	0.536
Population Density	No	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	Yes	Yes	Yes
Observations	683	683	683	683	680	680

Table 2 reports OLS estimates associating regional development with pre-colonial ethnic institutions, as reflected in Murdock's (1967) index of jurisdictional hierarchy beyond the local community. The dependent variable is log (0.01 + light density at night from satellite) at the ethnicity-country level. In columns (5) we control for national institutions augmenting the specification with the rule of law index (in 2007). In column (6) we control for the overall level of economic development augmenting the specification with the log of per capita GDP (in 2007). In columns (2)-(6) we control for log (0.01 + population density). In columns (3)-(6) we control for location augmenting the specification with distance of the centroid of each ethnicity-country area from the respective capital city, distance from the closest sea coast, and distance from the national border. The set of geographic controls in columns (4)-(6) includes log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator.

The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and ethno-linguistic family dimensions. We also report in brackets Conley's (1999) standard errors that account for 2-dimensional spatial auto-correlation. ***, **, and * indicate statistical significance with the most conservative standard errors at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 3A: Pre-colonial Ethnic Institutions and Regional Development within African Countries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Jurisdictional Hierarchy	0.3260***	0.2794***	0.2105***	0.1766***								
	(0.0852)	(0.0853)	(0.0554)	(0.0502)								
Binary Political Centralization					0.5264***	0.5049***	0.3413***	0.3086***				
					(0.1492)	(0.1575)	(0.0898)	(0.0974)				
Petty Chiefdoms									0.1538	0.1442	0.1815	0.1361
									(0.2108)	(0.1739)	(0.1542)	(0.1218)
Paramount Chiefdoms									0.4258*	0.4914*	0.3700**	0.3384**
									(0.2432)	(0.2541)	(0.1628)	(0.1612)
Pre-Colonial States									1.1443***	0.8637***	0.6809***	0.5410***
									(0.2762)	(0.2445)	(0.1641)	(0.1486)
Adjusted R-squared	0.409	0.540	0.400	0.537	0.597	0.661	0.593	0.659	0.413	0.541	0.597	0.661
Observations	682	682	682	682	682	682	682	682	682	682	682	682
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 3B: Pre-colonial Ethnic Institutions and Regional Development within African Countries
Intensive Margin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Jurisdictional Hierarchy	0.3279***	0.3349***	0.1651**	0.1493**								
	(0.1240)	(0.1120)	(0.0705)	(0.0729)								
Binary Political Centralization					0.4819**	0.6594***	0.2649**	0.2949**				
					(0.2386)	(0.2090)	(0.1234)	(0.1394)				
Petty Chiefdoms									0.1065	0.1048	0.0987	0.0135
									(0.2796)	(0.2363)	(0.1791)	(0.1733)
Paramount Chiefdoms									0.2816	0.6253*	0.2255	0.2374
									(0.3691)	(0.3373)	(0.2261)	(0.2392)
Pre-Colonial States									1.2393***	0.9617***	0.5972***	0.4660**
									(0.3416)	(0.3224)	(0.2226)	(0.2208)
Adjusted R-squared	0.424	0.562	0.416	0.562	0.638	0.671	0.636	0.671	0.431	0.564	0.639	0.672
Observations	517	517	517	517	517	517	517	517	517	517	517	517
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

Table 3 reports within-country OLS estimates associating regional development with pre-colonial ethnic institutions. In Panel A the dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. In Panel B the dependent variable is the log (light density at night from satellite) at the ethnicity-country level; as such we exclude areas with zero luminosity. In columns (1)-(4) we measure pre-colonial ethnic institutions using Murdock's (1967) jurisdictional hierarchy beyond the local community index. In columns (5)-(8) we use a binary political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local community variable. Following Gennaioli and Rainer (2007), this index takes on the value of zero for stateless societies and ethnic groups that were part of petty chiefdoms and one otherwise (for ethnicities that were organized as paramount chiefdoms and ethnicities that were part of large states). In columns (9)-(12) we augment the specification with three dummy variables that identify petty chiefdoms, paramount chiefdoms, and large states. The omitted category consists of stateless ethnic groups before colonization. All specifications include a set of country fixed effects (constants not reported).

In even-numbered columns we control for location and geography. The set of control variables includes the distance of the centroid of each ethnicity-country area from the respective capital city, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4: Examining the Role of Other Pre-colonial Ethnic Features

	Specification *A*		Specification *B*		
	Additional Variable	Obs.	Additional Variable	Jurisdictional Hierarchy	Obs.
	(1)	(2)	(3)	(4)	(5)
Gathering	-0.0937 (0.1689)	749	-0.0771 (0.1842)	0.2082*** (0.0552)	682
Hunting	-0.0292 (0.1217)	749	-0.0167 (0.1238)	0.2099*** (0.0563)	682
Fishing	0.2385* (0.1315)	749	0.2359* (0.1269)	0.2087*** (0.0552)	682
Animal Husbandry	0.0549 (0.0407)	749	0.0351 (0.0432)	0.2008*** (0.0618)	682
Milking	0.1782 (0.1395)	702	0.0872 (0.1443)	0.2016*** (0.0581)	680
Agriculture Dependence	-0.1058** (0.0436)	749	-0.1032** (0.0454)	0.2078*** (0.0558)	682
Agriculture Type	0.0237 (0.1015)	703	-0.0131 (0.1022)	0.2092*** (0.0549)	680
Polygyny	0.0744 (0.1197)	735	0.0796 (0.1290)	0.2140*** (0.0562)	677
Polygyny Alternative	-0.019 (0.1588)	749	0.007 (0.1482)	0.2106*** (0.0544)	682
Clan Communities	-0.1294 (0.1479)	617	-0.0079 (0.1404)	0.2158*** (0.0537)	567
Settlement Pattern	-0.0159 (0.0350)	701	-0.0057 (0.0378)	0.2103*** (0.0572)	679
Complex Settlements	0.2353 (0.1597)	701	0.2154 (0.1609)	0.1991*** (0.0554)	679

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 4: Examining the Role of Other Pre-colonial Ethnic Features, cont.

	Specification *A*		Specification *B*		
	Additional Variable	Obs.	Additional Variable	Jurisdictional Hierarchy	Obs.
	(1)	(2)	(3)	(4)	(5)
Hierarchy of Local Community	0.0197 (0.0825)	686	-0.0009 (0.0869)	0.2085*** (0.0566)	680
Patrilineal Descent	-0.1880 (0.1313)	733	-0.2011 (0.1309)	0.1932*** (0.0500)	671
Class Stratification	0.1268** (0.0493)	608	0.0672 (0.0581)	0.1556** (0.0697)	570
Class Stratification Indicator	0.3970** (0.1817)	608	0.2757 (0.1900)	0.1441** (0.0563)	570
Elections	0.3459 (0.2462)	536	0.2764 (0.2583)	0.2217*** (0.0583)	500
Slavery	-0.0236 (0.1559)	661	-0.1192 (0.1583)	0.2016*** (0.0619)	616
Inheritance Rules for Property Rights	-0.1963 (0.2172)	569	-0.1788 (0.2283)	0.2196*** (0.0691)	529

Table 4 reports within-country OLS estimates associating regional development with pre-colonial ethnic features as reflected in Murdock's (1967) Ethnographic Atlas. The dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. All specifications include a set of country fixed effects (constants not reported). In all specifications we control for log (0.01 + population density at the ethnicity-country level). In specification A (in columns (1)-(2)) we regress log (0.01 + light density) on various ethnic traits from Murdock (1967). In specification B (columns (3)-(5)) we regress log (0.01 + light density) on each of Murdock's additional variables and the jurisdictional hierarchy beyond the local community index. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 5: Pre-colonial Ethnic Institutions and Regional Development: Pixel-Level Analysis

	Lit/Unlit Pixels					Ln (0.01 + Luminosity)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Jurisdictional Hierarchy beyond the Local Community Level										
Jurisdictional Hierarchy	0.0673**	0.0447**	0.0280***	0.0308***	0.0265***	0.3619**	0.2362**	0.1528***	0.1757***	0.1559***
Double-clustered s.e.	(0.0314)	(0.0176)	(0.0081)	(0.0074)	(0.0071)	(0.1837)	(0.1035)	(0.0542)	(0.0506)	(0.0483)
Adjusted R-squared	0.034	0.272	0.358	0.375	0.379	0.045	0.320	0.418	0.448	0.456
Panel B: Pre-colonial Institutional Arrangements										
Petty Chiefdoms	0.0285	0.0373	0.0228	0.0161	0.0125	0.1320	0.1520	0.0796	0.0642	0.0531
Double-clustered s.e.	(0.0255)	(0.0339)	(0.0220)	(0.0175)	(0.0141)	(0.1192)	(0.1832)	(0.1271)	(0.0976)	(0.0837)
Paramount Chiefdoms	0.0685**	0.0773	0.0546*	0.0614**	0.0519***	0.3103**	0.3528	0.2389	0.3054**	0.2802***
Double-clustered s.e.	(0.0334)	(0.0489)	(0.0295)	(0.0266)	(0.0178)	(0.1560)	(0.2472)	(0.1498)	(0.1347)	(0.0964)
Pre-Colonial States	0.2013**	0.1310**	0.0765***	0.0798***	0.0688***	1.0949**	0.6819**	0.4089***	0.4544***	0.3994***
Double-clustered s.e.	(0.0956)	(0.0519)	(0.0240)	(0.0216)	(0.0235)	(0.5488)	(0.2881)	(0.1432)	(0.1430)	(0.1493)
Adjusted R-squared	0.033	0.271	0.357	0.375	0.379	0.046	0.319	0.417	0.448	0.456
Country Fixed Effects	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Controls at the Pixel Level	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Controls at the Ethnic-Country Level	No	No	No	No	Yes	No	No	No	No	Yes
Observations	66570	66570	66570	66173	66173	66570	66570	66570	66173	66173

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 5 reports OLS estimates associating regional development, as reflected in satellite light density at night, with pre-colonial ethnic institutions. The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees (around 12 x 12 kilometers). In columns (1)-(5) the dependent variable is a dummy variable that takes on the value of one if the pixel is lit and zero otherwise. In columns (6)-(10) the dependent variable is the log (0.01 + light density at night from satellite). In Panel A we measure pre-colonial ethnic institutions with Murdock's (1967) jurisdictional hierarchy beyond the local community index. In Panel B we examine in a flexible way the role of each pre-colonial political structure on regional development, augmenting the specification with three dummy variables that identify ethnic groups that were organized as petty chiefdoms, large paramount chiefdoms, and large states before colonization. The omitted category consists of stateless upon colonization ethnic groups.

In columns (3)-(5) and (8)-(10) we control for ln (pixel population density). In columns (4), (5), (9) and (10) we control for a set of geographic and location variables at the pixel level. The set of controls includes the distance of the centroid of each pixel from the respective capital city, the distance of each pixel from the sea coast, the distance of each pixel from the national border, an indicator for pixels that have water (lakes, rivers, and other streams), an indicator for pixels with diamond mines, an indicator for pixels with oil fields, pixel's land suitability for agriculture, pixel's mean elevation, pixel's average value of a malaria stability index, and the log of the pixel's area. In columns (5) and (10) we also control for location and geographic features at the ethnic-country level. Specifically, the set of control variables includes the distance of the centroid of each ethnicity-country area from the respective capital city, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 6: Pre-colonial Ethnic Institutions and Geographic Characteristics within Contiguous Ethnic Homelands in the Same Country

	Dependent variable is:								
	Diamond Indicator	Oil Indicator	Water Indicator	Distance to the Capital	Distance to the Sea	Distance to the Border	Malaria Stability	Land Suitability	Mean Elevation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Jurisdictional Hierarchy	0.0011	0.0063	-0.0058	-9.1375	9.4628	-3.7848	-0.001	-0.0059	21.3826
Double-clustered s.e.	(0.0008)	(0.0051)	(0.0077)	(20.1494)	(6.3349)	(10.0488)	(0.0181)	(0.0060)	(19.5522)
Adjusted R-squared	0.508	0.019	0.126	0.915	0.944	0.660	0.629	0.835	0.767
Mean of Dependent Variable	0.004	0.036	0.125	521.899	643.984	157.596	0.754	0.377	743.446
Observations	78139	78139	78139	78139	78139	78139	77985	77983	78139
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6 reports OLS estimates associating various geographical, ecological, and other characteristics with pre-colonial ethnic institutions within pairs of adjacent ethnicities. The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees (around 12 x 12 kilometers). Every pixel falls into the historical homeland of ethnicity i in country c that is adjacent to the homeland of another ethnicity j in country c, where the two ethnicities differ in the degree of political centralization.

The dependent variable in column (1) is a binary index that takes on the value of one if there is a diamond mine in the pixel; in column (2) a binary index that takes on the value of one if an oil/petroleum field is in the pixel; in column (3) a binary index that takes on the value of one if a water body falls in the pixel. In columns (4)-(6) the dependent variable is the distance of each pixel from the capital city, the sea coast, and the national border, respectively. In column (7) the dependent variable is the average value of a malaria stability index; in column (8) the dependent variable is land's suitability for agriculture; in column (9) the dependent variable is elevation. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 7: Pre-colonial Ethnic Institutions and Regional Development within Contiguous Ethnic Homelands in the Same Country

	All Observations			Difference in Jurisdictional Hierarchy Index > | 1 |			One Ethnic Group was Part of a Pre-colonial State		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Jurisdictional Hierarchy	0.0253*	0.0152**	0.0137**	0.0280*	0.0170**	0.0151**	0.0419**	0.0242**	0.0178***
Double-clustered s.e.	(0.0134)	(0.0073)	(0.0065)	(0.0159)	(0.0079)	(0.0072)	(0.0213)	(0.0096)	(0.0069)
Adjusted R-squared	0.329	0.391	0.399	0.338	0.416	0.423	0.424	0.501	0.512
Observations	78139	78139	77833	34180	34180	34030	16570	16570	16474
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Controls at the Pixel Level	No	No	Yes	No	No	Yes	No	No	Yes

Table 7 reports adjacent-ethnicity (ethnic-pair-country) fixed effects OLS estimates associating regional development, as reflected in satellite light density at night with pre-colonial ethnic institutions, as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community index within pairs of adjacent ethnicities with a different degree of political centralization in the same country. The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees (around 12 x 12 kilometers). Every pixel falls into the historical homeland of ethnicity i in country c that is adjacent to the homeland of another ethnicity j in country c, where the two ethnicities differ in the degree of political centralization. The dependent variable is a dummy variable that takes on the value of one if the pixel is lit and zero otherwise.

In columns (4)-(6) we restrict estimation to adjacent ethnic groups with large differences in the 0-4 jurisdictional hierarchy beyond the local level index (greater than one point). In columns (7)-(9) we restrict estimation to adjacent ethnic groups in the same country where the one of the two ethnicities was part of a large state before colonization (in this case the jurisdictional hierarchy beyond the local level index equals 3 or 4). In columns (2), (3), (5), (6), (8), and (9) we control for ln (pixel population density). In columns (3), (6), and (9) we control for a set of geographic and location variables at the pixel level. The set of controls includes the distance of the centroid of each pixel from the respective capital, the distance of each pixel from the sea coast, the distance of each pixel from the national border, an indicator for pixels that have water (lakes, rivers, and other streams), an indicator for pixels with diamond mines, an indicator for pixels with oil fields, pixel's land suitability for agriculture, pixel's mean elevation, pixel's average value of a malaria stability index, and the log of the pixel's area. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 8A: Pre-colonial Ethnic Institutions and Regional Development within Contiguous Ethnic Homelands in the Same Country
Pixel-Level Analysis in Areas Close to the Ethnic Border

	All Observations Adjacent Ethnicities in the Same Country			Difference in Jurisdictional Hierarchy Index > | 1 |			One Ethnic Group was Part of a pre-colonial Centralized State		
	<100 km of ethnic border	<150 km of ethnic border	<200 km of ethnic border	<100 km of ethnic border	<150 km of ethnic border	<200 km of ethnic border	<100 km of ethnic border	<150 km of ethnic border	<200 km of ethnic border
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Border Thickness (total: 50KM; 25KM from each side of the ethnic boundary)									
Jurisdictional Hierarchy	0.0194*	0.0230**	0.0231**	0.0269***	0.0285***	0.0280***	0.0240***	0.0297***	0.0300***
Double-clustered s.e.	(0.0102)	(0.0106)	(0.0102)	(0.0092)	(0.0088)	(0.0084)	(0.0090)	(0.0067)	(0.0069)
Adjusted R-squared	0.463	0.439	0.429	0.421	0.430	0.434	0.485	0.500	0.501
Observations	6830	10451	13195	3700	5421	6853	2347	3497	4430
Panel B: Border Thickness (total: 100KM; 50KM from each side of the ethnic boundary)									
Jurisdictional Hierarchy	0.0227**	0.0278**	0.0274**	0.0318***	0.0331***	0.0312***	0.0317***	0.0367***	0.0350***
Double-clustered s.e.	(0.0114)	(0.0117)	(0.0108)	(0.0094)	(0.0083)	(0.0076)	(0.0092)	(0.0057)	(0.0068)
Adjusted R-squared	0.467	0.433	0.423	0.458	0.451	0.452	0.525	0.526	0.521
Observations	4460	8081	10825	2438	4159	5591	1538	2688	3621
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Pixel Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 8B: Pre-colonial Ethnic Institutions and Regional Development within Contiguous Ethnic Homelands in the Same Country
Pixel-Level Analysis in Areas Close to the Ethnic Border
Controlling for an RD-type 4th-order Polynomial in Distance to the Ethnic Border

	All Observations Adjacent Ethnicities in the Same Country			Difference in Jurisdictional Hierarchy Index > | 1 |			One Ethnic Group was Part of a pre-colonial Centralized State		
	<100 km of ethnic border	<150 km of ethnic border	<200 km of ethnic border	<100 km of ethnic border	<150 km of ethnic border	<200 km of ethnic border	<100 km of ethnic border	<150 km of ethnic border	<200 km of ethnic border
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Border Thickness (total: 50KM; 25KM from each side of the ethnic boundary)									
Jurisdictional Hierarchy	0.0472***	0.0425***	0.0428***	0.0470**	0.0517***	0.0480***	0.0407*	0.0419*	0.0384**
Double-clustered s.e.	(0.0156)	(0.0146)	(0.0147)	(0.0222)	(0.0192)	(0.0180)	(0.0231)	(0.0234)	(0.0190)
Adjusted R-squared	6830	10451	13195	3700	5421	6853	2347	3497	4430
Observations	0.464	0.441	0.431	0.422	0.432	0.436	0.489	0.504	0.504
Panel B: Border Thickness (total: 100KM; 50KM from each side of the ethnic boundary)									
Jurisdictional Hierarchy	0.0475***	0.0336**	0.0343***	0.0525***	0.0426***	0.0376***	0.0495**	0.0244	0.0147
Double-clustered s.e.	(0.0166)	(0.0146)	(0.0120)	(0.0179)	(0.0131)	(0.0094)	(0.0245)	(0.0190)	(0.0192)
Adjusted R-squared	0.468	0.434	0.424	0.460	0.453	0.453	0.530	0.530	0.525
Observations	4460	8081	10825	2438	4159	5591	1538	2688	3621
RD-type Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjacent-Ethnic-Groups Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Pixel Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Table 8A and Table 8B report adjacent-ethnicity (ethnic-pair-country) fixed effects OLS estimates associating regional development with pre-colonial ethnic institutions, as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community index within pairs of adjacent ethnicities. The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees (around 12.5 x 12.5 kilometers). Every pixel falls into the historical homeland of ethnicity i in country c that is adjacent to the homeland of another ethnicity j in country c, where the two ethnicities differ in the degree of political centralization. The dependent variable is a dummy variable that takes on the value of one if the pixel is lit and zero otherwise. All columns report local-linear-regression estimates restricting estimation to pixels close to the ethnic border. In columns (1), (4), and (7) we focus on pixels within 100 kilometers on each side of the national border. In columns (2), (5), and (8) we focus on pixels within 150 kilometers from the ethnic border. In columns (3), (6), and (9) we focus on pixels within 200 kilometers from the ethnic border. To account for population mixing in the areas along the ethnic boundary, measurement error on Murdock's map, and bleeding/blooming in the luminosity data we exclude from the estimation pixels very close to the ethnic border. In Panel A we exclude areas 25km from each side of the ethnic border (total 50km). In Panel B we exclude areas 50km from each side of the ethnic border (total 100km).

In all specifications we control for ln (pixel population density) and a rich set of geographic and location variables at the pixel level. The set of controls includes the distance of the centroid of each pixel from the respective capital city, the distance of each pixel from the sea coast, the distance of each pixel from the national border, an indicator for pixels that have water (lakes, rivers, and other streams), an indicator for pixels with diamond mines, an indicator for pixels with oil fields, pixel's land suitability for agriculture, pixel's mean elevation, pixel's average value of a malaria stability index, and the log of the pixel's area. All specifications in Table 8B include a fourth-order regression discontinuity type polynomial in distance of each pixel to the "thick" ethnic border, where we allow for the coefficients on the distance terms to be different for the relatively high and the relatively low institutional quality homelands. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Supplementary Appendix

(not for publication)

Pre-colonial Ethnic Institutions and Contemporary African Development *

Stelios Michalopoulos
Brown University and NBER

Elias Papaioannou
Dartmouth College, NBER and CEPR

June 29, 2012

Abstract

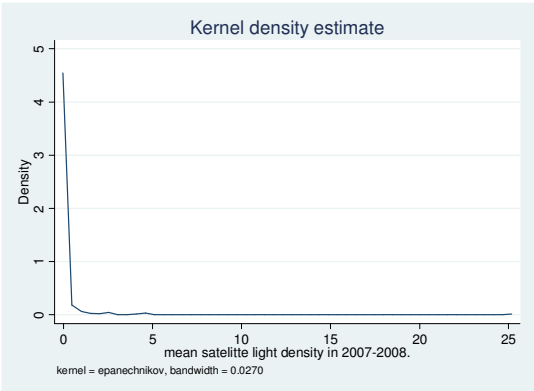
The Supplementary Appendix reports (*i*) descriptive statistics, (*ii*) further cross-validation tests of the luminosity data, (*iii*) numerous sensitivity checks and (iv) the detailed Data Appendix.

*We thank 4 anonymous referees and the Editor for proposing many of the sensitivity checks reported in this Supplementary Appendix. All errors are our sole responsibility.

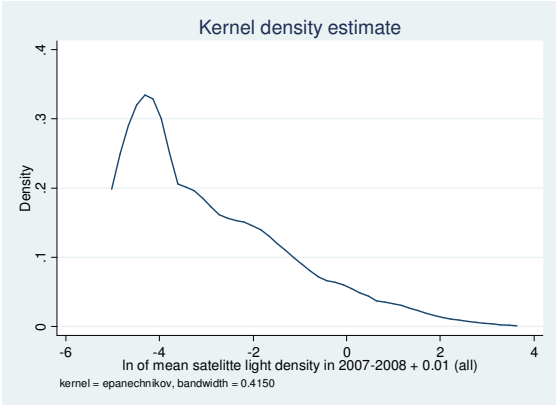
Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

1 Descriptive Analysis

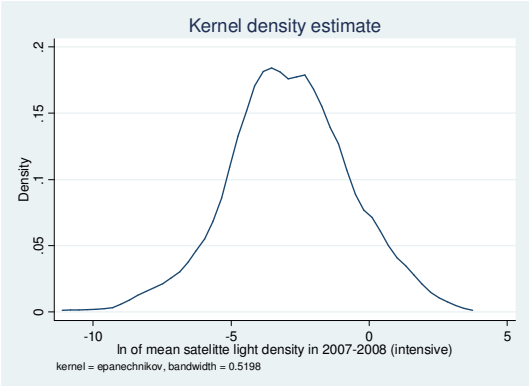
Appendix Figure 1a reports the distribution of satellite light density at night across ethnic homelands. Appendix Figure 1b plots the distribution of the log of light density adding a small number ($y_{i,c} \equiv \ln(0.01 + Light\ Density_{i,c})$. The latter is the dependent variable used in Tables 2, 3A, and 4. This transformation assures that (i) we use all observations and (ii) we minimize the problem of outliers. Appendix Figure 1c graphs the distribution of log light density across lit ethnic-country areas ($y_{i,c} \equiv \ln(LightDensity_{i,c})$). We use log luminosity as the dependent variable in Table 3B so as to show that our results at the ethnic homeland level are not sensitive to the transformation of the luminosity data.



Appendix Figure 1a



Appendix Figure 1b



Appendix Figure 1c

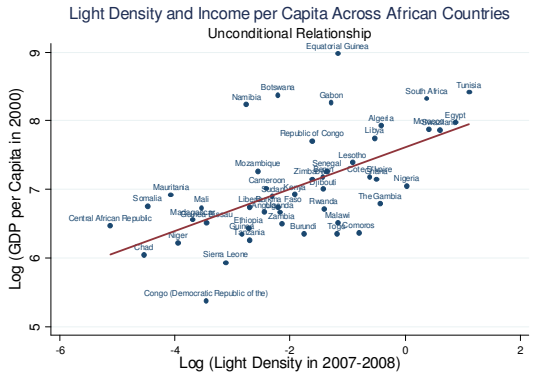
Appendix Table 2 Panel A reports summary statistics for all control variables at the ethnic-country homeland level. Appendix Table 2 Panel B reports summary statistics for the control variables at the pixel level.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

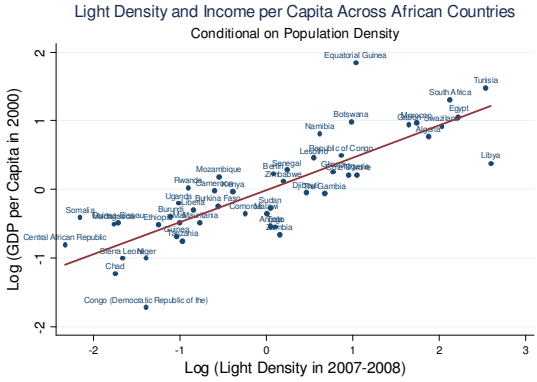
2 Further Cross-Validation Checks

We performed several cross-validation exercises investigating the relationship between luminosity and various economic indicators across African countries and regions within countries.

Cross-country analysis We start by examining whether luminosity correlates with development across African countries. Appendix Figure 2*a* illustrates the unconditional correlation between log light density and log GDP per capita in 2000. The R^2 is 0.35 and the estimate is more than 6 standard errors larger than zero. Besides economic performance, light density also reflects urbanization. Appendix Figure 2*b* shows the relationship between log GDP per capita and log light density partialling out log population density. The relationship between log light density and log GDP p.c. is now stronger (the coefficient increases from 0.31 to 0.47 and the t-stat jumps to 10).



Appendix Figure 2*a*

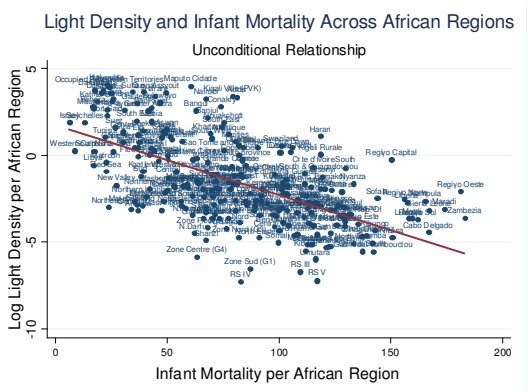


Appendix Figure 2*b*

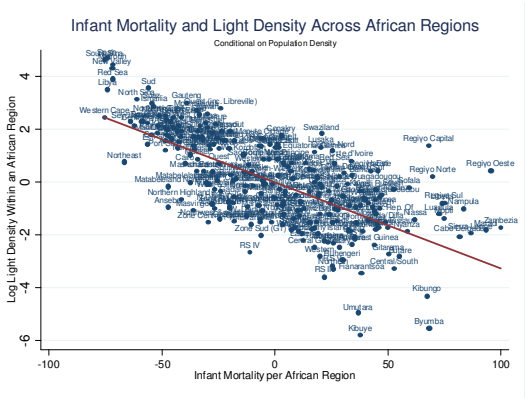
Cross-region analysis Appendix Figures 3*a* and 3*b* plot luminosity and infant mortality across African administrative regions, using data from the Center for International Earth Science Information Network (CIESIN) of Columbia University's Earth Institute which are comparable across African countries.[1] The figures illustrate the significantly negative correlation between log light density and infant mortality across 264 African regions. The estimate is -9.44 with a t-stat of 9; when we condition on log population density, the estimate increases in absolute value (-14.89) retaining its significance at the 99% confidence level.

[1] The data is available at: http://sedac.ciesin.columbia.edu/povmap/

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9



Appendix Figure 3*a*



Appendix Figure 3*b*

Analysis across DHS Enumeration Areas Appendix Table 1 reports the unconditional and conditional on population density correlations between light density and the DHS household composite wealth index, presence of electricity and average schooling, for four large African countries across all African regions. Across all countries studied, Nigeria, Tanzania, Zimbabwe, and the Democratic Republic of Congo, there is highly significant (within country) correlation between luminosity and these proxy measures of development (see also Figures $2a - 2d$ in the main body of the paper for a graphical illustration).

3 Further Sensitivity Checks

We performed various sensitivity checks to explore the robustness of our results.

Outliers In column (1) of Appendix Table 3 we report estimates dropping observations where the dependent variable exceeds the $99th$ percentile.[2] This is useful because there are some extreme values in the luminosity data (see discussion in Section 2 and the descriptive statistics in Table $1A$). The coefficient on the jurisdictional hierarchy index (0.183) is unchanged compared to the analogous specification in column (4) of Table $3A$ (0.177). We also repeated estimation dropping ethnic areas where capital cities fall. This is an appropriate check because capitals are likely to be populated by people from several ethnic groups and thus the ethnic-specific index of jurisdictional hierarchy index may be inappropriate. Moreover, by excluding areas where capitals fall, we further account for outliers in the dependent variable. The positive and highly significant correlation between pre-colonial political centralization and regional development remains intact (column (2)).

[2] In the previous drafts of the paper we also estimated specifications excluding also the top 2% or the top 5% of the dependent variable. If anything, the coefficient on the jurisdictional hierarchy index increases and becomes more accurately estimated.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Regional Heterogeneity We also repeated our analysis excluding each time a different African region (using the classification of Nunn (2008)). Doing so allows us to account for the fact that colonization took different forms across Africa and this may have affected the role of local institutions during the colonial era and after independence. Appendix Table 3, columns (3)-(7) report the results. The estimates show that the strong positive association between pre-colonial ethnic political centralization and regional development is not driven by a particular African region.

Population Density In Appendix Table 4*A* we report cross–sectional and within-country estimates associating pre-colonial ethnic institutions with contemporary population density at the ethnic-country level. Appendix Table 4*B* reports analogous specifications at the pixel level. While urbanization has been linked to the emergence of complex political institutions before colonization in Africa and thus these estimates may suffer from endogeneity, it is useful to see whether the complexity of pre-colonial ethnic institutions correlates with contemporary population density, since the latter may be capturing (to some extent at least) the level of development. There is a significant and positive correlation between the jurisdictional hierarchy beyond the local community index and contemporary population density. However, the estimates indicate that ethnic institutions are less strongly related to population density as compared to luminosity; for example the standardized "beta" coefficient for the models in columns (1) and (3) is 0.08 and 0.10; the analogous specifications with log luminosity as the dependent variable (column (1), Table 2 and column (1) Table 3*A*) yield beta coefficients of 0.22 and 0.19, respectively. These comparisons suggest that while political centralization does interact with population density, its relationship with economic development is stronger.

Results with Afrobarometer's Living Condition Index and Education While we consider the use of the satellite images at night as a proxy for development part of our contribution, we also estimated specifications with alternative proxies of regional development using the Afrobarometer Surveys. Specifically, in Appendix Table 5 we report specifications associating pre-colonial ethnic institutions with two measures from the 2005 Afrobarometer Surveys that capture economic well-being, a living conditions index and years of schooling. The Afrobarometer surveys are based on interviews conducted in a random sample of either $1,200$ or $2,400$ individuals of voting age in 17 Sub-Saharan African countries.[3] As a result, the sample drops significantly not only because the number of countries and ethnicities falls, but also because in many countries we do not have adequate spatial coverage across all ethnic areas.

[3] These countries are: Benin, Botswana, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Namibia, Nigeria, Senegal, South Africa, Tanzania, Uganda, Zambia, and Zimbabwe.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

We assign each respondent's current residence area (Afrobarometer's enumeration area) to one of the 834 ethnic homelands in Murdock's (1959) ethnolinguistic map and then aggregated the living conditions index and schooling at the ethnic homeland level in each country. We then estimated our baseline specification (equation (1)) at the ethnic-country homeland level using the living conditions index and average years of schooling as the dependent variable.[4] The results with these alternative proxies of development are in line with our estimates with luminosity. On average respondents residing in ethnic homelands with centralized pre-colonial political institutions report better living conditions. Similarly, the average level of education is higher in ethnic areas of politically centralized societies. Yet it should be noted that due to the limited number of respondents within each ethnic area, these results should be cautiously interpreted.

[4]The living conditions index reflects respondent's view of their present living conditions which can be: (i) very bad, (ii) fairly bad, (iii) neither good nor bad, (iv) fairly good, or (v) very good.

The schooling data in constructed aggregating respondent's answers on the following categories: (i) no formal schooling, (ii) informal schooling only, (iii) some primary schooling, (iv) primary school completed, (v) some secondary school/high school, (vi) secondary school completed/high school, (vii) post-secondary qualifications, but no university, (viii) some university, (ix) university completed, and (x) post-graduate.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

4 Data Appendix

4.1 Variables at the ethnicity-country level / pixel level

Light Density at Night: Light Density is the average luminosity across pixels that fall within the unit of analysis. We use the average of the values in 2007 and 2008. In the regressions at the ethnic homeland we use Log (0.01 + Average Luminosity) or Log (Average Luminosity). In the regressions at the pixel level we use as the dependent variable a dummy variable that takes on the value of one if the area is lit and zero otherwise; we also use the Log (0.01 + Average Luminosity)
Source: Available at http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html.

Population Density: Log (0.01 + population density per sq. km. in 2000). For the pixel-level analysis we use Log (population density per sq. km. in 2000). *Source: Nelson, Andy,* 2004. *African Population Database Documentation, UNEP GRID Sioux Falls. Available at:* http://na.unep.net/siouxfalls/datasets/datalist.php

Water Area: Log (1 + total area covered by rivers or lakes in sq. km.). For the pixel-level analysis we use an indicator that equals 1 if there is some water body at the pixel. *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, CO, USA. Global Ministry Mapping System.*

Elevation: Average elevation in km in each ethnicity-country or in each pixel. *Source: National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colorado.*
http://www.sage.wisc.edu/atlas/data.php?incdataset=Topography

Land Suitability for Agriculture: Average land quality for cultivation within each ethnicity-country or within each pixel. The index is the product of two components capturing the climatic and soil suitability for farming. *Source: Michalopoulos (2012); Original Source: Atlas of the Biosphere.*
Available at http://www.sage.wisc.edu/iamdata/grid_data_sel.php.

Malaria Stability Index: The index takes into account the prevalence and type of mosquitoes indigenous to a region, their human biting rate, their daily survival rate, and their incubation period. We use the average value within each ethnicity-country or within each pixel. *Source: Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004)*

Distance to the Capital City: The geodesic distance from the centroid of each ethnic group in a country or of each pixel to the capital city of the country it belongs to, measured in 1000*s* of km's. *Source: Calculated using the Haversine formula.*

Distance to the Sea Coast: The geodesic distance from the centroid of each ethnic group in a country or of each pixel to the nearest coastline, measured in 1000*s* of km's. *Source:*

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Global Mapping International, Colorado Springs, Colorado, USA. Series name: Global Ministry Mapping System. Series issue: Version 3.0

Distance to the Border: The geodesic distance from the centroid of each ethnic group in a country or from each pixel to the nearest border, measured in 1000*s* of km's. *Source: Calculated using ArcGis.*

Petroleum: Indicator variable that equals one if there is an oil field in the homeland or pixel of ethnic group i in country c. *Source: The Petroleum Dataset v.1.1 contains information on all known on-shore oil and gas deposits throughout the world. http://www.prio.no/CSCW/Datasets/Geographical-and-Resource/Petroleum-Dataset/Petroleum-Dataset-v11/*

Diamond: Indicator variable that equals one if there is a diamond mine in the homeland or pixel of ethnic group i in country c. *Source: Map of Diamond Resources. www.prio.no/CSCW/Datasets/Geographical-and-Resource/Diamond-Resources/*

4.2 Country-level variables

Rule of Law: The index is "capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence." The standardized index ranges from -2.5 to $+2.5$ with higher values indicating better functioning institutions. We use the value of the index in 2007. *Source: World Bank Governance Matters Indicators Database (Kaufman, Kraay, and Mastruzzi (2005)). available at: http://info.worldbank.org/governance/wgi/index.asp*

Income: Log of per capita GDP in PPP in 2007. *Source: Penn World Tables.*

4.3 Pre-colonial Ethnicity-level variables

Jurisdictional Hierarchy beyond Local Community: Ordered variable ranging from 0 to 4 indicating the number of jurisdictional levels (political complexity) in each society above the local level**.** A 0 indicates stateless societies, 1 and 2 indicate petty and large paramount chiefdoms (or their equivalent), 3 and 4 indicate large states. *Source: Murdock (1967); variable code in the Ethnolinguistic Atlas v33; A revised version of Murdock's Atlas has been made available by J. Patrick Gray at: http://eclectic.ss.uci.edu/~drwhite/worldcul/EthnographicAtlasWCRevisedByWorldCultures.sav.*

Centralization Indicator: This binary index takes the value 0 if the Jurisdictional Hierarchy beyond Local Community variable equals 0 or 1. The index takes on the value 1 if the Jurisdictional Hierarchy Beyond Local Community variable equals 2, 3, and 4. This

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

aggregation follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967).*

Gathering: Binary index that indicates the reliance of the economy on "*the collection of wild plants and small land fauna.*" The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5% dependence. *Source: Murdock (1967); variable code in the Ethnographic Atlas v1.*

Hunting: Binary index that indicates the intensity in hunting (including trapping and fowling). The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5%. *Source: Murdock (1967); variable code in the Ethnographic Atlas v2.*

Fishing: Binary index that indicates the intensity in fishing (including shell fishing and the pursuit of large aquatic animals). The index equals zero when the dependence is between 0% and 5%; the index equals one when dependence is greater than 5%. *Source: Murdock (1967); variable code in the Ethnographic Atlas v3.*

Animal Husbandry: $0-9$ scale index reflecting the percentage of subsistence coming from animal husbandry for each ethnicity at the time of colonization. The index equals 0 when there $0\%-5\%$ dependence; 1 when there is $6\%-15\%$ dependence; 2 when there is $16\%-25\%$ dependence; 3 when there is $26\%-35\%$ dependence; 4 when there is $36\%-45\%$ dependence; 5 when there is $46\%-55\%$ dependence; 6 when there is $56\%-65\%$ dependence; 7 when there is $66\%-75\%$ dependence; 8 when there is $76\%-85\%$ dependence; and 9 when there is $86\%-100\%$ dependence. *Source: Murdock (1967); variable code in the Ethnographic Atlas v4.*

Milking: Binary index that equals one when "*domestic animals are milked more often that sporadically*" and zero when "*little or no milking*". *Source: Murdock (1967); variable code in the Ethnographic Atlas v41.*

Agriculture Dependence: $0-9$ scale index reflecting the intensity of agriculture. "*It includes penetration of the soil, planting, tending the growing crops, and harvesting but not subsequent food preparation*". The index equals 0 when there $0\%-5\%$ dependence; 1 when there is $6\%-15\%$ dependence; 2 when there is $16\%-25\%$ dependence; 3 when there is $26\%-35\%$ dependence; 4 when there is $36\%-45\%$ dependence; 5 when there is $46\%-55\%$ dependence; 6 when there is $56\%-65\%$ dependence; 7 when there is $66\%-75\%$ dependence; 8 when there is $76\%-85\%$ dependence; and 9 when there is $86\%-100\%$ dependence. *Source: Murdock (1967); variable code in the Ethnographic Atlas v5.*

Agriculture Type: $0-4$ scale index reflecting the type of agriculture. The index equals 0 when there is "*no agriculture*"; 1 when there is "*causal agriculture*"; 2 when there is "*extensive or shifting agriculture*"; 3 when there is "*intensive agriculture*"; and 4 when there is "*intensive irrigated agriculture.*" *Source: Murdock (1967); variable code in the Ethnographic*

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Atlas v28.

Polygyny: Indicator that equals one when polygyny is practiced and zero otherwise. *Source: Murdock (1967); variable code in the Ethnographic Atlas v8.*

Alternative Polygyny: Alternative indicator that equals one when polygyny is practiced and zero otherwise, based on Murdock's domestic organization variable. The indicator equals zero when "*independent nuclear, monogamous families*" are the norm and one when polygyny is common. *Source: Murdock (1967); variable code in the Ethnographic Atlas v9; the construction of the variable follows Fenske (2009).*

Clan Communities: Indicator that equals one when Murdock's community marriage organization variable indicates that "*clan communities or clan barrios*" are present and zero otherwise. *Source: Murdock (1967); variable code in the Ethnographic Atlas v15.*

Settlement Pattern: Ordered variable ranging from 1 to 8 quantifying *"settlement pattern of each group".* 1 indicates fully nomadic (migratory) groups, 2 indicates semi-nomadic groups, 3 indicates semi-sedentary groups, 4 identifies groups that live in compact and impermanent settlements, 5 indicates societies those in neighborhoods of dispersed family homes, 6 indicates for groups in separated hamlets forming a single community, 7 indicates societies living in compact and relatively permanent settlements, and 8 denotes the groups residing in complex settlements. *Source: Murdock (1967); variable code in the Ethnographic Atlas v30.*

Complex Settlements: Indicator that equals one for ethnicities living in compact and relatively permanent settlements ($v30 = 7$) or in complex settlements ($v30 = 8$), and zero otherwise. *Source: Murdock (1967); variable code in the Ethnographic Atlas v30.*

Jurisdictional Hierarchy of Local Community: Ordered variable ranging from 0 to 2 reflecting the hierarchy of local community organization. A zero score indicates the theoretical minimum of two (e.g., family and band), while a score of 2 indicates the theoretical maximum of four levels (e.g., nuclear family, extended family, clan barrio, village levels). *Source: Murdock (1967); variable code in the Ethnographic Atlas v32.*

Patrilineal Descent Type: Indicator that equals one if the society is characterized by patrilineal descent and zero otherwise. *Source: Murdock (1967); variable code in the Ethnographic Atlas v43.*

Class Stratification: Ordered variable ranging from 0 to 4 quantifying "*the degree of class differentiation, excluding purely political and religious statuses*". A zero score indicates "*absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom.*" A score of 1 indicates "*the presence of wealth distinctions, based on possession or distribution of property, which however have not crystallized into distinct and hereditary social classes.*" A score of 2 indicates "*elite stratification*

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

in which an elite class derives its superior status from control over scarce resources, particularly land, and is thereby differentiated from a property-less proletariat or serf class". A score of 3 indicates a "*dual stratification into a hereditary aristocracy and a lower class of ordinary commoners or freemen, where traditionally ascribed noble status is at least as decisive as control over scarce resources.* A score of 4 indicates "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses.*" *Source: Murdock (1967); variable code in the Ethnographic Atlas v67.*

Class Stratification Indicator: Following Gennaioli and Rainer (2007) we define a dummy stratification index that equals zero when Murdock's variable equals zero indicating "*absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom,*" and one when Murdock's class stratification measure equals 1, 2, 3, or 4. *Source: Murdock (1967); variable code in the Ethnographic Atlas v67.*

Elections: Indicator that equals 1 when succession to the office of the local headman is conducted via "*election or other formal consensus, nonhereditary*" and zero otherwise. *Source: Murdock (1967); variable code in the Ethnographic Atlas v72.*

Slavery: Indicator that equals one when some type of slavery (hereditary, incipient, or significant) is present and zero when there it is absent or near absent. *Source: Murdock (1967); variable code in the Ethnographic Atlas v70; the construction of the index follows Fenske (2009).*

Inheritance Rules for Property: Indicator that equals one when some form of inheritance rule of real property (land) is present; the binary indicator equals zero when there is "*absence of individual property rights*". *Source: Murdock (1967); variable code in the Ethnographic Atlas v74.*

4.4 Data for the Cross-Validation Analysis

DHS Composite Wealth Index: The DHS wealth index is composed taking into account consumer durables, electricity, toilet facilities, source of drinking water, dwelling characteristics and some country-specific attributes like whether there is a domestic servant, for example. The measure is derived by the DHS using principal component analysis to assign indicator weights resulting in a composite standardized index for each country. *Source: DHS; available at: http://www.measuredhs.com*

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

References

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs (2004): "A Global Index of the Stability of Malaria Transmission," *American Journal of Tropical Medicine and Hygiene*, 70(5), 486–498.

Michalopoulos, S. (2012): "The Origins of Ethnolinguistic Diversity," *American Economic Review*, 102(4), 1508–1539.

Murdock, G. P. (1967): *Ethnographic Atlas*. University of Pittsburgh Press, Pittsburgh, PA.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Appendix Table 1: Satellite Light Density at Night and DHS Measures of Household Wealth, Electrification, and Schooling

	Tanzania 2007		Zimbabwe 2005		Nigeria 2008		Congo, DRC 2007	
	Unconditional Correlation	Conditional on PD	Unconditional Correlation	Conditional on PD	Unconditional Correlation	Conditional on PD	Unconditional Correlation	Conditional on PD
DHS Composite Wealth Index	0.7828	0.7194	0.8809	0.7540	0.7219	0.5496	0.8599	0.7424
Electrification	0.7045	0.6296	0.8423	0.7012	0.6064	0.4188	0.7733	0.6461
Years of Schooling	0.5575	0.5035	0.7171	0.5191	0.5603	0.5385	0.6344	0.5378
Enumeration Areas	466	466	398	398	886	886	293	293

The table reports unconditional and conditional on log population density correlations between satellite light density at night (log (0.01 + light density)) and three proxy measures of development from the Demographic and Health Surveys (DHS) for Tanzania, Zimbabwe, Nigeria, and the Democratic Republic of Congo. The unit of analysis is the enumeration area of the Demographic and Health Surveys. As within each enumeration area there are several households, we average household responses across the enumeration area and use the respective mean values. Light density at night is the average light density in a radius of 10 km from the centroid of the DHS enumeration area. The DHS composite wealth index is derived using principal component analysis taking into account access of households to electricity, toilet facilities, clean drinking water, dwelling characteristics, etc. Schooling years denotes average years of schooling of the household's head. Electrification denotes whether households have access to electricity. All correlations are significant at the 1% confidence level.

Appendix Table 2: Summary Statistics for the Control Variables

variable	Obs.	mean	st. dev.	p25	median	p75	min	max
				Panel A: Ethnic Homeland Sample				
Ln (0.01 + Population Density)	683	2.90	2.02	2.18	3.24	4.20	-4.61	7.43
Ln (1 + Water Area)	683	0.29	0.44	0.00	0.12	0.36	0.00	3.12
Ln (Surface Area)	683	2.16	1.60	1.14	2.23	3.25	-2.21	6.20
Mean Elevation	683	0.63	0.47	0.28	0.49	1.00	0.00	2.18
Land Suitability For Agriculture	683	0.43	0.24	0.28	0.45	0.59	0.00	0.98
Malaria Stability Index	683	0.72	0.33	0.50	0.87	1.00	0.00	1.00
Oil Deposit Indicator	683	0.08	0.27	0.00	0.00	0.00	0.00	1.00
Diamond Mine Indicator	683	0.10	0.30	0.00	0.00	0.00	0.00	1.00
Distance to the Capital City	683	0.50	0.39	0.24	0.40	0.63	0.01	1.88
Distance to the Sea Coast	683	0.59	0.42	0.20	0.56	0.92	0.00	1.70
Distance to the National Border	683	0.10	0.11	0.02	0.06	0.15	0.00	0.62
Rule of Law in 2007	680	-0.88	0.53	-1.35	-0.96	-0.42	-2.63	0.65
Log GDP p.c. in 2007	680	7.12	0.90	6.67	7.10	7.57	5.07	9.89
Jurisdictional Hierarchy	683	1.22	0.97	0.00	1.00	2.00	0.00	4.00
Political Centralization	683	0.36	0.48	0.00	0.00	1.00	0.00	1.00
				Panel B: Pixel Sample				
Ln (Population Density)	66570	2.37	1.85	1.02	2.48	3.72	-2.20	10.10
Water Body Indicator	66570	0.14	0.34	0.00	0.00	0.00	0.00	1.00
Ln (Area)	66570	5.11	0.27	5.12	5.23	5.25	3.91	5.26
Mean Elevation	66570	690.65	523.95	293.89	547.61	1064.89	-219.48	4314.56
Land Suitability For Agriculture	66314	0.38	0.28	0.16	0.35	0.57	0.00	1.00
Malaria Stability Index	66428	0.70	0.41	0.30	1.00	1.00	0.00	1.00
Oil Deposit Indicator	66570	0.03	0.18	0.00	0.00	0.00	0.00	1.00
Diamond Mine Indicator	66570	0.00	0.06	0.00	0.00	0.00	0.00	1.00
Distance to the Capital City	66570	507.02	372.43	236.65	403.13	667.76	1.46	1923.01
Distance to the Sea Coast	66570	576.12	423.34	200.35	509.78	905.03	0.55	1701.14
Distance to the National Border	66570	141.82	124.59	44.65	106.35	206.35	0.67	755.60
Jurisdictional Hierarchy	66570	1.49	1.03	1.00	1.00	2.00	0.00	4.00
Political Centralization	66570	0.50	0.50	0.00	0.00	1.00	0.00	1.00

The table reports descriptive statistics for all the control variables employed in the empirical analysis. Panel A reports summary statistics for all control variables in the country-ethnicity sample. Panel B reports summary statistics for all control variables in the pixel sample. The Data Appendix gives detailed variable definitions and data sources.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Appendix Table 3: Sensitivity Analysis at the Ethnic-Country Level

	Excluding						
	Outliers	Capitals	North	South	Central	East	West
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Jurisdictional Hierarchy	0.1825***	0.1754***	0.1834***	0.1676***	0.1327**	0.2218***	0.1655**
	(0.0529)	(0.0576)	(0.0526)	(0.0525)	(0.0560)	(0.0703)	(0.0652)
Adjusted R-squared	0.64	0.657	0.609	0.656	0.697	0.689	0.669
Observations	675	651	642	619	523	495	449
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports within-country OLS estimates associating regional development with pre-colonial ethnic institutions, as reflected in Murdock's (1967) index of jurisdictional hierarchy beyond the local community index. The dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. In column (1) we exclude from the estimation observations where the dependent variable exceeds the 99th percentile. In column (2) we exclude observations where capital cities fall. In column (3) we exclude observations that fall in North Africa. In column (4) we exclude observations that fall in Southern Africa. In column (5) we exclude observations that fall in Central Africa. In column (6) we exclude observations that fall in East Africa. In column (7) we exclude observations that fall in West Africa. The regional classification follows Nunn (2007).

All specifications include a set of country fixed effects (constants not reported). In all specifications we control for population density, location and geography. Specifically the set of control variables includes the log (0.01 + population density), the distance of the centroid of each ethnicity-country area from the respective capital city, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), log (surface area) land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Appendix Table 4A: Pre-colonial Ethnic Institutions and Regional Population Density
Ethnic Homeland Analysis

	Jurisdictional Hierarchy (0-4)				Political Centralization (0-1)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Institutions	0.1687	0.2854***	0.2455**	0.2348**	0.2903	0.5689***	0.3890*	0.4461**
	(0.1104)	(0.0769)	(0.1173)	(0.1114)	(0.2729)	(0.1880)	(0.2351)	(0.2215)
Adjusted R-squared	0.008	0.391	0.403	0.557	0.006	0.391	0.400	0.557
Observations	683	683	682	682	683	683	682	682
Country Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes

The table reports cross-sectional and within-country OLS estimates associating regional development, as reflected in log population density, with pre-colonial ethnic institutions. In columns (1)-(4) we measure pre-colonial ethnic institutions using Murdock's (1967) jurisdictional hierarchy beyond the local community index. In columns (5)-(8) we use a binary political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local community variable. Following Gennaioli and Rainer (2007), this index takes on the value of zero for stateless societies and ethnic groups that are part of petty chiefdoms and one otherwise (for ethnicities that are organized as large paramount chiefdoms and ethnicities that are part of large states).

The specifications in columns (3), (4), (7), and (8) include a set of country fixed effects (constants not reported). In even-numbered columns we control for location and geography. The set of control variables includes the distance of the centroid of each ethnicity-country area from the respective capital city, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Appendix Table 4B: Pre-colonial Ethnic Institutions and Regional Population Density Pixel-Level Analysis

	Jurisdictional Hierarchy (0-4)				Political Centralization (0-1)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Institutions	0.3183***	0.2425*	0.2107**	0.2105***	0.4487**	0.3315	0.3957**	0.3985**
	(0.1095)	(0.1325)	(0.1005)	(0.0811)	(0.2275)	(0.2426)	(0.1939)	(0.1612)
Adjusted R-squared	0.031	0.26	0.335	0.345	0.015	0.254	0.334	0.344
Observations	66570	66570	66173	66173	66570	66570	66173	66173
Country Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Controls Pixel Level	No	No	Yes	Yes	No	No	Yes	Yes
Controls Ethnic-Country Level	No	No	No	Yes	No	No	No	Yes

The table reports OLS estimates associating regional development, as reflected in log population density, with pre-colonial ethnic institutions. In columns (1)-(4) we measure pre-colonial ethnic institutions using Murdock's (1967) jurisdictional hierarchy beyond the local community index. In columns (5)-(8) we use a binary political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local community variable. Following Gennaioli and Rainer (2007), this index takes on the value of zero for stateless societies and ethnic groups that are part of petty chiefdoms and one otherwise (for ethnicities that are organized as large paramount chiefdoms and ethnicities that are part of large states). The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees (around 12 x 12 kilometers).

Specifications in (2)-(4) and (6)-(8) include a set of country fixed effects (constants not reported). In columns (3), (4), (7) and (8) we control for a set of geographic and location variables at the pixel level. The set of controls includes the distance of the centroid of each pixel from the respective capital city, the distance of each pixel from the sea coast, the distance of each pixel from the national border, an indicator for pixels that have water (lakes, rivers, and other streams), an indicator for pixels with diamond mines, an indicator for pixels with oil fields, pixel's land suitability for agriculture, pixel's mean elevation, pixel's average value of a malaria stability index, and the log of the pixel's area. In columns (4) and (8) we also control for location and geographic features at the ethnic-country level. Specifically, the set of control variables includes the distance of the centroid of each ethnicity-country area from the respective capital city, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Appendix Table 5: Pre-colonial Ethnic Institutions and Regional Development within African Countries Results with Afrobarometer Survey Data on Living Conditions and Schooling

	Living Conditions Index				Years of Schooling			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Binary Political Centralization	0.0693**	0.0692**	0.0696**	0.0592**	0.1180*	0.1134*	0.1359**	0.1359**
Double-clustered s.e.	(0.0282)	(0.0286)	(0.0338)	(0.0292)	(0.0605)	(0.0665)	(0.0621)	(0.0621)
Adjusted R-squared	0.358	0.358	0.372	0.395	0.496	0.497	0.523	0.523
Observations	194	194	194	194	193	193	193	193
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes

The table reports within-country OLS estimates associating regional development, as reflected in a living conditions index and years of schooling, with pre-colonial ethnic institutions. We measure pre-colonial ethnic institutions using a binary political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local community variable. Following Gennaioli and Rainer (2007), this index takes on the value of zero for stateless societies and ethnic groups that are part of petty chiefdoms and one otherwise (for ethnicities that are organized as large paramount chiefdoms and ethnicities that are part of large states). All specifications include a set of country fixed effects (constants not reported). In columns (1)-(4) the dependent variable is the log of a living conditions index, as reported in the 2005 Afrobarometer Surveys. In columns (5)-(8) the dependent variable is the log of years of schooling, as reported in the 2005 Afrobarometer Surveys.

In columns (2), (3), (4), (6), (7), and (8) we control for log (0.01 + population density). In columns (3), (4), (7), and (8) we control for location augmenting the specification with distance of the centroid of each ethnicity-country area from the respective capital city, distance from the closest sea coast, and distance from the national border. The set of geographic controls in columns (4) and (8) includes log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Discussion Publishers edition DPR-2026-1409 https://discussionpublishers.com/research/38e31119-d9fa-4a6f-97e0-bb75df17b3f9

Appendix Table 6: Pre-colonial Ethnic Institutions and Regional Development
Pixel-Level Analysis

	Lit/Unlit Pixels					Ln (0.01 + Luminosity)				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Binary Political Centralization	0.1000**	0.0690**	0.0460***	0.0565***	0.0486***	0.5183*	0.3524**	0.2369***	0.3094***	0.2806***
Double-clustered s.e.	(0.0489)	(0.0317)	(0.0149)	(0.0150)	(0.0122)	(0.2723)	(0.1785)	(0.0913)	(0.0903)	(0.0808)
Adjusted R-squared	0.018	0.269	0.357	0.375	0.379	0.022	0.315	0.416	0.447	0.456
Country Fixed Effects	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Controls Pixel Level	No	No	No	Yes	Yes	No	No	No	Yes	Yes
Controls Ethnic-Country Level	No	No	No	No	Yes	No	No	No	No	Yes
Observations	66570	66570	66570	66173	66173	66570	66570	66570	66173	66173

The table reports OLS estimates associating regional development, as reflected in satellite light density at night, with pre-colonial ethnic institutions. We measure pre-colonial ethnic institutions using a binary political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local community variable. Following Gennaioli and Rainer (2007), this index takes on the value of zero for stateless societies and ethnic groups that are part of petty chiefdoms and one otherwise (for ethnicities that are organized as large paramount chiefdoms and ethnicities that are part of large states). The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees. In columns (1)-(5) the dependent variable is a dummy variable that takes on the value of one if the pixel is lit and zero otherwise. In columns (6)-(10) the dependent variable is the log (0.01 + light density at night from satellite).

In columns (3)-(5) and (8)-(10) we control for ln (pixel population density). In columns (4), (5), (9) and (10) we control for a set of geographic and location variables at the pixel level. The set of controls includes the distance of the centroid of each pixel from the respective capital city, the distance of each pixel from the sea coast, the distance of each pixel from the national border, an indicator for pixels that have water (lakes, rivers, and other streams), an indicator for pixels with diamond mines, an indicator for pixels with oil fields, pixel's land suitability for agriculture, pixel's mean elevation, pixel's average value of a malaria stability index, and the log of the pixel's area. In columns (5) and (10) we also control for location and geographic features at the ethnic-country level. Specifically, the set of control variables includes the the distance of the centroid of each ethnicity-country area from the respective capital city, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), log (surface area), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.