NBER WORKING PAPER SERIES

DIVIDE AND RULE OR THE RULE OF THE DIVIDED? EVIDENCE FROM AFRICA

Stelios Michalopoulos
Elias Papaioannou

Working Paper 17184
http://www.nber.org/papers/w17184

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2011

We would like to thank 4 referees and the Editor for their invaluable comments. We thank seminar participants at Dartmouth, Tufts, Oxford, Vienna, Brown, Harvard, NYU, the CEPR Development Economics Workshop, Stanford, UC-Berkeley, UC-Davis, 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. A Supplementary Appendix with additional sensitivity checks is available online at: http://www.dartmouth.edu/~elias/ and http://sites.google.com/site/steliosecon/. 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.

Divide and Rule or the Rule of the Divided? Evidence from Africa
Stelios Michalopoulos and Elias Papaioannou
NBER Working Paper No. 17184
June 2011
JEL No. N0,N17,O1,O11,O55

ABSTRACT

We investigate jointly the importance of contemporary country-level institutional structures and local ethnic-specific pre-colonial institutions in shaping comparative regional development in Africa. We utilize information on the spatial distribution of African ethnicities before colonization and regional variation in contemporary economic performance, as proxied by satellite light density at night. We exploit the fact that political boundaries across the African landscape partitioned ethnic groups in different countries subjecting identical cultures to different country-level institutions. Our regression discontinuity estimates reveal that differences in countrywide institutional arrangements across the border do not explain differences in economic performance within ethnic groups. In contrast, we document a strong association between pre-colonial ethnic institutional traits and contemporary regional development. While this correlation does not necessarily identify a causal relationship, this result obtains conditional on country fixed-effects, controlling for other ethnic traits and when we focus on pairs of contiguous ethnic homelands.

Stelios Michalopoulos
Tufts University
Department of Economics
Stelios.Michalopoulos@tufts.edu

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

1 Introduction

In recent years there has been a surge of empirical research on the determinants of African and more generally global underdevelopment. The predominant institutional view suggests that poorly performing national institutional structures, such as lack of constraints on the executive and poor property rights protection, as well as inefficient legal and court systems are the ultimate causes of underdevelopment (see Acemoglu, Johnson, and Robinson (2005) for a review). This body of research puts an emphasis on the impact of colonization on contemporary country-level institutions and in turn on economic development. Yet in the African context many downplay the importance of colonial and contemporary institutional structures. Recent works on weak and strong states emphasize the limited state capacity of most African states and their inability to provide public goods, collect taxes, and enforce contracts (Acemoglu (2005); Besley and Persson (2009, 2010)). The inability of African governments to broadcast power outside the capital cities has led many influential African scholars to highlight the role of pre-colonial ethnic-specific institutional and cultural traits. This body of research argues that the presence of the Europeans in Africa was (with some exceptions) quite limited both regarding timing and location. As a result of the negligible penetration of Europeans in the mainland and the poor network infrastructure that has endured after independence, it is local ethnic-level, rather than national institutional structures, that shape African development today (see Herbst (2000) for a summary of the arguments).

In this paper we contribute to the literature on the determinants of African development tackling these two distinct, though interrelated, questions. First, do contemporaneous nationwide institutions affect economic performance across regions once we account for hard-to-observe ethnicity-specific traits, culture, and geography? Second, do pre-colonial institutional ethnic characteristics correlate with regional development once we consider country-specific attributes, like economic/institutional performance and national post-independence policies?

In contrast to most previous works that have relied on cross-country data and methods, we tackle these questions exploiting both within-country and within-ethnicity regional variation across African ethnic regions. We utilize data from the pioneering work of Murdock (1959, 1967), who combining various sources has produced a map portraying the spatial distribution of ethnicities (Figure 1a) as well as quantitative information on the economy, institutions, and cultural traits of several ethnic groups around colonization. To overcome the paucity of economic indicators across African ethnicities, we measure regional economic development at the ethnicity-country level using satellite images of light density at night which are available at a fine level of aggregation.

After showing that light density correlates strongly with various measures of economic

development at different levels of aggregation (namely across African countries, administrative regions and villages within countries, as well as within ethnic areas across national boundaries), we examine the impact of contemporary national institutions on economic performance. In line with cross-country studies, we find a positive correlation between rule of law (or control of corruption) and luminosity across African ethnic regions. Yet due to omitted variables and other potential sources of endogeneity this correlation does not imply a causal relationship. To isolate the one-way effect of contemporaneous institutions on regional development we exploit differences in country-level institutional quality within ethnicities partitioned by national boundaries, as identified by intersecting Murdock's ethnolinguistic map with the 2000 Digital Chart of the World (Figure 1*b*).

Figure 1*a*: Ethnic Boundaries

Figure 1*b*: Ethnic and Country Boundaries

The artificial design of African borders, which took place in European capitals in the late 19th century (mainly in the Berlin Conference in $1884-5$ and subsequent treaties in the $1890s$), well before independence and when Europeans had hardly settled in the regions whose borders were designing, offers a nice (quasi)-experimental setting to address this question.[1] The drawing of political boundaries partitioned in the eve of African independence more than 200 ethnic groups across different countries. Taking advantage of this historical accident, we compare economic performance in regions belonging to the historical homeland of the same ethnic group, but subject to different contemporary national institutions. The regression discontinuity (RD)

[1] There is no ambiguity among African scholars and historians that almost all African borders were artificially drawn. See Asiwaju (1985) for examples and Michalopoulos and Papaioannou (2011) for additional references on the drawing of borders in Africa.

approach allows us to account for differences in geography, the disease environment, and other ecological features. Moreover, by comparing development across border regions that belong to the historical homeland of the same ethnic group (see Figures $2a - 2d$ for examples), allows us to also neutralize biases coming from cultural and other ethnic-specific differences. Our results show that there is no systematic relationship between countrywide differences in institutions and regional economic performance within partitioned ethnicities in Africa.

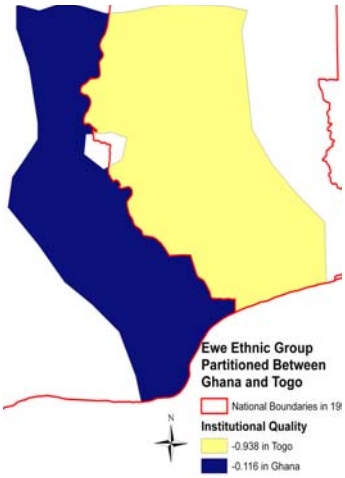


Figure $2a$

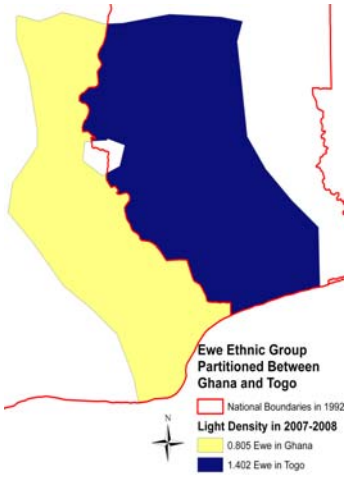


Figure $2b$

Figure $2c$

Figure $2d$

We then turn our focus on the economic impact of pre-colonial ethnic institutions. Our analysis shows that political complexity before the advent of European colonizers correlates significantly with contemporary development, even when we account for national policies and other country-specific features. 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 related to land endowments or the ecology drive the association between pre-colonial ethnic institutional traits and development. Yet the positive association between

historical institutions and luminosity prevails numerous permutations. First, it is robust to an array of controls related to the disease environment, land endowments, and natural resources among others. Second, regressing luminosity on a variety of alternative pre-colonial ethnicity-specific economic and cultural traits reported by Murdock (1967), we find that political centralization is the strongest correlate of regional economic development. Third, we find that the positive correlation between ethnic historical political complexity and regional development obtains across pairs of adjacent ethnic homelands where groups with different pre-colonial institutions reside. Thus, although we do not have random assignment in ethnic institutions and it is therefore hard to establish causality, the results clearly point out that, unlike national institutions, traits manifested in differences in the pre-colonial institutional legacy of each ethnic group matter for contemporary African development.

Related Literature Our research nests and advances over many strands of literature that examine the historical roots of economic development in Africa.

First, an influential body of research asserts that through persistence, the institutions that European powers established in the eve of colonization are the deep roots of contemporary development. While there is ambiguity on the exact mechanisms via which colonization affected African (and more generally non-European) development, there is an agreement that the type of colonization and the identity of the colonizing power had long-lasting effects on institutional quality (e.g. La Porta *et al.* (1997, 1998, 1999); Acemoglu *et al.* (2001, 2002); Feyrer and Sacerdote (2009)). Yet in spite of the ingenious instrumental variables identification schemes employed in the cross-country literature omitted variables and heterogeneity are always major concerns (e.g. Glaeser *et al.* (2004), La Porta *et al.* (2008), Nunn (2009)). The micro approach of our study enables us to overcome problems inherent to the cross-country analysis adding to a vibrant body of research that examines the within-country impact of institutions (e.g. Banerjee and Iyer (2005); Huillery (2009); Iyer (2010); Acemoglu and Dell (2010) and Gennaioli *et al.* (2011)).

Second, our identification scheme on the impact of the national institutions explores discontinuities across the border within partitioned ethnicities. As such our work relates to case studies that examine the effect of national policies at the border. In an early contribution Miguel (2004) compares public policies in health and education across the Kenya-Tanzania border to examine the effect of Tanzanian nation-building policies. Using a similar to ours methodology, Bubb (2009) investigates how differences in de jure property rights between Ghana and Ivory Coast affect development in border areas. He finds that despite large differences in de jure property rights between the two countries, there are no differences in de

facto property rights across the border. Berger (2009) and Arbesu (2011) explore discontinuities across administrative colonial boundaries within Nigeria to study the long-run effects of the different colonial tax systems. The political science literature provides anecdotal evidence based on case studies at the border. For example, Miles (1994) studies the development of the Hausa after their partitioning (at the Niger-Nigeria border), documenting that different French and British policies (mainly on the role of local chiefs) endured after independence and had long-lasting effects. Posner (2005) examines ethnic policies in Zambia and associates them with national representation. Our study, rather than focusing on the effect of national policies across a particular border or within a single partitioned ethnic group, combines satellite light density images with Murdock's ethnolinguistic map to examine in a systematic way the effect of national institutions on development across Africa's partitioned ethnicities.

Third, our findings advance the literature on the role of pre-colonial institutional and cultural features in African development (Fortes and Evans-Pritchard (1940), Schapera (1967), Stevenson (1968), Goody (1971), Bates (1983), Robinson (2002), Boone (2003); see Herbst (2000) for an overview). There were marked differences in institutional and social traits across African regions at the time of colonization (Murdock (1959, 1967); Bates (1983)). There were noteworthy differences on political centralization, land rights, and the power of local chiefs among others. As colonizers did not expand their power in remote areas far from the capital cities and the coastline, such local institutions were preserved and were instrumental during the half century period of colonial rule (roughly $1890-1940$).[2] Along the same lines, Mamdani (1996) argues that the indirect rule of the Europeans, if anything, increased the role of local chiefs during the colonial era. Moreover, several African case studies stress the ongoing crucial role of ethnic institutions and traditions (Englebert (2000); Miguel and Gugerty (2005); Franck and Rainer (2009); Glennerster, Miguel, and Rothenberg (2010)). For example, in an early contribution Douglas (1962) compares development between the neighboring groups of Lele and Bushong in the Democratic Republic of Congo providing evidence that the two ethnicities have different local institutions which are manifested in their levels development. Gennaioli and Rainer (2006, 2007) construct a country-level measure of pre-colonial political centralization and show that it is positively correlated with various contemporary measures of economic and institutional development.

The African historiography has proposed various channels via which ethnic institutions matter today. Herbst (2000) and Boone (2003) argue that in centralized societies there is a high degree of political accountability of local chiefs. Others argue that centralized societies were quicker in adopting growth-enhancing Western technologies and habits, because the colonizers

[2] For example, Acemoglu, Johnson, and Robinson (2003) partly attribute the economic success of Botswana to the limited impact of colonization and the inclusive character of pre-colonial ethnic institutions.

collaborated more strongly with politically and socially complex ethnic groups with strong chiefs (Schapera (1967, 1970)). Herbst (2000) emphasizes the role of ethnic class stratification and political centralization in establishing well-defined and secure land rights (see also Goldstein and Udry (2008)). Furthermore, complex tribal societies with strong political institutions seem to have been more successful in getting concessions both from colonial powers and from national governments after independence. 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.

We improve upon this body of research showing that pre-colonial institutions are positively and systematically linked to regional development even when we control for local geography and country-specific effects. Accounting for common-to-all-ethnicities country factors is central, as Gennaioli and Rainer (2006) show a positive cross-country correlation between pre-colonial centralization and current measures of institutional development. Moreover, controlling for geography is important as studies on African institutional development argue that pre-colonial political centralization was driven by land suitability for agriculture and population density (e.g. Bates (1981); Fenske (2009)). Although our results do not necessarily identify causal effects, they offer support to those emphasizing the importance of pre-colonial ethnicity-specific institutions in current times. In this regard our findings are in line with recent empirical studies showing that historically determined socioeconomic and political factors have persistent effects on comparative development (examples include the forced labor practices of Spanish colonizers in Peru (Dell (2010)); the formation of city-states in Italy during the late period of the Middle Ages (Guiso, Sapienza, and Zingales (2008)); 19th century inequality in Colombia (Acemoglu, Bautista, Querubin, and Robinson (2008)); the type of colonization and early inequality in Brazil (Naritomi, Soares, and Assunção (2009)).

Moreover, the uncovered evidence on the limited role of national institutions on regional development relates to works on state capacity (e.g. Tilly (1985); Migdal (1988); Acemoglu (2005); Besley and Persson (2009, 2010); Acemoglu, Ticchi, and Vindigni (2011)) that emphasize the inability of weak states to broadcast power. Likewise, the finding that the positive correlation between national institutions and regional development weakens in border areas has implications for the literature on optimal state formation (e.g. Alesina and Spolaore (2003); Spolaore and Wacziarg (2005)). Finally, this study contributes to a large body of work on the historical causes of contemporary African development. Nunn (2008) stresses the importance of the slave trade, while Englebert, Tarango, and Carter (2002), Alesina, Easterly, and Matuszeski (2011) and Michalopoulos and Papaioannou (2011) show a significant negative impact

on economic development from the improper colonial border design.

Paper Structure In the next section we first discuss the luminosity data that we use to proxy regional development and present the pre-colonial ethnic institutional measures. We then lay down the empirical design. In section 3 we report our results on the effect of contemporary national institutions on regional development. Section 4 presents our findings on the role of pre-colonial ethnic institutions in shaping regional development. Section 5 concludes.

2 Data and Identification

2.1 Data on partitioning

The starting point in compiling our dataset is George Peter Murdock's (1959) Ethnolinguistic map that portrays the spatial distribution of ethnicities across Africa around the European colonization in the mid/late 19th century. Murdock's map (reproduced in Figure 1*a*) 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 not considered in our analysis. In the empirical analysis we also eliminate the Guanche, a small 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 $100-150$ years. Reassuringly, using individual data from the Afrobarometer Nunn and Wantchekon (2011) show a strong correlation (around 0.62) between the location of the respondents in 2005 and the historical homeland of their ethnicity as identified in Murdock's (1959) map. In the same vein, Glennerster, Miguel, and Rothenberg (2010) document in Sierra Leone that after the massive displacement of the $1991-2002$ civil war there has been a systematic movement of individuals towards the areas of their ethnic group's historical homeland.

We project on top of Murdock's ethnolinguistic map the 2000 Digital Chart of the World (Figure 1*b*) that portrays contemporary national boundaries yielding $1,247$ country-tribe observations. This allows identifying in a systematic way partitioned ethnicities across Africa. Appendix Table 8 reports split groups, defined as groups where at least 10% of their historical homeland belongs to more than one contemporary states. In the empirical analysis we focus on partitions of at least 100 square kilometers as tiny partitions are likely due to the lack of precision in the underlying mapping of ethnicities. Our procedure identifies 526 partitions that belong to 227 ethnic groups. For example, the Maasai were partitioned between Kenya and Tanzania (shares 62% and 38% respectively), the Anyi between Ghana and the Ivory Coast (shares 58% and 42%), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). We also checked whether our codification of partitioned ethnicities is in line with

Asiwaju (1985), who provides the only systematic codification (to our knowledge) of split ethnicities in Africa. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned. Our procedure reveals that the median country in Africa has 43% of its population belonging to partitioned ethnicities. This estimate is of the same order of magnitude to that of Englebert *et al.* (2002) and Alesina *et al.* (2011), who using alternative sources and techniques estimate that on average 40% of the African population comes from partitioned ethnic groups.

2.2 Satellite Light Density at Night

The nature of our study requires detailed data on economic development at the grid level. To the best of our knowledge, geocoded high resolution measures of economic development spanning all Africa are not readily available. To overcome this issue we use satellite data on light density at night to proxy for local economic activity.

The luminosity data come from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) that reports images of the earth at night captured from 20 : 00 to 21 : 30 local time. The satellite detects lights from human settlements, fires, gas flares, heavily lit fishing boats, lightning, and the aurora. The measure is a six-bit digital number (ranging from 0 to 63) calculated for every 30-second area pixel (approximately 1 square kilometer). The resulting annual composite images are created by overlaying all images captured during a calendar year, dropping images where lights are shrouded by cloud cover or overpowered by the aurora or solar glare (near the poles), and removing ephemeral lights like fires, lightning and other noise. The result is a series of global images of time stable night lights. Using these data we construct average light density per square kilometer for 2007 and 2008 at the desired level of aggregation (ethnicity-country) averaging across pixels that fall within the historical homeland of each ethnic group in each country.

This high resolution dataset makes the data uniquely suited to spatial analyses of economic development in Africa for several reasons. First, most African countries have low quality income statistics, even at the national level (for example, the codebook of the Penn World Tables assigns the lowest scores on data quality to all African countries). Second, we lack data on regional income or value added for most African countries. And while there are some regional proxies of poverty and health, these data do not map to our unit of analysis and they are not usually comparable across countries (as survey methods differ). Third, by using light density we also capture the economic activities of the underground economy, which are not reflected in the aggregate statistics. As the share of the shadow economy is high in Africa (La-Porta and Shleifer (2008)), the usage of luminosity data is particularly desirable in our setting.

The use of luminosity data as a proxy for development builds on the recent contribution

by Henderson, Storeygard, and Weil (2011) and others (e.g. Elvidge, Baugh, Kihn, Kroehl, and Davis (1997); Doll, Muller, and Morley (2006); Sutton, Elvidge, and Ghosh (2007)) who show that light density at night is a good proxy of economic activity. These works 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 Elvidge, Baugh, Kihn, Kroehl, and Davis (1997) and Min (2008)). In ongoing work Pinkovskiy (2011) documents large and statistically significant jumps in GDP per capita as forecasted by luminosity across borders (both globally and across African countries); he further shows that these discontinuities correlate with differences in countrywide GDP per capita. Even Chen and Nordhaus (2010) 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.

Satellite data on lights are subject to saturation and blooming. Saturation, which occurs at a level of light similar to that in the urban centers of rich countries, results in top-coded values. Blooming occurs as 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) that occurs due to the diffusion of lights is not a major problem. Additionally, to account for overglow over water, area under water is a standard geographic control in our regressions. Finally, variation in satellite light density across countries and regions may arise because of: (i) cultural differences in use of lights and geographic differences; (ii) the composition of income between consumption and investment; (iii) the division of economic activity between night and day, and (iv) the satellite used. By including country and ethnicity fixed effects, and conditioning on a rich set of climatic and geographic control variables, we account for these problems.

2.2.1 Satellite Light and Development

In spite of the works showing a strong correlation between luminosity and GDP per capita we performed cross-validation exercises investigating the relationship between luminosity and various economic indicators across African countries, regions, and areas across the border.

Luminosity and GDP per Capita across Countries We start by examining whether luminosity correlates with development across African countries. Figure 3*a* illustrates the unconditional correlation between log light density and log GDP per capita in 2000. There is a clear positive relationship. 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. Figure 3*b* shows the relationship between log GDP per capita and log light density partialling out the effect of log population density. The relationship between log light density and log GDP per capita is now stronger (the coefficient increases from 0.31 to 0.47 and the t-stat jumps to 10).

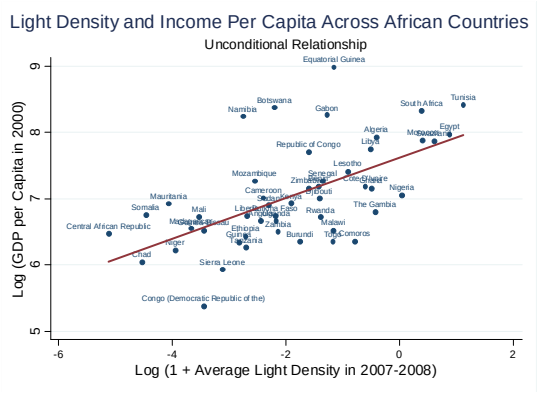


Figure 3*a*

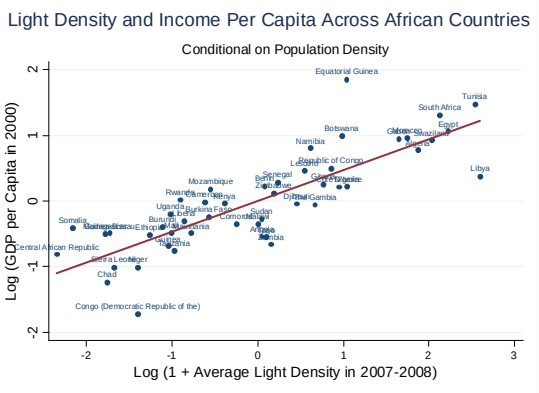


Figure 3*b*

Luminosity and Infant Mortality across Administrative Regions

We also examine the correlation between satellite light density and infant mortality, as an alternative proxy of development, at the regional level, using comparable across African countries data from the Center for International Earth Science Information Network (CIESIN) of Columbia University's Earth Institute.[3] Figures 4*a* and 4*b* 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.

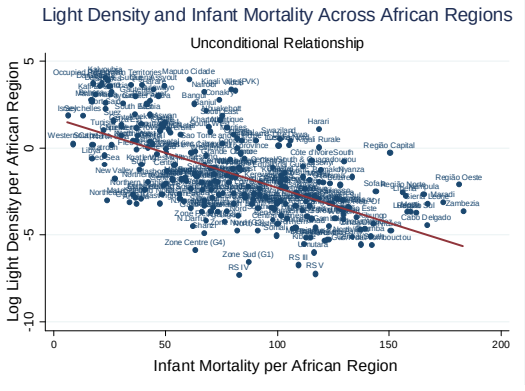


Figure 4*a*

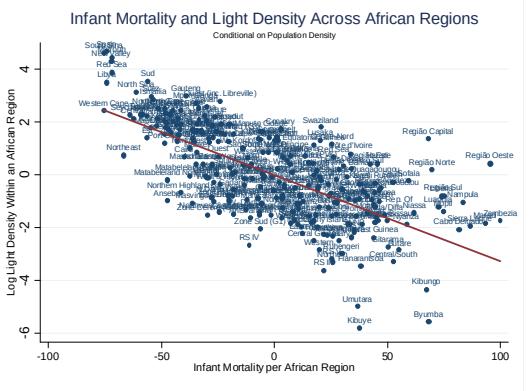


Figure 4*b*

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

Luminosity and Household Wealth across Clusters (DHS) within Countries

To further illustrate how well satellite light density reflects comparative economic development at finer levels of aggregation, we used geocoded data from the Demographic and Health Surveys (DHS). Conducting household questionnaires on a variety of economic indicators the DHS team in each country produces a composite wealth index.[4] We calculated average light density in the surrounding area of each DHS cluster using a radius of $10km$ and we then examined the association between luminosity and the average wealth index across all households interviewed in the cluster. Figures $5a - 5d$ plot the correlation between household wealth and light density in four large countries from different parts of Africa (namely Tanzania, the Democratic Republic of Congo (Zaire), Zimbabwe, and Nigeria). The correlation is strong, around 0.80, further showing that luminosity is a robust predictor of economic development. Note that the scatter plots partial out the effect of population density so the observed relationship does not mask underlying differences in urbanization rates across the DHS clusters.

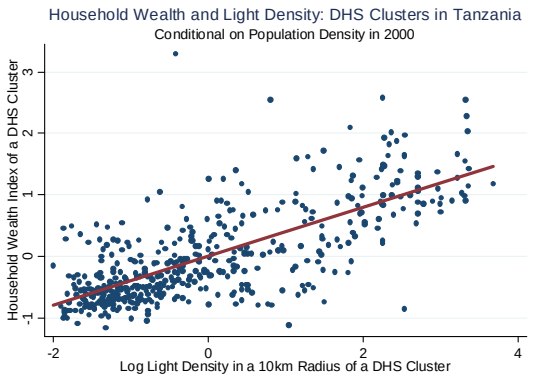


Figure $5a$

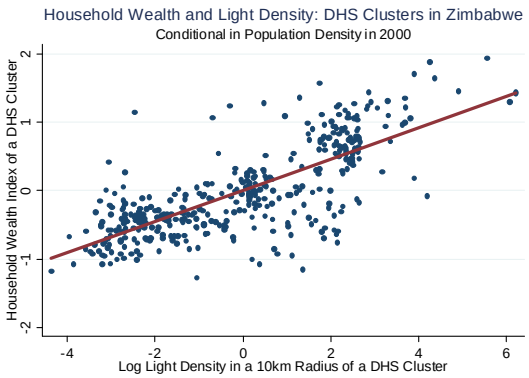


Figure $5b$

[4]The DHS wealth index is composed taking into account consumer durables, electricity, toilet facilities, source of drinking water, dwelling characteristics and other country-specific attributes like whether there is a domestic servant, for example. The measure is derived using principal component analysis to assign indicator weights resulting in a composite index which is standardized for each country. DHS data are available at: http://www.measuredhs.com

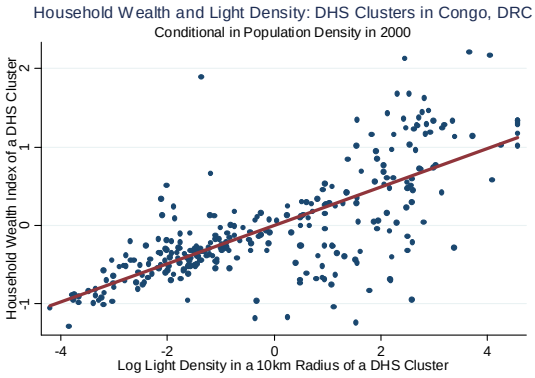


Figure 5*c*

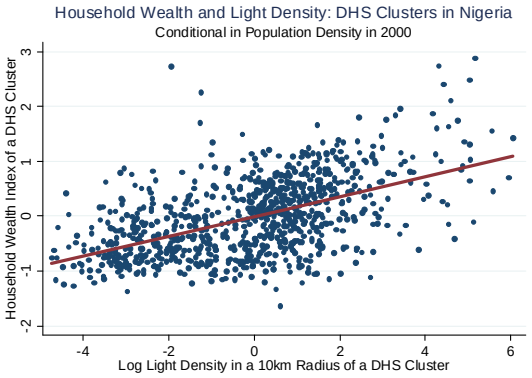


Figure 5*d*

Luminosity and Development within Partitioned Ethnicities (Afrobarometer)

Although the strong correlation between luminosity and the DHS wealth index reveals that satellite images of light density reflect economic activity at a fine level, one may still wonder whether luminosity captures differences in development across the national border. We thus used data on access to clean water and education from the Afrobarometer to explore whether within-ethnicity (across the border) differences in luminosity correlate with differences in these two development proxies. 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.[5] Following Nunn and Wantchekon (2011) for each country we assign the current location of the respondents to the respective historical ethnic homeland and we then estimate the average of the responses across individuals at the ethnicity-country level. Figure 6*a* shows the relationship between log light density and average education of the respondents partialling out differences in population density and ethnicity fixed effects. Partitions of the same ethnic homeland across the border exhibiting higher light density are inhabited by more educated people. Similarly, Figure 6*b* depicts the relationship between log light density and access to piped water.

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

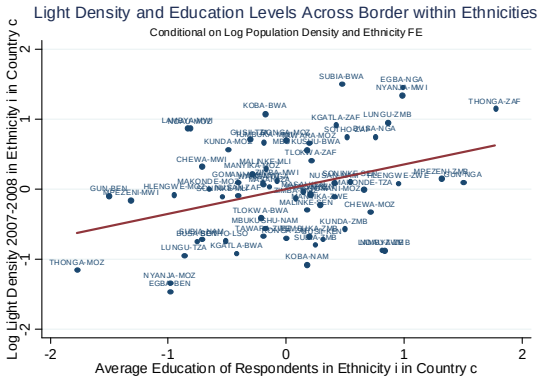


Figure 6*a*

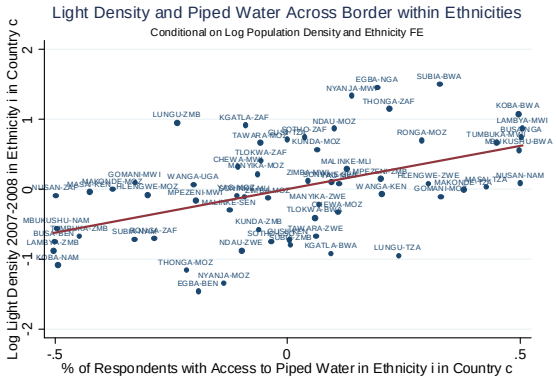


Figure 6*b*

The uncovered robust association between luminosity and various development proxies across African countries, administrative regions and villages within countries, as well as areas across national boundaries belonging to the historical homeland of the same ethnicity, suggests that in absence of alternative comprehensive measures of regional economic performance, light density is an informative proxy of local economic conditions for African regions.

2.3 Ethnic Institutional Traits

In work following the mapping of African ethnicities, Murdock (1967) produced an Ethnographic Atlas (published 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 835 African ethnicities of Murdock's Map of 1959 to the ethnic groups in his Ethnographic Atlas of 1967. As the two sources do not always use the same name for identifying ethnic groups we employed several sources and the updated version of Murdock's Atlas produced by J. Patrick Gray to match a total of 534 ethnicities.

Following Gennaioli and Rainer (2006, 2007) we proxy pre-colonial political institutions using Murdock's (1967) "Jurisdictional Hierarchy beyond the Local Community Level" index. This is an ordered variable, ranging from 0 to 4, that describes the number of political jurisdictional hierarchies above the local (usually village) level for each society. A zero score indicates stateless societies "*lacking any form of centralized political organization*". Scores 1 and 2 designate petty and larger-paramount chiefdoms, while 3 and 4 indicate groups that were part of large organized states. Murdock (1967) explicitly excludes colonial regimes (such as protectorates) and attempts to capture political complexity before Europeans started the settlement of Africa. Figure 7*a* illustrates the significant heterogeneity in political centralization 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 small chiefdoms are the Mende in Sierra Leone and the Ibo of Nigeria. The Mbundu in Angola and the Zerma in Niger were part of large paramount chiefdoms, while the Yoruba in Nigeria and the Mossi in Burkina Faso are societies that were parts of large states before colonization. The Bubi in Equatorial Guinea and the Beduin Arabs in Morocco and Tunisia are classified as having been part of large complex states.

Cross-validation of Murdock's Jurisdictional Hierarchy Index We also performed a cross-validation of Murdock's data on the degree of political centralization of ethnicities going over numerous works in African historiography and political science. Our reading of the literature suggests that the jurisdictional hierarchy index 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 all major pre-colonial African states. For example, the Ankale and the Buganda, which were the central ethnic groups in the strong states of Eastern Africa (in contemporary Uganda) get a score of 3 in the jurisdictional hierarchy index. The same applies for other ethnic groups that were part of large empires, such as the Zulu and the Swazi in South Africa, the Yoruba in Nigeria, and the Shongai in Mali (see Fortes and Evans-Pritchard (1940); Goody (1971).

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 use them as examples of stateless 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 seems also to identify quite accurately societies with intermediate levels of political centralization (paramount chiefdoms and small states). The Ashanti that were part of a loose confederation get a score of 2; likewise the Nupe in Northern Nigeria, the Bemba in Zambia, and the Ngwato in Botswana which were also part of small states get a score of 2 (see Fortes and Evans-Pritchard (1940)).

2.4 National Institutions

For national institutions we rely on the World Bank's Governance Matters Database (Kaufmann, Kraay, and Mastruzzi (2005)). The World Bank assembles numerous de facto institutional quality measures (originally compiled by various non-governmental organizations and

risk assessment agencies) and aggregates them into six categories via a principal component analysis. For our benchmark estimates we use the rule of law index that reflects the effectiveness of the judiciary and the quality of property rights protection. As many studies on African development focus on graft, we also report results using the control of corruption index. Both variables have a theoretical minimum and maximum of -2.5 and $+2.5$, respectively, with higher values indicating better functioning institutions and less corruption. Results are qualitatively similar if we use alternative measures of national institutions, like the ICRG risk of expropriation, the Polity's executive constraints index, or measures from World Bank's Doing Business around the World Project. In our sample the countries with the lowest rule of law are Somalia (-1.91) and the Democratic Republic of Congo (-1.84), while Namibia (0.64) and Botswana (0.71) are the most institutionally developed countries (see Figure 7*b*).

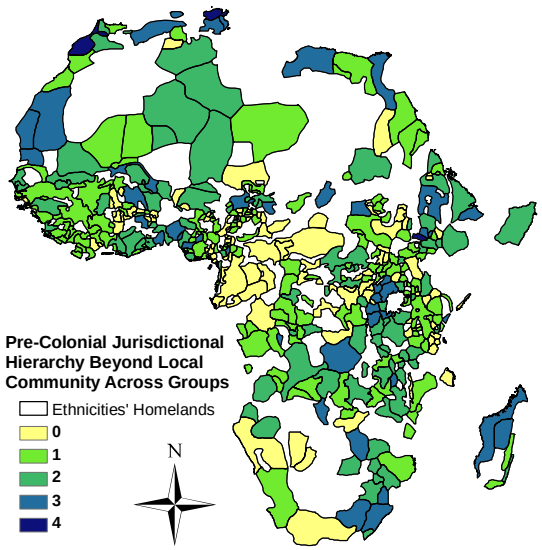


Figure 7*a*

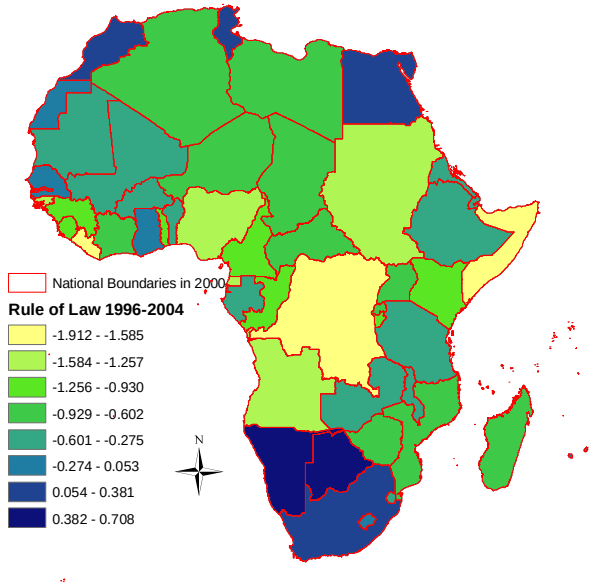


Figure 7*b*

2.5 General Empirical Framework

Our analysis on the relationship between contemporary national and pre-colonial ethnic institutions and regional development is based on variants of the following specification:

$$y_{i,c} = a_0 + \gamma IQL_c + \delta LOCINST_i + X'_{i,c}\Phi + \lambda PD_{i,c} + [a_c + a_i] + \varepsilon_{i,c} \quad (1)$$

The dependent variable, $y_{i,c}$, is the level of local economic activity in the historical homeland of ethnic group i in country c, as proxied by light density at night. IQL_c denotes institutional quality of country c (as reflected in the rule of law and the control for corruption

measures). For ethnicities that fall into more than one country, each area of the partitioned group is assigned to the corresponding country c. For example, regional light density in the part of the Ewe in Ghana is matched to the institutional quality of Ghana, while the adjacent region of the Ewe in Togo is assigned the value of Togo. $LOCINST_i$ denotes local ethnic institutions as reflected in the degree of jurisdictional hierarchy beyond the local level.

Since the correlation between luminosity and proxies of development strengthens when we condition for urbanization in many specifications we control for log population density ($PD_{i,c}$) though population density is likely endogenous to national or/and ethnic institutional development. Moreover, when we control for population density the regression estimates capture the relationship between institutions and economic development per capita.

A potential merit of our regional focus is that we can account properly for local geography and other factors (captured in vector $X_{i,c}$). This is non-trivial as there is a fierce debate in the literature on the institutional origins of development on whether the strong correlation between institutional and economic development is driven by geographical features and the disease environment (e.g. Gallup, Sachs, and Mellinger (1999), Easterly and Levine (2003)). In many specifications we include a rich set of geographic controls, reflecting land endowments (elevation, area under water), ecological features (malaria stability index, land suitability for agriculture), and natural resources (diamond mines and petroleum fields).

There are several studies that suggest the inclusion of these variables. First, Nunn and Puga (2011) show that elevation and terrain ruggedness have affected African development both via goods and slave trades. Second, the inclusion of surface under water accounts for blooming in the light image 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 large negative effect of malaria on development. Fourth, there is a vast literature linking natural resources like oil and diamonds to development and civil conflict (e.g. Ross (2006)). Fifth, Michalopoulos (2011) shows that differences in land suitability and elevation across regions lead to the formation of ethnic groups, whereas Fenske (2009) and Ashraf and Galor (2011) show that land quality is strongly correlated with pre-colonial population densities. We also control for the distance of the centroid of each ethnic group i in country c from the capital city, the national border, and the coast. As most contemporary capital cities in Africa were established by Europeans during the colonial period, the coefficient on distance from the capital may reflect the impact of colonization and the limited penetration of national institutions due to the poor infrastructure (we formally explore this possibility below). Distance to the border captures the potentially lower level of development in border areas. Distance to the sea coast captures the effect of trade, but to some extent also the

penetration of colonization. This is because during the colonial era (and the slave trades) Europeans mainly settled in coastal areas.

In our analysis on the impact of national institutions we include ethnicity fixed effects (a_i) to effectively control for cultural and unobserved geographical features of ethnic homelands, whereas in Section 4 where we examine the role of ethnic institutions we include country fixed effects (a_c) to account for national policies and institutions, as well as other countrywide factors.

2.6 Technical Remarks

Estimation First, the distribution of luminosity is not normal, as (i) a significant fraction (around 30%) of the observations takes on the value of zero[6] and (ii) we have a few extreme observations in the right tail of the distribution (see Appendix Figure 1a). While the mean of satellite light density is 0.364 the median is more than twenty times smaller, 0.017. This occurs because there are a few ethnic areas where light density is extremely high. For example, we have 13 observations (1%) where light density exceeds 6.4 and 26 observations (2%) where light density exceeds 4.67.

To account for both issues we use as the dependent variable the log of light density adding a small number ($(y_{i,c} \equiv \ln(0.01 + LightDensity_{i,c})$, see Appendix Figure 1$b$).[7] This transformation assures 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 (see Appendix Figure 1c). Looking at the "intensive margin" also guarantees that we investigate the role of (national and ethnic) institutions in explaining variation in economic performance across densely populated regions displaying non-trivial economic activity (non-lit areas have a median population density of 8.82 people per square kilometer whereas regions with positive light density have a median of 27.63).[8]

Inference Second, in all specifications we employ the approach of Cameron, Gelbach, and Miller (2006) and cluster standard errors both at the country and at the ethnic-family level (Murdock assigns the 835 groups into 96 ethnolinguistic clusters/families), as this accounts for two main concerns related to non-adjusted standard errors. First, within each country

[6] A zero level of light density occurs either because the area is extremely sparsely populated without any electricity or because the satellite sensors cannot capture dimly lit areas.

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

[8] The results are similar ignoring the top 1%, 2% or 5% of the luminosity data. Moreover, in the previous version of the paper we reported Poisson ML specifications finding analogous results. We also estimated OLS models in levels, Tobit specifications that account for censoring in the dependent variable and also performed least absolute deviation (median) regressions using all data to account for outliers. These alternative estimation techniques deliver quite similar results.

we have several ethnicities where the country-level rule of law and the control-of-corruption measures take the same value and thus clustering at the country-level is required (Moulton (1986)). Likewise, partitioned ethnicities appear more than one time in our sample and thus clustering at the ethnic family accounts for unobserved features within each ethnolinguistic family.[9] As we report specifications using the ethnicity-level indicators that exhibit within-ethnic-family correlation, it is appropriate to also cluster standard errors at the ethnic-family level. Second, the multi-way clustering method accounts for arbitrary residual correlation within both dimensions and thus accounts for spatial correlation (Cameron, Gelbach, and Miller (2006) 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 very similar standard errors; and if anything the two-way clustering produces somewhat larger standard errors yielding the most conservative inference. Moreover, as in many specifications we include country or ethnicity fixed effects this soaks up further the spatial correlation at each dimension.

2.7 Preliminary Evidence

Table 1 reports summary statistics for the variables employed in the empirical analysis. Table 2 reports cross-sectional LS specifications that associate regional development with contemporary national and pre-colonial ethnic institutions. Below the estimates we report both double-clustered (in parentheses) and Conley's (in brackets) standard errors.[10] Column (1) shows that there is a positive and significant correlation between the rule of law index and regional development. In column (2) we add population density, whereas in column (3) we control for distance to the capital city, distance to the border, and distance to the coast. While all distance terms enter with significant coefficients, the estimate on rule of law retains its economic and statistical significance. In (4) we control for population density, location, and a rich set of geographic controls. Conditioning on geography reduces the magnitude of the coefficient but the estimate retains significance at conventional levels.[11] Overall, the correlations in (1)-(4) echo

[9] Clustering at the ethnicity level rather than at the ethnic family level produces similar standard errors. We prefer to cluster at the broader ethnic-family level, because the consistency of the standard errors improves with the number of within-cluster observations (Cameron, *et al.* (2006)).

[10] Conley's correction method requires a cut-off distance beyond which the spatial correlation is assumed to be zero; we experimented with various cutoff values between 100km and 3000km choosing the cutoff of 2000km which delivers the largest in magnitude standard errors.

[11] Land suitability for agriculture, which reflects climatic (temperature and precipitation) and soil conditions, enters most models with a positive and significant estimate. The malaria stability index enters in all specifications with a statistically negative estimate. The coefficient on land area under water is positive and in many specifications statistically significant. Elevation enters with a negative estimate which is significant in some models. The petroleum dummy enters always with a positive and significant coefficient, most likely because the satellite captures fires from oil facilities. The diamond dummy enters in most specifications with a negative coefficient.

the findings of cross-country works; although the association between institutional quality and development weakens somewhat when one accounts for geography, it remains highly significant.

In columns (5) to (8) we associate regional development with ethnic pre-colonial political institutions. Column (5) reports unconditional estimates. The coefficient on jurisdictional hierarchy index is positive and significant at the 99% confidence level. Controlling for population density, location, and the rich set of geographic controls (in columns (6)-(8)) has a noticeable effect on the coefficient which nevertheless remains at least two standard deviations above zero in all permutations.

In columns (9)-(12) we regress regional light density on both national and ethnic institutions. Given the positive correlation (0.16) between rule of law and jurisdictional hierarchy, it is useful to investigate the stability of the previous results. Column (9) introduces both the rule of law index and the jurisdictional hierarchy measure. The unconditional estimate of rule of law in the sample of 680 ethnicity-country observations is 0.14 (specification not shown). Once we control for the degree of jurisdictional hierarchy the estimate on rule of law retains its significance and falls by 15%. Likewise, the coefficient on jurisdictional hierarchy is positive and highly significant, though its magnitude is somewhat smaller compared to the analogous specification in (5).[12] A similar pattern obtains when we control for location (distance to the border, the sea coast, and the capital city), population density, and the set of geographic-ecological controls (in (10)-(12)).

The coefficient in column (12) implies that a one point increase in the rule of law index (moving approximately from the institutional quality level of Angola to that of Gabon) is associated with a 73% increase in regional luminosity. Turning now to the magnitude of pre-colonial institutions, the most conservative LS estimate (0.17) implies that regional development increases by approximately 50% as one moves from areas where stateless societies reside to regions with ethnic groups featuring centralized pre-colonial institutions (i.e. have a jurisdictional hierarchy index equal to 3). The preliminary results in Table 2 are informative about the broad data patterns. Yet these estimates do not identify the one-way effect of neither contemporary national institutions nor ethnic historical institutional traits on regional development. This is the task of the next two sections.

3 National Institutions and Regional Development

Identifying the causal impact of contemporary institutions on regional development is a demanding task, because, among other challenges, there are rarely otherwise identical cultures

[12] Compared to the specifications in columns (5)-(8), we lose three observations when we include the rule of law index, because we lack data on Western Sahara (the results are almost identical if we assign the rule of law index of Morocco to the Western-Saharan ethnic regions).

exposed to different institutional settings. The arbitrary design of borders in Africa offers an ideal setting to isolate the effect of nationwide institutions from cultural traits and ethnic institutions.

There is significant variation in both the rule of law and luminosity across African borders. Sharp border discontinuities in rule of law appear in several parts of Africa. For example, in the Botswana and Zimbabwe border (where the Hiechware, the Subia, and the Tlokwa are partitioned); across the Namibia and Angola border (where the Ambo are split); between Kenya and Somalia (where the Bararetta group resides); or between Gabon and Congo (where the Duma live). Likewise, there are changes in luminosity across the border within the historical homeland of the same partitioned ethnic group (see Figure 8). On the one hand, in around 30% of the sample there are no differences in light density across the border within ethnic groups; on the other hand in about 40% of the partitions there are more than one log point differences in luminosity. For example, large jumps in luminosity appear in the Egypt-Sudan border (where the Ababda and the Barabra groups are partitioned), in the border between Ghana and the Ivory Coast (where the Assini reside), and between Benin and Togo (where the Popo are split).

To the extent that national institutions affect regional development, one should expect to find that the part of the ethnic group that belongs to the high institutional quality country would outperform economically the adjacent region of the same historical homeland that falls into the country with the relatively worse national institutions.

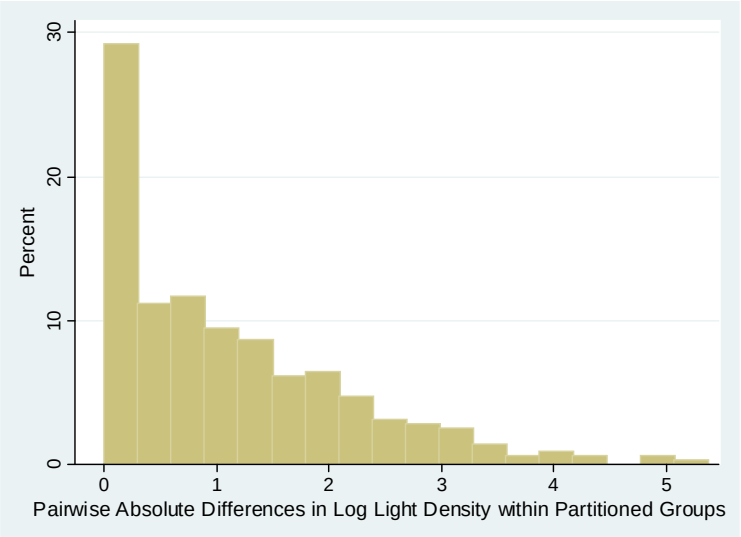


Figure 8

3.1 Within-Ethnicity Results

In this section we examine the cross-sectional and the within-ethnicity correlation between national institutions and luminosity in the group of partitioned ethnicities. In the odd-numbered columns of Table 3 we report cross-sectional estimates, while in the even-numbered specifications we add ethnicity fixed effects to account for local geography and culture. The cross-sectional estimates in columns (1) and (3) echo the findings of Table 2. Along border regions partitions of ethnic groups that belong to countries with higher institutional quality display systematically higher levels of development. Yet when we solely exploit the within-ethnicity variation (i.e. estimating equation (1) with a_i), the coefficients on rule of law and control of corruption drop sizably and become statistically indistinguishable from zero. The insignificance is not driven by a decrease in the precision of the estimated coefficients since the standard errors remain largely unchanged. In both permutations (as well in most subsequent specifications), two-standard-error bands in the within-ethnicity estimates exclude the cross-sectional ones.

In columns (5)-(8) we repeat estimation across lit partitions. The cross-sectional estimates show that across lit border areas there is a strong correlation between institutional quality and regional development. Yet once we include ethnicity fixed effects the coefficients on rule of law and control of corruption drop by more than a half and turn insignificant. Figures 9*a* and 9*b* below illustrate the lack of a systematic within-ethnicity correlation between light density and institutional quality at the national level.

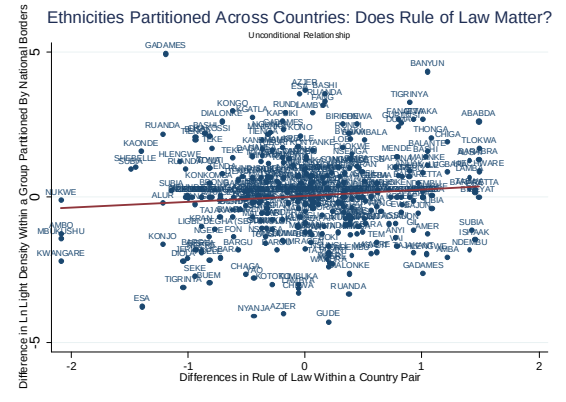


Figure 9*a*

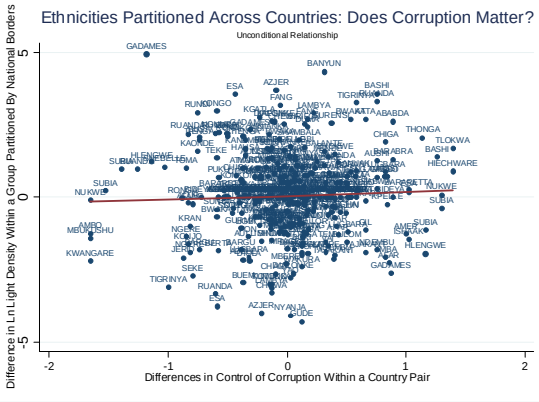


Figure 9*b*

In light of the results reported in Table 2 that the correlation between rule of law and light density weakens but remains significant when one controls for ethnic institutions, the findings in Table 3 suggest that the ethnicity fixed effects capture on the top of the measured tribal institutions other cultural and unobserved ethnic features. To gauge, for example, how much of the decline in the fixed-effects estimate of rule of law is due to the inclusion of the

pre-colonial jurisdictional hierarchy consider the following: In the specification of column (1) in Table 3 when we restrict the estimation on partitioned groups for which information on tribal institutions is available the OLS estimate on rule of law is 0.72. If we introduce ethnicity fixed effects (Table 3-column 2), then the coefficient drops by 54% (0.33) and becomes insignificant. If we add the jurisdictional hierarchy index instead, the estimate on rule of law drops by 14% to 0.61. So, from the 54% drop in the magnitude of rule of law that is due to ethnicity fixed effects a quarter of the decline can be ascribed to variation in pre-colonial ethnic institutions.

3.2 Validity of the Regression Discontinuity Design

The partitioning of several African groups in different countries enables us to investigate the role of national institutions in a regression discontinuity framework by exploiting changes in the quality of national institutions at the border. Before presenting the results it is necessary to check the validity of the regression discontinuity design. The RD design requires that all relevant factors besides the treatment -national institutions in our application- vary smoothly at the border. So a concern is that the geography or historical development of ethnic partitions in the relatively low institutional quality countries are systematically different from the partitions of the same group falling into the relatively high institutional quality ones. In this case the two (or more) partitions of each ethnicity might not be appropriate counterfactuals.

To investigate whether differences in institutional quality across the border correlate with differences in various observable characteristics we estimate simple ethnicity fixed-effects regressions associating variables reflecting geography, the disease environment, natural resources, urbanization at independence, and location with rule of law and control of corruption in the ethnic groups partitioned by the national border.[13]

Table 4 reports the coefficient on the rule of law and the control of corruption index with the double-clustered (at the country and the ethnic family-level) standard error for each specification. The results support our identification design. First, differences in geography, the disease environment, and natural resources across the border are small and most importantly not systematically linked to differences in national institutions. Second, although estimates of 1960 population density have to be interpreted with caution, there is no statistical difference in population density around independence across adjacent partitions of the same ethnic group. In a Malthusian regime where richer areas are more densely populated, this implies that there

[13] We thank an anonymous referee for suggesting this implicit test for our RD design. We also investigated whether partitions of ethnic groups differ systematically from non-partitioned ethnicities with respect to their location and geographic traits. Except for the fact that the partitions of ethnic groups are on average 42.5 kilometers further from the capital cities and 92.50 kilometers closer to the borders there are no other significant differences regarding geography, ecology, and population density in 1960. Overall, ethnic partitions are similar to ethnic homelands unaffected by the border design along a host of observable characteristics.

were no systematic differences in economic performance within split ethnicities whose partitions following independence would come to be subject to different national level institutions. Third, the only covariate that is significantly correlated with the treatment is distance from the capital city. Partitions falling in the relatively high rule of law countries are closer to the capital city of that country. This correlation is driven by Sudan and the Democratic Republic of Congo, two of the largest in terms of size countries in Africa that score low in institutional development.

3.3 Regression Discontinuity

In this subsection we report the results from the regression discontinuity approach that identifies the effect of national institutions at the border. In our context the assignment (running/forcing) variable is the distance from the centroid of an ethnic area to the national border ($BDIST_{i,c}$). We thus estimate specifications adding a cubic RD-polynomial in distance to the border ($f(DIST_{i,c})$); the polynomial takes on positive values for the partition that falls in the relatively more institutionally developed country and negative values for the partition of the same ethnic group that falls in the relatively less institutionally developed one. We also add interaction terms between the proxy measures of national institutions (IQL) and the polynomial terms. Rewriting equation (1), our specification reads:

$$y_{i,c} = a_0 + \gamma IQL_{i,c} + f(BDIST_{i,c}) + f(BDIST_{i,c}) * IQL_{i,c} + \lambda PD_{i,c} + X'_{i,c}\Gamma + a_i + \varepsilon_{i,c} \quad (2)$$

This RD-type design (see Imbens and Lemieux (2008); Lee and Lemieux (2010)) exploits the discontinuity in the quality of national institutions at the border to identify institutions' local average treatment effect (LATE).[14] The underlying idea is that by comparing regional development in the historical homeland of the same ethnicity exactly at the border, where only the quality of national institutions differs, one accounts for all characteristics that may affect regional development. The results from Table 4 suggest that there are no differences in geography, ecology, and natural resources within ethnic partitions. Previous research has employed variants of the regression equation (2) using different control functions of the running variable (distance to the border in our application) and limiting estimation close to the discontinuity using different neighborhoods (bandwidth). At the one end, some researchers use all observations, both far and close to the discontinuity, as this approach maximizes the sample and is more efficient (e.g. Lee, Moretti, and Butler (2004)). On the other end, others estimate local

[14] Strictly speaking we do not have a discrete "treatment" group since the national institutions variable is continuous. We also estimated specifications defining indicator variables that take on the value 1 in the country where the rule of law and the control of corruption measures are relatively higher. The results (not reported for brevity) are similar.

linear regressions limiting their analysis to an area close to the discontinuity (e.g. Angrist and Lavy (2001)). Table 5 presents estimates with both approaches. Panel A reports estimates with a global polynomial control function using all pixels belonging to partitioned ethnicities. Panel B reports local linear regression estimates narrowing the analysis in areas close to the border using a cutoff of $25km$ and $50km$ on each side of the border.

Global Polynomial Function Approach In Table 5 Panel A the sample size is 454 partitions since we focus on the two major partitions of each of the 227 split groups. The cross-sectional estimates in columns (1) and (5) on the rule of law and the control for corruption indicators are positive and more than 3 standard errors larger than zero. The respective coefficients are comparable to the analogous estimates in Table 3 where we did not include the RD polynomial in distance to the border and its interactions with the national institutions. Nevertheless, once we include ethnicity fixed effects (in columns (2) and (6)) the estimates drop and become statistically indistinguishable from zero. The same pattern applies when we control for geography, natural resources, location, and the ecology of the terrain (in (3), (4), (7), and (8)). Across all permutations the within-ethnicity estimate on national institutions is insignificant and in all but one instance the two-standard-error band excludes the between-ethnicity estimates.

Local Linear Regressions In Panel B we restrict estimation in the ethnic areas close to the border using either a $25km$ buffer zone (in columns (3)-(4) and (7)-(8)) or a $50km$ buffer zone (in (1)-(2) and (5)-(6)) on each side of the border. When we limit our attention to these regions the polynomial terms and their interactions with the national institutions measures are insignificant; thus we omit them from the specification noticing that this has no effect on the results.

The local-regression estimates yield results similar to those in Tables 3 and $5A$ where we used all pixels of the ethnic homelands partitioned by national boundaries. While in the cross section there is a strong positive association between national institutions and regional development, once we account for unobserved cultural and geographic features focusing on the sides of the border populated by the same ethnic group, the within-ethnicity coefficients of rule of law and control for corruption become statistically indistinguishable from zero. For example, the cross-sectional estimate in the rule of law index is around 0.65 when we focus on the areas within $25km$ and a $50km$ from the borders. Its magnitude drops by more than two thirds once we add ethnicity fixed effects and becomes insignificant casting doubt on the causal interpretation of the simple cross-sectional association between national institutional quality and regional development. Similar to our previous estimates the insignificance is not

driven by an increase in the standard errors. In all permutations two-standard-error bands in the within-ethnicity estimates reported in the even-numbered columns exclude the analogous cross-sectional ones in the odd-numbered columns.

3.4 Sensitivity Analysis

3.4.1 Migration

African scholars and anecdotal evidence suggest that national boundaries across Africa are poorly enforced; this is due to poor demarcation, geographic conditions (desert areas in the North, rainforest in Central Africa), and lack of border patrolling. This poses a threat to our identification strategy. If people migrate to take advantage of higher incomes in regions with higher levels of institutional quality, mobility across national boundaries may attenuate income differences across the border. Furthermore, to the extent that mobility barriers are lower within the same ethnicity compared to moving across different ethnic homelands, once we focus within the same ethnic group, then the ease of mobility would further attenuate any differences in regional development caused by changes in the national level institutions. This scenario predicts that if institutions matter, then as a result of the migration towards the partitions located in the high institutional quality country, population density should be systematically higher in the latter. Alternatively it may be the case that population in the low institutional quality country clusters near the border to migrate to the high institutional quality country every day to work. Although we have controlled for population density, we explore in detail the issue of migration with two different approaches.

First, in Table 6*A* we directly examine the effect of national institutions on population density using data from the United Nations Environmental Programme.[15] There seems to be no systematic association between the quality of national institutions and log population density in border areas where partitioned ethnicities reside; this applies both to the cross-sectional and the within-ethnicity estimation, suggesting that the insignificant within-ethnicity relationship between country-level institutions and regional development is not driven by migration. Note also that the coefficient is small and even changes sign across permutations. Moreover, to the extent that population density reflects regional development, the insignificant within-ethnicity

[15]The UNEP dataset imputes population at the grid level using information on roads, railroads and navigable rivers as well as information on urban centers (see details at: http://na.unep.net/siouxfalls/globalpop/africa/part2.html#construct). To further assuage concerns related to measurement error in the population estimates, we also experimented with an alternative measure from the Gridded Population of the World (http://sedac.ciesin.columbia.edu/gpw/) which does not reallocate the population, finding again no association between national institutional quality and population density. Concerns of measurement error in the population density are further toned down because in line with a large body of work in African historiography we do find a positive association between pre-colonial ethnic institutions and contemporary population density (see Supplementary Appendix Table 7).

coefficient on rule of law and control of corruption provides additional evidence that national level institutions are not systematically related to regional economic performance.

Second, we re-estimated our benchmark specifications associating light density with national institutions across partitioned ethnic homelands excluding areas very close to the border. Doing so we account for potential temporary migration flows across the border as well as for local trade around border regions. By excluding pixels close to the national border we also account for potential stealing and bleeding (and/or blooming) in the luminosity data. Table $6B$ reports cross-sectional (in odd-numbered columns) and within-ethnicity specifications (in even-numbered columns) associating satellite light density and national institutions across partitioned ethnic groups excluding the areas close to the border. In columns (1), (2), (5), and (6) we exclude areas within 25 kilometers from each side of the national border (total 50 kilometers), while in columns (3), (4), (7), and (8) we exclude from each side regions within 50 kilometers of the national boundary (total 100 kilometers). The results are similar to the estimates in Tables 3 and 5. While in the cross section there is a strong positive correlation between institutional quality and luminosity, the coefficient drops sizably and becomes statistically indistinguishable from zero when we add ethnicity fixed effects.

3.4.2 Penetration of National Institutions

African historians and political scientists (e.g. Herbst (2000)) have long argued that the European's presence in Africa with some exceptions was limited to the coastline and the capital cities. Hence, colonial institutional arrangements, reflected through persistence on today's institutional quality, would have limited reach far from the capital cities. Along the same lines, several scholars have argued that due to the lack of the necessary infrastructure (roads, transportation system) and limited state capacity, nationwide institutions have minimal impact far from the capital cities.[16] We explore this hypothesis searching for potential heterogeneous effects of national institutions in ethnic areas far and close from the capital cities.

Table 7 reports estimates with the two RD approaches, the global polynomial control function approach in columns (1) and (2) and the local linear regression method in columns (3) and (4). In Panel A we focus on two-way split ethnic groups which have both their partitions close to or far from the capital city (using as a threshold the median distance from the respective capital city). The cross-sectional correlation between rule of law and luminosity is twice as

[16] Herbst (2000; pp. 16) notes that "*rather systematically, Europeans created capitals that moved power toward the ocean and away from the interior centers of power that Africans had slowly created*". Herbst lists many examples where colonizers decided to ignore local needs and established capital cities outside preexisting polities. As extreme examples he lists Mauritania and Bechunaland (Botswana) that were ruled during colonization by capitals outside their nominal territories (Saint-Louis and Mafeking, respectively). Moving the location of the capital was a key question for African leaders at independence. Yet with a few exceptions (Tanzania, Malawi, and Nigeria), most countries did not relocate the capital city.

strong for partitions close to the capital vis a vis partitions far from the capital. When we add ethnicity fixed effects in columns (2) and (4) the coefficients of interest drop by more than 50% and become insignificant at conventional levels. Similar pattern obtains when we use all partitioned groups in Panel B. In columns (1) and (2) we focus on the two major partitions of all split groups and in columns (3) and (4) we use all 526 partitions. The cross-sectional coefficient on rule of law is positive and significant whereas the interaction between capital distance and rule of law is negative. For example, combining the coefficients of column (1) of Panel B we get that for ethnic partitions that are further than $1,070$ kilometers from the capital city (this is the case for 10% of the partitions) the cross-sectional correlation is statistically insignificant.

The evidence is consistent with works from the African historiography that stress the limited penetration of national institutions far from the capital cities and it complements the literature on optimal country size (e.g. Alesina and Spolaore (2003)) and state capacity (e.g. Acemoglu (2005); Besley and Persson (2009, 2010)). Finally, the diminishing effect of national institutions cautions against generalizing the finding on the non-effect of national institutions for areas around the capital cities in Africa.[17]

3.4.3 Further Robustness Checks

We performed several sensitivity checks to explore the robustness of the sharp difference between the cross-sectional and the within-ethnicity estimates.

Sub-Saharan Africa only First, we repeat estimation excluding North Africa from the analysis to account for the different timing and type of colonization. The Europeans had established relationships from the ancient times with North Africa while contacts with most Sub-Saharan regions were limited till mid-19th century. Appendix Table 1 reports the results. The cross-sectional coefficient on national institutions is quite similar to the estimate in the full sample (Table 3). Yet once we include ethnicity fixed effects, the coefficients on rule of law and control of corruption decline significantly and become statistically indistinguishable from zero.

Large differences in institutional quality Second, we explore whether the lack of within-ethnicity correlation between national institutions and regional development is driven by the small differences in institutional quality among African countries. In Appendix Table 2

[17] Measurement error in the institutional quality index may also explain the weakening correlation between rule of law and regional development in areas far from the capital. Since most institutional variables are measured in the capital cities capturing the rules governing activities of the formal economy, they might not reflect very accurately the institutional features in rural areas that depend on agriculture (Pande and Udry (2006)).

we report specifications estimated for two-way partitioned ethnic groups residing across country pairs with large (defined as higher than the median) differences in rule of law. Once again while the cross-sectional correlation between rule of law or control of corruption and log light density is positive and significant, it weakens considerably and becomes insignificant when we include ethnicity fixed effects.

Cross-validation with Ethnologue's language mapping Third, we repeat the analysis using the mapping of languages by Ethnologue (2005)'s WLMS (2006) database that reports the spatial distribution of linguistic groups in the early/mid 1990s.[18] The advantage of using a contemporary dataset is that it is likely to contain less error than Murdock's pre-colonial map. The disadvantage is that the current location of ethnic groups is likely to have been affected by the border drawing and the quality of (colonial and national) institutions. Appendix Table 3 reports cross-sectional and within-ethnicity specifications; we control always for a third-order RD polynomial of distance to the border and we also condition on log population density. Specifications (1)-(4) are estimated across ethnic groups that are partitioned between two countries.[19] Ethnic partitions belonging to the relatively more institutionally developed country are also more economically developed. However, once we condition on ethnicity fixed effects the coefficients on the rule of law and the control for corruption measures drop substantially and become statistically insignificant. The pattern is similar in columns (5)-(8), where we include all partitioned ethnic groups and focus on the two major partitions.

Cross-validation with Afrobarometer data Fourth, using household data from the Afrobarometer survey we examined whether differences across the border on household's access to piped water and education are correlated with differences in national institutions. Unfortunately data coverage is limited and we end up having only 32 ethnic groups that are partitioned across the 17 countries that Afrobarometer covers. Nevertheless, it is useful repeating the analysis to cross-validate our estimates based on satellite images with micro-data. Appendix Table 4 reports the results. In columns (1)-(4) the dependent variable takes the value of 1 if a household located in the historical homeland of a partitioned ethnic group i in country c has access to piped water. In columns (5)-(8) the dependent variable equals 1 if the respondent reports having some formal education. The cross-sectional estimates in columns (1) and (3) reveal that household's access to clean water is greater in countries scoring higher in the two institutional quality proxies. However, once we add ethnicity fixed effects in columns

[18] We thank an anonymous referee for proposing this cross-validation check.

[19] According to the Ethnologue, there are a few regions where linguistic groups overlap. Hence, the statistics for each partition are derived excluding such regions. Including these areas yields similar results.

(2) and (4) to exploit across-the-border variation among households of the same ethnic group, this relationship weakens considerably and even reverses sign. We obtain a similar pattern, albeit less precisely estimated, once we examine the within-ethnicity association between education and national institutions. The micro-based results with the Afrobarometer data suggest that our benchmark estimates with satellite light density as the dependent variable are not an artifact of the luminosity data.

4 Pre-colonial Ethnic Institutions and Development

4.1 Benchmark Estimates

The preliminary results in Table 2 reveal a significant association between the complexity of pre-colonial political structures and regional development. The correlation between pre-colonial ethnic institutions and regional development retained significance when we controlled for geography, the disease environment and urbanization. Yet the positive correlation between local institutions and regional development may be driven by country-level characteristics, reflecting national policies, or the type of colonization, etc. In Table 8 we estimate within-country specifications associating regional development with pre-colonial ethnic institutions. Panel A reports estimates using all observations, while in Panel B we focus on the intensive margin.

Jurisdictional Hierarchy beyond the Local Community Level The coefficient on the jurisdictional hierarchy index in column (1) is 0.33 and more than three standard deviations larger than zero.[20] The estimates are only slightly smaller than the analogous unconditional specifications (reported in Table 2 column (5)), suggesting that common-to-all-ethnicities country factors are not driving the positive correlation. In column (2) we augment the specification with distance to the coast, distance to the border, distance from the capital and the rich set of geographic controls. The coefficient on the jurisdictional hierarchy beyond the local community level retains its statistical and economic significance. In column (3) we control for population density only, while in column (4) we control jointly for geography, location and population density. The estimate on the jurisdictional hierarchy index retains statistical significance, although it falls by almost a half; this is not surprising as according to the African historiography (see Stevenson (1968), Herbst (2000)) there is a strong interplay between

[20] Note that in the pre-colonial sample when we add country fixed effects we lose one observation (instead of 683 observations of column (5) in Table 2 we have 682 ethnicity-country observations). This is because for Swaziland we have only one group with information on pre-colonial institutions, the Swazi.

geography, population density and political complexity.[21] The most conservative estimate in Panel B of Table 8 implies that across lit ethnic area, regional development increases by approximately 45 percent when one moves from the homeland of a stateless society (e.g. Luo in Uganda) to the historical homeland of an ethnicity with a hierarchical political structure beyond the local level (e.g. Ganda in Uganda).

Political Centralization In columns (5) and (8) we use an alternative binary indicator of pre-colonial institutions, based on the jurisdictional hierarchy index. Following Gennaioli and Rainer (2006, 2007) we construct a dummy variable of pre-colonial political centralization (statehood) that takes the value of zero when Murdock's jurisdictional hierarchy indicates that the tribe lacks a centralized political organization or is part of a small chiefdom. Experimenting with the re-scaled index is useful, because the aggregation may account for measurement error in the jurisdictional hierarchy index.[22] Moreover, the binary classification is more in line with the distinction of African ethnic political systems into strong ones, "*which have centralized authority, administrative machinery, and judicial institutions*" and societies lacking such structures (Fortes and Evans-Pritchard (1940)).[23] The within-country coefficient on the political centralization indicator variable is positive and highly significant. The estimate retains significance, when we control for geography (in (6)), current levels of population density (in (7)) or both (in (8)). The magnitude of political centralization in column (8) in Panel B suggests that luminosity is 30 percent ($exp(0.26) - 1 = 0.30$) 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).

Overall, these results advance the findings of Gennaioli and Rainer (2007), by showing that even when one accounts for regional geographic endowments and country fixed effects, the correlation between pre-colonial political centralization and regional development remains

[21] Note that since population density may be both a cause and an effect of ethnic institutions, the specifications where we also control for population density (in (3), (4), (7), and (8)) should be carefully interpreted. Following Angrist and Pishcke's (2008) recommendation we also used lagged (at independence) population density as a control. In these models (not reported) the estimates on the ethnic institutions measures are somewhat larger (and always significant at the 95% confidence level).

[22] To further examine the impact of precolonial local institutions, we estimated specifications with four indicator variables that take on the value one when the jurisdictional hierarchy index takes the value 1, 2, 3, and 4 respectively and zero otherwise (the omitted category consists of stateless societies). In line with the results of Table 8 the unrestricted specifications with the four indicator variables show that the higher the degree of pre-colonial centralization the higher light density is today.

[23] 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.*"

strong. The estimates are supportive of an old conjecture among African scholars that dates at least back to Fortes and Evans-Pritchard (1940) on the importance of pre-colonial ethnic institutions in the process of African development. Moreover, these results are in line with the cross-country evidence of Bockstette, Chanda, and Putterman (2002) on the role of statehood experience on contemporary development.

4.2 Institutions or Other Ethnic Traits?

One concern with the estimates in Table 8 is that some other deeply-rooted ethnic feature, related to the organization of the economy, social norms, early development, or societal structure, is driving the positive correlation between luminosity and pre-colonial political centralization. We thus examined whether political centralization rather than some other ethnic trait correlates with contemporary development utilizing information from Murdock (1967). In Table 9 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 definitions);[24] 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 some proxies of local institutional arrangements (namely an indicator for the presence of property rights, elections for local headman, ethnic class stratification and jurisdictional hierarchy within the local community).

In Specification A (reported in the first two columns) 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 enter with statistically indistinguishable from zero estimates. An indicator that takes the value 1 for societies that fishing contributes more than 5% in the pre-colonial economy enters with a positive coefficient as current economic development is higher in areas close to the sea and other streams and potentially because of blooming. 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 intensity. More importantly, the results in columns (1)-(2) show that ethnic class stratification, a societal trait that has been linked to property rights protection (e.g. Rudmin (1995)), correlates significantly with luminosity, suggesting that regional development is higher in areas where historically highly stratified societies reside. The

[24] Murdock (1967) reports data on around 60 variables. We include in the analysis around 20 because, there are a lot of missing observations for the remaining variables. We are grateful to an anonymous referee for proposing this test.

positive association between class stratification and regional development, though surprising at first glance, is in line with recent works in Latin America (e.g. Acemoglu, Bautista, Querubin, and Robinson (2008), Naritomi, Soares, and Assunção (2009), Dell (2010)). The main explanation is that in weakly institutionalized societies inequality may lead to some form of property rights protection, as the elite has the incentive to establish constraints (see Goldstein and Udry (2008) who show a positive link between political power inside the local community (class stratification) and land tenure (property rights) in rural Gambian communities).

In Specification B, reported in columns (3)-(5), we add the jurisdictional hierarchy beyond the local community index to test whether pre-colonial ethnic political centralization correlates with current regional development conditional on the other historical ethnic traits. In all specifications the jurisdictional hierarchy beyond the local community level index enters with a positive and quite stable coefficient (around 0.20). The estimate is always significant at standard confidence levels (usually at the 99% level). The results in Table 9 imply that the positive within-country association between the complexity of pre-colonial political institutions and contemporary regional development is not driven by other deeply rooted ethnic traits.

4.3 Contiguous-Ethnic-Homeland Analysis

Another concern with the within-country estimates in Tables 8 and 9 is that in spite of employing a rich conditioning set, some unobservable geographic conditions are driving the results. To account for this in Table 10 we report results from a contiguous-homelands analysis where we compare regional development across adjacent ethnicities with different degrees of pre-colonial political centralization.[25] This approach is conceptually similar to the framework employed to estimate the role of national institutions across the border. The idea behind this type of adjacent-region analysis is that it will neutralize the effect of hard-to-account-for geographical/ecological features that vary smoothly across contiguous areas. Our specification reads:

$$\Delta y_{i,j} = a_0 + \delta \Delta LOCINST_{i,j} + \Delta X'_{i,j}\Phi + \zeta_{i,j}$$

The dependent variable is the difference in log luminosity between the historical homeland of ethnicity i and the adjacent ethnic homeland of ethnicity j.[26] The estimate of interest is the δ coefficient, i.e. the difference in the jurisdictional hierarchy beyond the local community index between ethnicity i and the contiguous ethnic group j. Panel A reports estimates without

[25] In some sense this approach extends the pioneering case study of Douglas (1962), who attributed the large differences in well-being between the Bushong and the Lele to their local institutions. We are thankful to Jim Robinson for providing us with this reference.

[26] Note that in the adjacent-ethnic-group regressions the unit of analysis are pairs of ethnic homelands for which there is information on pre-colonial jurisdictional hierarchy, see Figure 7*a*. So, all relevant statistics are derived at the ethnicity level. For example, $\Delta y_{i,j} = \ln(0.01 + y_i) - \ln(0.01 + y_j)$

conditioning on (differences in) log population density while in Panel B we control for differences in log population density.

In columns (1) and (2) we compare centralized to non-centralized adjacent ethnic groups to examine how differences in jurisdictional hierarchy translate into differences in regional development. The estimates show that luminosity is significantly higher in the historical homeland of ethnicities with more complex pre-colonial political institutions. Controlling in column (2) for differences in location, geography/ecology, and natural resources has no effect on the estimate in differences in pre-colonial ethnic institutions. In columns (3) and (4) we repeat the analysis requiring that both contiguous ethnic areas (i and j) are in the same country. While we lose a quarter of our sample the results remain intact. Regional development is significantly higher in ethnic homelands where tribes with centralized pre-colonial institutions reside, as compared to their neighboring groups within the same country lacking strong pre-colonial institutions.

In columns (5)-(6) and (7)-(8) we restrict our analysis to contiguous areas where ethnicities with large differences in jurisdictional hierarchy reside. This allows to account for potential measurement error in Murdock's data; it is also useful checking whether large differences in pre-colonial institutions translate in significant differences in contemporary development. In columns (5)-(6) we require a difference of at least two levels in the jurisdictional hierarchy index between neighboring ethnicities, while in (7)-(8) we compare development between adjacent ethnic groups with the maximum difference in the respective index of pre-colonial institutions (i.e. $\Delta LOCINST_{i,j} = 3, -3$). The results show that there is a strong positive correlation between differences in luminosity and differences in the degree of pre-colonial political institutions. A couple of examples illustrate the results. In South Africa the average luminosity in the historical homeland of the Sotho, a highly centralized group is 0.68, while luminosity in the neighboring ethnic homeland of the Xam, a stateless society (Murdock's measure equals zero) is 0.13. Likewise, in Burkina Faso the average luminosity in the homeland of the Mossi, the dominant ethnic group of the pre-colonial Mossi empire that gets a score of 3 in the jurisdictional hierarchy index, is 0.27, while it is ten times smaller (0.027) in the adjacent ethnic area of the Samo that lack any level of political authority beyond the local community. Similarly, average luminosity in the homeland of the Ganda, the central ethnic group of a strong kingdom in Uganda that had a centralized authority (under the kabaka/king), is much higher (0.48) compared to the neighboring territory of the stateless Lango (0.023).

It is important to note that in both Panels of Table 10 the introduction of geographic and location controls across specifications does not affect the stability of the coefficient on the differences in political complexity. This pattern assuages (albeit not entirely resolves) concerns that the positive correlation between pre-colonial ethnic institutions and regional development

is driven by unobserved characteristics related to local endowments and geography.

4.4 Sensitivity Checks

We performed several sensitivity checks to explore the robustness of the strong positive correlation between pre-colonial political complexity and contemporary regional development. In Appendix Tables 5, 6 and 7 the first four columns focus on the ethnic-country sample (and are thus comparable to Table 8) whereas columns (5)-(8) employ pairs of adjacent ethnicities (hence comparable to Table 10).

Sub-Saharan Africa only First, we repeat estimation dropping North Africa. Appendix Table 5 reports a series of specifications illustrating that the positive correlation between the jurisdictional hierarchy index and luminosity is not driven by groups located in North Africa.

Capital Cities Second, we estimated specifications excluding ethnic homelands where capital cities are located. This robustness check is motivated by the observation that capital cities are 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 account for outliers in the dependent variable. The results reported in Appendix Table 6 suggest that the positive correlation between pre-colonial political centralization and regional development remains intact.

Population Density Third, we examined the effect of pre-colonial ethnic institutions on population density. Appendix Table 7 reports the results. While urbanization has been linked to the emergence of complex political institutions before colonization (see Hopkins (1973)) and thus these estimates may suffer from endogeneity, it is useful to see whether ethnic institutions correlate with contemporary population density for a couple of reasons. First, in the African context one could think of population density as an alternative to luminosity proxy of development. Second, in light of the lack of systematic relationship between national institutions and population density (Table 6*A*), one may worry that classical-error-in variables or systematic biases in the construction of the population estimates at a fine grid somewhat explain our results. The estimates in Appendix Table 7 show that there is a strong correlation between pre-colonial ethnic institutions and contemporary population density across groups as well as within pairs of adjacent ethnicities.

5 Conclusion

We study the role of institutional quality in shaping contemporary comparative development in Africa focusing both on formal nationwide structures and informal ethnic-specific arrangements. We perform our analysis at the regional level utilizing anthropological and historical data on the spatial distribution and local institutions of African ethnicities at the time of colonization. To circumvent data unavailability on regional development in Africa we use satellite data on light density at night to measure economic performance across ethnic areas. Exploiting within-ethnicity across-country variation on contemporary country level institutions as well as within-country across-ethnicity variation in ethnic pre-colonial institutions, we document new empirical regularities on the role of institutional structures on African development.

First, our cross-sectional specifications reveal a positive correlation between contemporaneous nationwide institutions and regional development. Yet this correlation does not identify the one-way effect of the rule of law on development, as besides reverse causation there could be other country or local characteristics that affect both institutional and economic outcomes. To push on the identification front, we take advantage of the fact that the arbitrarily drawn national boundaries across the African landscape partitioned groups in different countries, thus subjecting identical cultures to different country-level institutions. The analysis uncovers that differences in economic performance within ethnic groups partitioned across different countries cannot be explained by countrywide differences in institutional quality. While this result does not necessarily generalize to areas far from the national borders or other parts of the world, it casts doubt on the causal interpretation of the cross-country positive correlation between institutional quality and economic development in Africa.

Second, we explore the significant heterogeneity of historical ethnic institutions and examine their role on regional development. In line with an influential conjecture among African scholars, we show that ethnic pre-colonial institutions correlate significantly with contemporary regional development, even when we control for geography at a fine level, country-characteristics, and other ethnic traits. Since we do not have random assignment on ethnic institutions, this correlation does not necessarily imply causation. Yet our result provides large-scale formal econometric evidence in support of the African historiography that emphasizes the importance of ethnic institutions and other historical features in shaping economic performance suggesting that research on African development needs to focus at the ethnicity rather than the country level.

Moreover, our codification of partitioned ethnic groups and the combination of high resolution proxies of development (such as satellite light density at night) with historical measures on culture and institutions provide a platform for subsequent research. One could employ

our approach to shed light on the perennial debate regarding the fundamental determinants of comparative economic development across countries; examining for example the effect of human capital, public policies, and democracy on economic performance in areas close to the national borders. Moreover, one could investigate how pre-colonial ethnic features interact with contemporary public policies to determine regional development.

6 Data Appendix

6.1 Variables at the ethnicity-country level

Light Density at Night: Light Density is calculated averaging light density observations across pixels that fall within the unit of analysis. We use the average of the values in 2007 and 2008. In the regressions we use Log (0.01 + Average Luminosity) or Log (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 and in 1960). *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 within an ethnic group district covered by rivers or lakes in sq. km.). *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, Colorado, USA. Global Ministry Mapping System.*

Elevation: Average elevation in km. *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 the area of each ethnic-country observation. The index is the product of two components capturing the climatic (mean monthly temperature and precipitation between 1961-1990) and soil suitability for cultivation. *Source: Michalopoulos (2011); 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. The index has been constructed for 0.5 degree by 0.5 degree grid-cells globally. *Source: Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004)*

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

Distance to the Sea Coast: The geodesic distance of the centroid of each ethnic group in a country from the nearest coastline, measured in 1000*s* of km's. *Source: 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 of the centroid of each ethnic group in a country from the closest border, measured in 1000*s* of km's. *Source: Calculated using the Haversine formula.*

Petroleum: Indicator variable that takes on the value of one if an oil field is found in the region 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 takes on the value of one if a diamond mine is found in the region of ethnic group i in country c. *Source: Map of Diamond Resources. www.prio.no/CSCW/Datasets/Geographical-and-Resource/Diamond-Resources/*

Education: Average education across respondents living in homeland i in country c. The education categories across which ethnic-level education is estimated are the following: (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) postsecondary qualifications, but no university, (viii) some university, (ix) university completed, and (x) postgraduate.

Source: Afrobarometer available at http://www.economics.harvard.edu/faculty/nunn/data_nunn

Access to piped water: Percentage of respondents with access to piped water living in homeland i in country c.

Source: Afrobarometer available at http://www.economics.harvard.edu/faculty/nunn/data_nunn

6.2 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. *Source: World Bank Governance Matters Indicators Database (Kaufman, Kraay, and Mastruzzi (2005)). available at: http://info.worldbank.org/governance/wgi/index.asp*

Control of Corruption: Index on the control of corruption "capturing perceptions of the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as capture of the state by elites and private interests." The standardized index ranges from -2.5 to $+2.5$ with lower values indicating a higher degree of corruption. *Source: World Bank Governance Matters Indicators Database (Kaufman, Kraay, and Mastruzzi (2005)). available at: http://info.worldbank.org/governance/wgi/index.asp*

6.3 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 Entholinguistic 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 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 plans 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: Binary index that indicates the "*care and tending of domestic animals, including milking*". 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 v4.*

Milking: Binary index that equals zero 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: $0-9$ 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 Atlas v28.*

Polygyny: Indicator that equals one when polygyny is practised and zero otherwise. The indicator equals one when the original variable indicates that polygyny is common and when large extended families are present. *Source: Murdock (1967); variable code in the Ethnographic Atlas v8.*

Alternative Polygyny: Alternative indicator that equals one when polygyny is practised 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 (*v15*) equals 6 ("*clan communities or clan barrios*") 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, 3 indicates semi-sedentary, 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.*

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.*

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.*

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 in which an elite class derives its superior status from control over scarce resources, particularly land, and is thereby differentiated from a propertyless 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 is absence or near absence. *Source: Murdock (1967); variable code in the Ethnographic Atlas v70; the construction of the index follows Fenske (2009).*

Property Rights: 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; the construction of the index follows Fenske (2009).*

References

Acemoglu, D. (2005): "Politics and Economics in Weak and Strong States," *Journal of Monetary Economics*, 52(7), 1199–1226.

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 M. Dell (2010): "Productivity Differences Between and Within Countries," *American Economic Journal: Macroeconomics*, 2(1), 169–188.

Acemoglu, D., S. Johnson, and J. A. Robinson (2003): "An African Success Story: Botswana," in *In Search of Prosperity: Analytical Narrative on Economic Growth*, ed. by D. Rodrik. Princeton University Press, Princeton, NJ.

——— (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., D. Ticchi, and A. Vindigni (2011): "Emergence and Persistence of Inefficient States," *Journal of the European Economic Association*, 9(2), 177–208.

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

Alesina, A., and E. Spolaore (2003): *The Size of Nations*. MIT Press.

Angrist, J., and V. C. Lavy (2001): "New Evidence on Classroom Computers and Pupil Learning," *Economic Journal*, 112, 735–765.

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.

Asiwaju, A. (1985): "The Conceptual Framework," in *Partitioned Africans*, pp. 1–18. St. Martin Press, New York.

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. (1981): *States and Markets in Africa*. University of California Press, Berkeley, CA.

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

——— (2010): "State Capacity, Conflict and Development," *Econometrica*, 78(1), 1–34.

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. (2009): "States, Law, and Property Rights in West Africa," Working Paper, New York University School of Law.

Cameron, C. A., J. Gelbach, and D. Miller (2006): "Robust Inference with Multi-Way Clustering," NBER Technical Working Paper No. T0327.

Chen, X., and W. D. Nordhaus (2010): "The Value of Luminosity Data as a Proxy for Economic Statistics," Working Paper, NBER, No. 16317.

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

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

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.

Easterly, W., and R. Levine (2003): "Tropics, Germs and Crops: the Role of Endowments in Economic Development," *Journal of Monetary Economics*, 50(1), 1203–1250.

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. (2000): "Precolonial Institutions, Post-Colonial States, and Economic Development in Tropical Africa," *Political Research Quarterly*, 53(7), 7–36.

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.

Feyrer, J., and B. Sacerdote (2009): "Colonialism and Modern Income: Islands as Natural Experiments," *Review of Economics and Statistics*, 91(2), 245–262.

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

Franck, R., and I. Rainer (2009): "Does the Leader's Ethnicity Matter? Ethnic Favoritism, Education and Health in Sub-Saharan Africa," Working Paper, mimeo Bar Ilan University.

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.

Gallup, J. L., J. D. Sachs, and A. Mellinger (1999): "Geography and Economic Development," Working Paper, CID Working Papers 1.

Gennaioli, N., R. L. Porta, F. L. de Silanes, and A. Shleifer (2011): "Human Capital and Regional Development," mineo, Harvard University.

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.

Glaeser, E. L., R. L. Porta, F. L. de Silanes, and A. Shleifer (2004): "Do Institutions Cause Growth?," *Journal of Economic Growth*, 9(3), 271–303.

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.

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

Guiso, L., P. Sapienza, and L. Zingales (2008): "Long Term Persistence," Working Paper, NBER, No. 14278.

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.

Hopkins, A. G. (1973): *An Economic History of West Africa*. Columbia University Press, New York, NY.

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.

Imbens, G. W., and T. Lemieux (2008): "Regression discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2), 615–635.

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

Kaufmann, D., A. Kraay, and M. Mastruzzi (2005): "Governance Matters IV : Governance Indicators for 1996-2004," *Policy Research Working Paper Series, The World Bank*.

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.

La-Porta, R., and A. Shleifer (2008): "The Unofficial Economy and Economic Development," Brookings Papers on Economic Activity.

Lee, D. S., and T. Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(2), 281–355.

Lee, D. S., E. Moretti, and M. J. Butler (2004): "Do Voters Affect Or Elect Policies? Evidence from the U. S. House," *The Quarterly Journal of Economics*, 119, 807–859.

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

Michalopoulos, S. (2011): "The Origins of Ethnolinguistic Diversity," American Economic Review, forthcoming.

Michalopoulos, S., and E. Papaioannou (2011): "The Long-Run Effects of the Scramble for Africa," mimeo Harvard University.

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

Migdal, J. S. (1988): *Strong Societies and Weak States: State-Society Relations and State Capabilities in the Third World*. Princeton University Press.

Miguel, E. (2004): "Tribe or Nation? Nation-Building and Public Goods in Kenya versus Tanzania," *World Politics*, 56(3), 327–362.

Miguel, E., and M. K. Gugerty (2005): "Ethnic Diversity, Social Santions, and Public Goods in Kenya," *Journal of Public Economics*, 89(11), 2325–2368.

Miles, W. F. (1994): *Hausaland Divided: Colonialism and Independence in Nigeria and Niger*. Cornell University Press, Ithaca and London.

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

Moulton, B. R. (1986): "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics*, 32, 385–397.

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

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

Naritomi, J., R. R. Soares, and J. J. Assunção (2009): "Institutional Development and Colonial Heritage within Brazil," Working Paper, IZA DP No. 4276.

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

——— (2009): "The Importance of History for Economic Development," *Annual Review of Economics*, 1(1), 65–92.

Nunn, N., and D. Puga (2011): "Ruggedness: The Blessing of Bad Geography in Africa," *Review of Economics and Statistics, forthcoming*.

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

Pande, R., and C. Udry (2006): "Institutions and Development: A View from Below," in *Proceedings of the 9th World Congress of the Econometric Society*, ed. by R. Blundell, W. Newey, and T. Persson. Cambridge University Press.

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

Posner, D. N. (2005): *Institutions and Ethnic Politics in Africa*. Cambridge University Press, Cambridge, UK.

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*.

Rudmin, F. W. (1995): "Cross-Cultural Correlates of the Ownership of Private Property: Two Samples of Murdock's Data," *Journal of Socio-Economics*, 24, 345–373.

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 (2005): "Borders and Growth," *Journal of Economic Growth*, 10(4).

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

Sutton, P. C., C. D. Elvidge, and T. Ghosh (2007): "Estimation of Gross Domestic Product at Sub-National Scales Using Nighttime Satellite Imagery," *International Journal of Ecological Economics and Statistics*, 8(7), 5–21.

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.

Tilly, C. (1985): "War Making and State Making as Organized Crime," in *Bringing the State Back In*, ed. by P. Evans, D. Rueschemeyer, and T. Skocpol. Cambridge University Press.

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.

WLMS (2006): World Language Mapping System, Version 3.2 Available on-line at http://www.gmi.org/wlms/.

Table 1: Summary Statistics

variable	Obs.	mean	st. dev.	p25	median	p75	min	max
			Panel A: Full Sample					
Light Density	1247	0.364	1.553	0.000	0.017	0.134	0.000	25.140
Ln (0.01 + Light Density)	1247	-3.027	1.719	-4.605	-3.597	-1.937	-4.605	3.225
Ln (Light Density)	898	-2.894	2.217	-4.299	-2.909	-1.491	-10.597	3.224
Ln (0.01 + Population Density)	1247	2.608	2.097	1.730	2.992	3.980	-4.605	7.432
Ln (1 + Water Area)	1247	0.273	0.427	0.004	0.101	0.340	0.000	3.119
Mean Elevation	1247	0.617	0.437	0.288	0.489	0.936	0.000	2.181
Land Suitability For Agriculture	1247	0.400	0.244	0.241	0.416	0.569	0.001	0.979
Malaria Stability Index	1247	0.708	0.352	0.500	0.876	1.000	0.000	1.000
Oil Deposit Indicator	1247	0.083	0.277	0.000	0.000	0.000	0.000	1.000
Diamond Mine Indicator	1247	0.091	0.287	0.000	0.000	0.000	0.000	1.000
Distance to the Capital City	1247	0.521	0.378	0.265	0.422	0.673	0.010	1.935
Distance to the Sea Coast	1247	0.608	0.434	0.225	0.563	0.935	0.000	1.739
Distance to the Border	1247	0.107	0.117	0.020	0.061	0.159	0.000	0.637
Rule of Law	1242	-0.843	0.586	-1.266	-0.888	-0.442	-1.912	0.708
Control of Corruption	1242	-0.771	0.497	-1.048	-0.873	-0.468	-1.590	0.722
Jurisdictional Hierarchy	683	1.215	0.965	0.000	1.000	2.000	0.000	4.000
Political Centralization	683	0.356	0.479	0.000	0.000	1.000	0.000	1.000
			Panel B: Pre-colonial Institutions Sample					
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
Ln (Light Density)	519	-2.887	2.173	-4.269	-2.909	-1.548	-10.597	3.224
			Panel C: Partitioned Ethnic Groups Sample					
Light Density	526	0.221	0.755	0.000	0.014	0.092	0.000	8.561
Ln (0.01 + Light Density)	526	-3.193	1.575	-4.605	-3.745	-2.281	-4.605	2.148
Ln (Light Density)	374	-3.176	2.183	-4.576	-3.116	-1.755	-10.597	2.147

Table 1 reports descriptive statistics for all variables employed in the empirical analysis. Panel A reports summary statistics for all variables in the maximum sample of country-ethnicity observations. Panel B reports summary statistics for the dependent variable in the sample with data availability on pre-colonial ethnic institutions (from Murdock (1967)). Panel C gives summary statistics for the dependent variable in the sample of partitioned ethnic groups. The Data Appendix gives detailed variable definitions and data sources.

Table 2 - Preliminary Evidence
National Contemporary Institutions, Pre-colonial Ethnic Institutions, and Regional Development

	National Contemporary Institutions				Ethnic Pre-colonial Institutions				National and Ethnic Institutions			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Rule of Law	0.8514***	0.8816***	0.9177***	0.7467***					0.8011***	0.7806***	0.8114***	0.7293***
Double-clustered s.e.	(0.2743)	(0.2427)	(0.2263)	(0.1692)					(0.2862)	(0.2552)	(0.2484)	(0.1903)
Conley's s.e.	[0.2181]	[0.2009]	[0.1756]	[0.1283]					[0.2238]	[0.2096]	[0.2051]	[0.1668]
Jurisdictional Hierarchy					0.4106***	0.3483**	0.3213***	0.2036***	0.3411***	0.2707**	0.2462***	0.1729***
Double-clustered s.e.					(0.1246)	(0.1397)	(0.1026)	(0.0735)	(0.0982)	(0.1067)	(0.0784)	(0.0627)
Conley's s.e.					[0.1294]	[0.1288]	[0.1014]	[0.0691]	[0.1094]	[0.1059]	[0.0810]	[0.0617]
Adjusted R-squared	0.085	0.251	0.362	0.463	0.056	0.246	0.361	0.461	0.123	0.308	0.415	0.498
Population Density	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Observations	1242	1242	1242	1242	683	683	683	683	680	680	680	680

Table 2 reports OLS estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1)-(4), (9)-(12)) and pre-colonial ethnic institutions, as reflected in Murdock's (1967) index of jurisdictional hierarchy beyond the local community index (in columns (5)-(12)). The dependent variable is log (0.01 + light density at night from satellite) at the ethnicity-country level. In the specifications in columns (2)-(4), (6)-(8), and (10)-(12) we control for log (0.01 + population density). In columns (3), (4), (7), (8), (11), and (12) we control for location augmenting the specification with distance of the centroid of each ethnicity-country area from the capital city of each country, distance from the closest sea coast, and distance from the national border. The set of geographic controls in columns (4), (8) and (12) includes log (1 + area under water (lakes, rivers, and other streams)), 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.

Table 3: Contemporary National Institutions and Regional Development across and within Partitioned Ethnic Groups

	All Observations		Intensive Margin (Lit Areas)		All Observations		Intensive Margin (Lit Areas)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	0.6610***	0.1170	0.8711***	0.2909				
Double-clustered s.e.	(0.2139)	(0.1825)	(0.2638)	(0.2730)				
Conley's s.e.	[0.1901]	[0.1089]	[0.2467]	[0.2121]				
Control of Corruption					0.7898***	0.0884	0.9997***	0.4068
Double-clustered s.e.					(0.2639)	(0.2375)	(0.3319)	(0.3801)
Conley's s.e.					[0.2371]	[0.1507]	[0.3111]	[0.2707]
Adjusted R-squared	0.225	0.774	0.287	0.765	0.227	0.773	0.287	0.766
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	526	526	313	313	526	526	313	313

Table 3 reports cross-sectional and within-ethnicity OLS estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1)-(4)) and control for corruption index (in columns (5)-(8)) in areas of partitioned ethnicities. The dependent variable in columns (1), (2), (5), and (6) is the log (0.01 + light density at night from satellite) at the ethnicity-country level The dependent variable in columns (3), (4), (7), and (8) is the log (light density at night from satellite) at the ethnicity-country level. In these specifications we thus ignore unlit areas focusing on the "intensive" margin of luminosity. Odd-numbered specifications report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density).

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.

Table 4: Validity of the Regression Discontinuity (RD) Design

	Rule of Law	Control for Corruption
	(1)	(2)
Dependent Variables		
Population Density in 1960	0.1552	0.0451
	(0.2087)	(0.2677)
Water Area	0.0434	0.0572
	(0.0497)	(0.0682)
Elevation	0.0156	0.0357
	(0.0253)	(0.0321)
Suitability for Agriculture	0.0115	0.0043
	(0.0200)	(0.0264)
Malaria Stability Index	-0.0187	-0.0419
	(0.0283)	(0.0328)
Distance to the Capital City	-0.2472**	-0.2293
	(0.1199)	(0.1573)
Distance to the Sea Coast	0.0000	0.0053
	(0.0150)	(0.0195)
Distance to the Border	-0.0037	-0.0046
	(0.0077)	(0.0098)
Diamond Mine Indicator	-0.0133	-0.0256
	(0.0457)	(0.0667)
Oil/Petroleum Deposit Indicator	0.0297	0.0515
	(0.0230)	(0.0385)
Observations	526	526

Table 4 reports within-ethnicity OLS estimates associating the set of conditioning variables in our subsequent empirical specifications with contemporary national institutions across partitioned ethnicities. Specifically we regress log (0.01 + population density in 1960 (around independence)), log (1 + area under water), mean elevation, land's suitability for agriculture, a malaria stability index, distance of the centroid of each partitioned ethnic area to the capital city, distance to the sea coast, distance to the national border, a diamond mine indicator, and an oil indicator on the rule of law index (in column (1)) and on the control of corruption index (in column (2)). The Data Appendix gives detailed variable definitions and data sources. All specifications include a vector of ethnicity fixed effects (constants not reported). Below the estimates we report in parentheses double-clustered standard errors at the country and ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 5 - RD Design: Contemporary National Institutions and Regional Development across and within Partitioned Ethnic Groups

Panel A: Global Polynomial Control Function Method

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	0.8153***	0.0644	0.6432**	0.0349				
	(0.2645)	(0.2795)	(0.2591)	(0.3294)				
Control of Corruption					0.9522***	-0.1235	0.6920**	-0.1762
					(0.2990)	(0.3214)	(0.2869)	(0.3700)
Adjusted R-squared	0.301	0.843	0.403	0.846	0.298	0.841	0.398	0.846
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	Yes	Yes	No	No	Yes	Yes
Observations	454	454	454	454	454	454	454	454

Panel B: Local Linear Regression Method

Bandwidth	100 km		50 km		100 km		50 km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	0.6544***	0.1464	0.6306***	0.1958				
	(0.2096)	(0.1819)	(0.1992)	(0.2014)				
Control of Corruption					0.8017***	0.1028	0.7609***	0.1774
					(0.2571)	(0.2411)	(0.2441)	(0.2627)
Adjusted R-squared	0.214	0.775	0.184	0.774	0.219	0.774	0.187	0.772
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	526	526	526	526	526	526	526	526

Table 5 - Notes

Table 5 reports cross-sectional and within-ethnicity estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1)-(4)) and control for corruption index (in columns (5)-(8)) in areas of partitioned ethnicities with two alternative regression discontinuity (RD) approaches, a global polynomial control function method (Panel A) and a local linear regression method (Panel B).

In Panel A we include a third-order RD polynomial in distance of the centroid of each partitioned ethnic area to the national border; distance takes positive values for partitioned areas that belong to the relatively high institutional quality country and negative values for partitioned areas that fall in the relatively low institutional quality country. All specifications include interactions between the polynomial terms and the proxy measure of national institutions. Estimation is performed across the two major partitions of each ethnic group.

In Panel B we report local linear regression estimates restricting estimation in the areas close to the national border. In columns (1), (2), (5), and (6) we focus of ethnic areas within 50 kilometers of each side of the national border (total 100 kilometers). In columns (3), (4), (7), and (8) we focus on ethnic areas within 25 kilometers of each side of the national border (total 50 kilometers). In both Panel A and Panel B odd-numbered specifications report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density). In columns (3), (4), (7), and (8) we control for location and geography. The set of control variables includes the distance of the centroid of each ethnicity-country area from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), 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. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 6: Accounting for Migration

Panel A: National Institutions and Regional Population Density

	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	-0.2453	0.1464	-0.1527			
	(0.4076)	(0.2399)	(0.2950)			
Control of Corruption				-0.3769	0.0865	-0.2043
				(0.4757)	(0.2923)	(0.3746)
Adjusted R-squared	0.006	0.815	0.851	0.009	0.815	0.851
Ethnicity Fixed Effects	No	Yes	Yes	No	Yes	Yes
Location Controls	No	No	Yes	No	No	Yes
Geographic Controls	No	No	Yes	No	No	Yes
Observations	526	526	526	526	526	526

Table 6 Panel A reports cross-sectional and within-ethnicity OLS estimates associating regional population density with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1)-(3)) and control for corruption index (in columns (4)-(6)) in areas of partitioned ethnicities. The dependent variable is log (0.01 + population density at the ethnicity-country level). Columns (1) and (4) report cross-sectional estimates. Columns (2), (3), (5), and (6) report within-country estimates that include a vector of ethnicity fixed effects (constants not reported). In columns (3) and (6) we control for location and geography. The set of control variables includes the distance of the centroid of each ethnicity-country area from the capital city of each country, the distance from the closest sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), 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 6: Accounting for Migration, Local Trade and Lights' Bleeding/Stealing

Panel B: Excluding Pixels Close to the National Border

Excluding Area within	50 km		100 km		50 km		100 km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	0.6100**	0.1997	0.7826***	0.1057				
	(0.2761)	(0.2194)	(0.1970)	(0.3196)				
Control of Corruption					0.7721**	0.3289	0.8795***	0.1452
					(0.3459)	(0.3514)	(0.2448)	(0.4861)
Adjusted R-squared	0.210	0.729	0.247	0.749	0.220	0.730	0.244	0.749
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	365	365	218	218	365	365	218	218

Table 6 Panel B reports cross-sectional and within-ethnicity OLS estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1)-(4)) and control for corruption index (in columns (5)-(8)) in areas of partitioned ethnicities. The dependent variable is the log (0.01 + light density at night from satellite at the ethnicity-country level). In columns (1), (2), (5), and (6) we exclude from the estimation ethnic areas 25 kilometers from each side of the national border (total 50 kilometers). In columns (3), (4), (7), and (8) we exclude from the estimation ethnic areas 50 kilometers from each side of the national border (total 100 kilometers). Odd-numbered specifications report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density).

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 7 - Differential Effect of National Institutions Far and Close to the Capital Cities

	Global Control Polynomial Function		Local Regression 100 km Bandwidth	
	(1)	(2)	(3)	(4)
	Panel A: Ethnic Partitions Both Close or Far from the Capital Cities			
Rule of Law	1.5424***	0.7247	1.4308***	0.4465
	(0.3599)	(0.6479)	(0.3807)	(0.5079)
Far Dummy Variable	-0.8029*	0.3771	-0.5251	-0.5473
	(0.4318)	(1.6135)	(0.4805)	(1.3905)
Rule of Law * Far Dummy Variable	-0.7935**	-0.3435	-0.7143*	-0.2120
	(0.3467)	(0.5948)	(0.3996)	(0.5706)
Adjusted R-squared	0.382	0.835	0.305	0.832
Observations	196	196	196	196
	Panel B: All Observations			
	Two Major Partitions		**All Partitions**	
Rule of Law	1.1672***	0.4542	1.0080**	0.5269
	(0.4104)	(0.6647)	(0.4450)	(0.3596)
Distance to the Capital City	-1.0067*	-0.5759	-0.8276	-0.7635
	(0.5811)	(0.6744)	(0.6328)	(0.5314)
Rule of Law * Distance to the Capital City	-0.6852*	-0.5361	-0.6490	-0.6388
	(0.4014)	(0.5667)	(0.4932)	(0.4409)
Adjusted R-squared	0.31	0.845	0.222	0.779
Observations	454	454	526	526
Ethnicity Fixed-Effects	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes

Table 7 reports cross-sectional and within-ethnicity OLS estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1)-(4)) and control for corruption index (in columns (5)-(8)) in areas of partitioned ethnicities. The dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. Odd-numbered specifications report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density). In Panel A we limit attention to two-way partitioned ethnic groups where either both partitions are close to the respective capital city or both are far from the respective capital city (using the median value of the distance to the capital city). Far is an indicator (dummy) variable that equals one for partitioned ethnic areas that are far from the capital. In Panel B we focus on all partitioned ethnic groups and explore the interaction of national institutions with distance to the capital city.
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 8: Pre-colonial Ethnic Institutions and Regional Development within African Countries

	Jurisdictional Hierarchy				Political Centralization			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Observations								
Ethnic Institutions	0.3260***	0.2900***	0.2105***	0.1878***	0.5264***	0.5330***	0.3413***	0.3385***
	(0.0852)	(0.0816)	(0.0554)	(0.0538)	(0.1492)	(0.1516)	(0.0898)	(0.1071)
Adjusted R-squared	0.409	0.534	0.597	0.654	0.400	0.531	0.593	0.652
Observations	682	682	682	682	682	682	682	682
Panel B: Intensive Margin (Lit Areas)								
Ethnic Institutions	0.3279***	0.3057***	0.1651**	0.1382*	0.4819**	0.5775***	0.2649**	0.2663*
	(0.1240)	(0.1130)	(0.0705)	(0.0715)	(0.2369)	(0.2142)	(0.1234)	(0.1399)
Adjusted R-squared	0.424	0.547	0.638	0.670	0.416	0.545	0.636	0.670
Observations	517	517	517	517	517	517	517	517
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes

Table 8 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). In these specifications we thus ignore unlit areas focusing on the "intensive" margin of 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 the binary Gennaioli and Rainer (2006, 2007) political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local community variable. This index takes on the value of zero for stateless societies and ethnic groups that are part of petty chiefdoms. All specifications include a set of country fixed effects (constants not reported). In columns (3), (4), (7), and (8) we control for log (0.01 + population density). 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 capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), 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 9: The Long-Run Effects of Pre-colonial Political Complexity Conditioning on Other Pre-colonial Ethnic Features

	Specification *A*		Specification *B*		
	Additional Variable	Observations	Additional Variable	Jurisdictional Hierarchy beyond the Local Community	Observations
	(1)	(2)	(3)	(4)	(5)
Gathering	-0.0941	742	-0.0783	0.2099***	676
	(0.1721)		(0.1850)	(0.0561)	
Hunting	-0.0311	742	-0.0133	0.2116***	676
	(0.1229)		(0.1245)	(0.0573)	
Fishing	0.2442*	742	0.2447*	0.2111***	676
	(0.1312)		(0.1254)	(0.0564)	
Animal Husbandry	0.0545	742	0.0346	0.2026***	676
	(0.0410)		(0.0428)	(0.0628)	
Milking	0.1782	695	0.0864	0.2032***	674
	(0.1389)		(0.1433)	(0.0589)	
Agriculture Dependence	-0.1066**	742	-0.1052**	0.2113***	676
	(0.0438)		(0.0459)	(0.0572)	
Agriculture Dependence Alternative	0.0237	695	-0.0116	0.2107***	674
	(0.1066)		(0.1039)	(0.0563)	
Polygyny	0.0382	742	0.0178	0.2110***	676
	(0.1111)		(0.1190)	(0.0582)	
Polygyny Alt	-0.2337	722	-0.2614	0.2211***	665
	(0.2851)		(0.3121)	(0.0543)	
Clan Communities	-0.0779	742	0.0357	0.2158***	676
	(0.1381)		(0.1285)	(0.0555)	
Settlement Pattern	-0.0171	694	-0.0073	0.2121***	673
	(0.0353)		(0.0377)	(0.0578)	
Complex Settlements	0.2341	694	0.2138	0.2006***	673
	(0.1618)		(0.1619)	(0.0562)	

Hierarchy of Local Community	0.0154	679	-0.0030	0.2102***	674
	(0.0789)		(0.0832)	(0.0577)	
Patrilineal Descent	-0.1879	726	-0.2010	0.1947***	665
	(0.1332)		(0.1307)	(0.0511)	
Class Stratification	0.1257**	603	0.0652	0.1576**	566
	(0.0513)		(0.0584)	(0.0745)	
Class Stratification Indicator	0.3970**	603	0.2758	0.1440**	566
	(0.1831)		(0.1896)	(0.0615)	
Elections	0.3372	533	0.2709	0.2199***	498
	(0.2518)		(0.2617)	(0.0601)	
Slavery	-0.0202	654	-0.1170	0.2030***	604
	(0.1600)		(0.1595)	(0.0620)	
Property Rights	-0.1967	562	-0.1788	0.2198***	523
	(0.2201)		(0.2295)	(0.0690)	

Table 9 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.

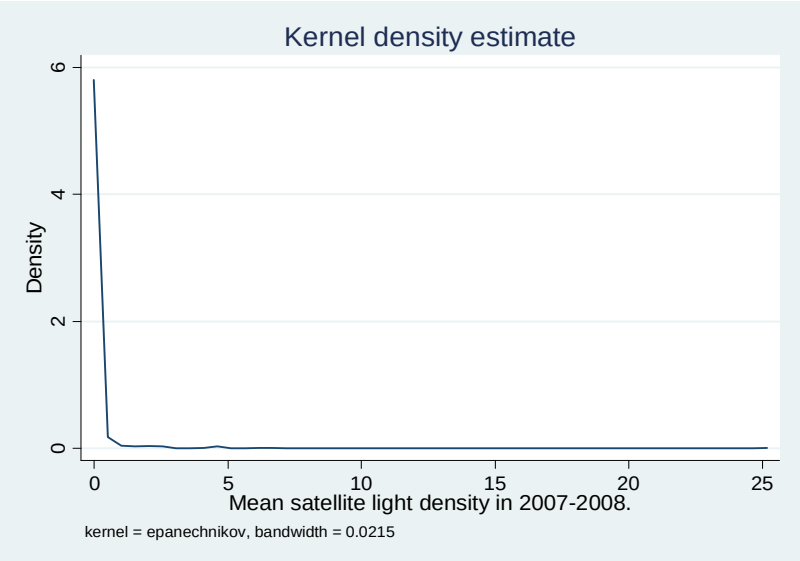
Table 10: Differences in Pre-colonial Ethnic Institutions and Differences in Regional Development across Contiguous Ethnicities

	Centralized versus Non-Centralized		Centralized versus Non-Centralized within the Same Country		Difference in Jurisdictional Hierarchy Index > | 1 |		Difference in Jurisdictional Hierarchy Index > | 2 |	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Without Conditioning on Differences in Population Density							
Differences in Jurisdictional Hierarchy	0.2476***	0.2653***	0.2252***	0.2500***	0.2146***	0.2349***	0.3332***	0.3353***
	(0.0625)	(0.0563)	(0.0667)	(0.0571)	(0.0698)	(0.0634)	(0.0961)	(0.0938)
Adjusted R-squared	0.084	0.184	0.100	0.270	0.065	0.170	0.451	0.786
	Panel B: Conditional on Differences in Population Density							
Differences in Jurisdictional Hierarchy	0.1236**	0.1449***	0.1129**	0.1258**	0.1291**	0.1434***	0.1998***	0.2798**
	(0.0495)	(0.0498)	(0.0510)	(0.0534)	(0.0543)	(0.0547)	(0.0768)	(0.1095)
Adjusted R-squared	0.366	0.381	0.506	0.533	0.336	0.356	0.724	0.801
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	258	258	125	125	186	186	21	21

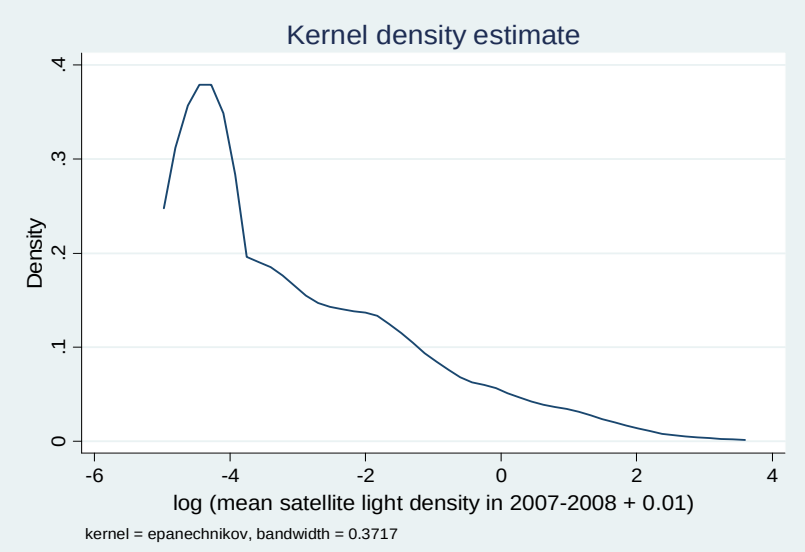
Table 10 reports OLS estimates associating differences in regional development (as reflected in differences in log (0.01 + satellite light density) between two adjacent ethnic homelands (of ethnicities i and j) with differences in pre-colonial ethnic institutions as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community index (that ranges from 0 to 4). In columns (1)-(2) we compare centralized to non-centralized adjacent ethnic groups. In columns (3)-(4) we compare centralized to non-centralized adjacent ethnic groups that are in the same country. In columns (5)-(6) we focus on pairs of adjacent ethnic homelands with at least 2-point difference in the four-scale index of the jurisdictional hierarchy beyond the local community index. In columns (7)-(8) we focus on pairs of adjacent ethnic homelands with at least 3-point difference in the four scale index of the jurisdictional hierarchy beyond the local community index.

In even numbered columns we control for differences in the location and differences in geography between the two adjacent ethnic homelands. The set of control variables includes the differences of the two adjacent ethnic homelands with respect to: distance of the centroid of each ethnicity from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1+ area under water (lakes, rivers, and other streams)), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. In Panel A we do not control for differences in log (0.01 + population density). In Panel B we control for differences in log (0.01 + population density) between the two adjacent ethnic homelands. The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the ethnicity *i* and the ethnicity *j* dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

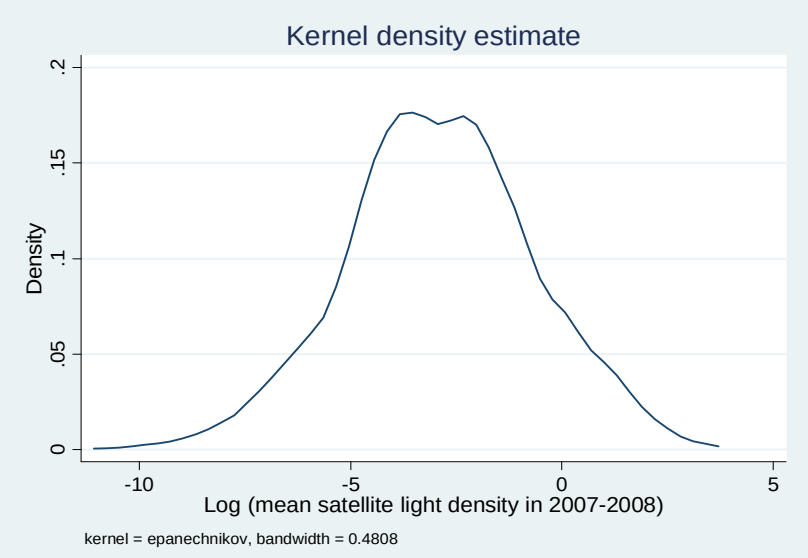
Supplementary Appendix Table Figure 1a



Supplementary Appendix Table Figure 1b



Supplementary Appendix Table Figure 1c



Supplementary Appendix Table 1: Contemporary National Institutions and Regional Development across and within Partitioned Ethnic Groups

Excluding North Africa

	Simple		Global Polynomial Control		Local Linear (100km)	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Rule of Law			
Rule of Law	0.5675**	0.1647	0.5958**	0.0724	0.5699**	0.1989
	(0.2475)	(0.1733)	(0.2616)	(0.1728)	(0.2437)	(0.1851)
adjusted R-squared	0.268	0.771	0.299	0.825	0.269	0.771
			Panel B: Control for Corruption			
Control for Corruption	0.6631**	0.1765	0.7533**	0.0211	0.6894**	0.1750
	(0.3057)	(0.2144)	(0.3320)	(0.2328)	(0.2985)	(0.2423)
adjusted R-squared	0.268	0.770	0.306	0.825	0.272	0.770
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	No	Yes	Yes	No	No
Observations	487	487	420	420	487	487

Appendix Table 1 reports cross-sectional and within-ethnicity OLS estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in Panel A) and control for corruption index (in Panel B) in areas of partitioned ethnicities located in Sub-Saharan African countries (excluding ethnic areas in North Africa). The dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. In columns (3)-(4) we include a third-order RD polynomial in distance of the centroid of each partitioned ethnic area to the national border; distance takes positive values for partitioned areas that fall into the relatively high institutional quality country and negative values for partitioned areas that fall into the relatively low institutional quality country. These specifications also include interactions between the polynomial terms and the proxy measure of national institutions.

Estimation is performed across the two-major partitions of each ethnic group. In columns (5)-(6) we report local linear regression estimates focusing on ethnic areas 50 kilometers within each side of the national border (total 100 kilometers). In both Panel A and Panel B odd-numbered columns report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density). 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 Table 2: Contemporary National Institutions and Regional Development across and within Partitioned Ethnic Groups

Large Differences in National Institutional Quality

	Simple		Global Polynomial Control		Local Linear (100km)	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: Rule of Law			
Rule of Law	0.7858***	0.1193	0.7159***	0.0174	0.7725***	0.1620
	(0.2091)	(0.1880)	(0.2601)	(0.3813)	(0.1976)	(0.2071)
Adjusted R-squared	0.299	0.854	0.429	0.864	0.282	0.862
			Panel B: Control for Corruption			
Control for Corruption	0.9369***	0.0162	0.6389**	-0.3907	0.9146***	0.0417
	(0.3009)	(0.2492)	(0.3006)	(0.4066)	(0.2942)	(0.2750)
Adjusted R-squared	0.226	0.855	0.359	0.865	0.218	0.859
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	No	Yes	Yes	No	No
Observations	230	230	230	230	230	230

Appendix Table 2 reports cross-sectional and within-ethnicity estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in Panel A) and control for corruption index (in Panel B) in areas of partitioned ethnicities across the national border where there are large differences in national institutional quality. The dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. In columns (3)-(4) we include a third-order RD polynomial in distance of the centroid of each partition to the national border; distance takes positive values for partitioned areas that fall into the relatively high institutional quality country and negative values for partitioned areas that fall into the relatively low institutional quality country. These specifications also include interactions between the polynomial terms and the proxy measure of national institutions.

Estimation is performed across the two-major partitions of each group. In columns (5)-(6) we report local linear regression estimates focusing on ethnic areas within 50 kilometers of each side of the national border (total 100 kilometers). In both Panel A and Panel B odd-numbered columns report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density). 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 Table 3: Contemporary National Institutions and Regional Development across and within Partitioned Ethnic Groups

RD Specifications (Global Polynomial Control Function Approach) with the Ethnologue Database

	Two-Way Partitioned Ethnic Groups				Two Major Partitions of All Split Groups			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	0.7051***	0.1360			0.7855***	0.1668		
	(0.2303)	(0.2474)			(0.2574)	(0.2312)		
Control for Corruption			0.7305**	0.1563			0.8447**	0.1694
			(0.3310)	(0.3097)			(0.3521)	(0.2633)
Adjusted R-squared	0.254	0.822	0.247	0.821	0.275	0.830	0.269	0.829
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	444	444	444	444	526	526	526	526

Appendix Table 3 reports cross-sectional and within-ethnicity estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1), (2), (5) and (6)) and control for corruption index (in columns (3), (4), (7) and (8)) in areas of partitioned ethnicities using the spatial distribution of ethnicities across Africa according to the Ethnologue database. Odd-numbered columns report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density). All specifications include a third-order RD polynomial in distance of the centroid of each partition to the national border; distance takes positive values for partitioned areas that fall into the relatively high institutional quality country and negative values for partitioned areas that fall into the relatively low institutional quality country. These specifications also include interactions between the polynomial terms and the proxy measure of national institutions. In columns (1), (2), (3), and (4) we focus on two-way partitioned ethnic groups.

In columns (5), (6), (7), and (8) we focus on the two major partitions of all split ethnic groups.

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 language group dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Supplementary Appendix Table 4: Contemporary National Institutions and Regional Development across and within Partitioned Ethnic Groups

Simple Cross-Sectional and Within-Ethnicity Specifications with Afrobarometer's Individual-level Data

	Household Access to Piped Water				Some Formal Schooling			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rule of Law	0.2836**	-0.1421			0.0364	-0.0675		
	(0.1373)	(0.1559)			(0.0473)	(0.1140)		
Control for Corruption			0.3483***	-0.1226			0.0429	-0.1170
			(0.0816)	(0.1725)			(0.0428)	(0.1175)
adjusted R-squared	0.049	0.189	0.098	0.188	0.002	0.076	0.003	0.078
Ethnicity Fixed-Effects	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4336	4336	4336	4336	4496	4496	4496	4496

Appendix Table 4 reports cross-sectional and within-ethnicity estimates associating regional development with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in columns (1), (2), (5), and (6)) and control for corruption index (in columns (3), (4), (7), and (8)) in areas of partitioned ethnicities using individual data from the Afrobarometer. In columns (1)-(4) the dependent variable is an indicator variable that takes on the value one if an individual located in a partitioned ethnic area reports having access to piped water and zero otherwise. In columns (5)-(8) the dependent variable is an indicator variable that takes on the value one if an individual located in a partitioned ethnic area reports having some formal education and zero otherwise. Odd-numbered columns report cross-sectional specifications. Even-numbered columns report within-ethnicity estimates, where we include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log (0.01 + population density).

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 Table 5: Pre-colonial Ethnic Institutions and Regional Development within African Countries Excluding North Africa

	Baseline Estimates				Contiguous-Ethnic-Homeland Analysis			
							Within the Same Country	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jurisdictional Hierarchy	0.1841**	0.1517**	0.2136***	0.1864***	0.1181**	0.1417***	0.1184**	0.1393**
	(0.0895)	(0.0712)	(0.0553)	(0.0528)	(0.0522)	(0.0534)	(0.0562)	(0.0580)
adjusted R-squared	0.282	0.416	0.530	0.609	0.31	0.329	0.291	0.316
Country Fixed Effects	No	No	Yes	Yes	No	No	No	No
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	640	640	640	640	243	243	178	178

Appendix Table 5 reports OLS estimates associating regional development with pre-colonial ethnic institutions, as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community index in ethnic homelands located in Sub-Saharan African countries (excluding ethnic areas in North Africa). In columns (1)-(4) the dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. The specifications in columns (3) and (4) include a set of country fixed effects (constants not reported). In columns (1)-(4) we control for log (0.01 + population density) at the ethnicity-country level. In columns (2) and (4) we also control for: the distance of the centroid of each ethnicity-country area from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. In columns (5)-(8) we examine how differences in pre-colonial ethnic institutions as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community level (that ranges from 0 to 4) affect regional development (as reflected in differences in log (0.01 + satellite light density) between two adjacent ethnic homelands (of ethnicities i and j). In columns (5)-(6) we compare centralized to non-centralized adjacent ethnic groups. In columns (7)-(8) we compare centralized to non-centralized adjacent ethnic groups that are in the same country. In columns (6) and (8) we control for differences in the location and differences in geography between the two adjacent ethnic homelands. The set of control variables includes the differences of the two adjacent ethnic homelands in distance of the centroid of each ethnicity from the capital city of each country, the distance from the closest sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), 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 dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Supplementary Appendix Table 6: Pre-colonial Ethnic Institutions and Regional Development within African Countries Excluding Ethnic Homelands where Capital Cities Fall

	Baseline Estimates				Contiguous-Ethnic-Homeland Analysis			
							Within the Same Country	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jurisdictional Hierarchy	0.3145**	0.1960**	0.1975***	0.1811***	0.1015*	0.1218**	0.1272**	0.1309**
	(0.1356)	(0.0774)	(0.0622)	(0.0568)	(0.0536)	(0.0529)	(0.0539)	(0.0581)
adjusted R-squared	0.212	0.442	0.589	0.657	0.248	0.273	0.240	0.274
Country Fixed-Effects	No	No	Yes	Yes	No	No	No	No
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	645	645	645	645	216	216	160	160

Appendix Table 6 reports OLS estimates associating regional development with pre-colonial ethnic institutions, as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community level excluding ethnic areas where capital cities fall. In columns (1)-(4) the dependent variable is the log (0.01 + light density at night from satellite) at the ethnicity-country level. The specifications in columns (3) and (4) include a set of country fixed effects (constants not reported). In columns (1)-(4) we control for log (0.01 + population density) at the ethnicity-country level. In columns (2) and (4) we also control the distance of the centroid of each ethnicity-country area from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. In columns (5)-(8) we examine how differences in pre-colonial ethnic institutions as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community level (that ranges from 0 to 4) affect regional development (as reflected in differences in log (0.01 + satellite light density) between two adjacent ethnic homelands (of ethnicities i and j) In columns (5)-(6) we compare centralized to non-centralized adjacent ethnic groups.
In columns (7)-(8) we compare centralized to non-centralized adjacent ethnic groups that are in the same country. In columns (6) and (8) we control for differences in the location and differences in geography between the two adjacent ethnic homelands. The set of control variables includes the differences of the two adjacent ethnic homelands in: distance of the centroid of each ethnicity from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), 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.

Supplementary Appendix Table 7: Pre-colonial Ethnic Institutions and Regional Population Density

	Baseline Estimates				Contiguous-Ethnic-Homeland Analysis			
							Within the Same Country	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jurisdictional Hierarchy	0.2153*	0.3088***	0.2550**	0.2469**	0.2356***	0.2465***	0.1693**	0.1975***
	(0.1105)	(0.0831)	(0.1189)	(0.1129)	(0.0723)	(0.0579)	(0.0824)	(0.0714)
Adjusted R-squared	0.012	0.386	0.391	0.548	0.070	0.245	0.039	0.211
Country Fixed-Effects	No	No	Yes	Yes	No	No	No	No
Location Controls	No	Yes	No	Yes	No	Yes	Yes	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	Yes	Yes
Observations	676	676	676	676	258	258	186	186

Appendix Table 7 reports OLS estimates associating regional development proxied by population density with pre-colonial ethnic institutions, as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community level. In columns (1)-(4) the dependent variable is the log (0.01 + population density). The specifications in columns (3) and (4) include a set of country fixed effects (constants not reported). In columns (2) and (4) we also control for the distance of the centroid of each ethnicity-country from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. In columns (5)-(8) we examine how differences in pre-colonial ethnic institutions as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local community level (that ranges from 0 to 4) affect regional development (as reflected in differences in log (0.01 + population density) between two adjacent ethnic homelands (of ethnicities i and j). In columns (5)-(6) we compare centralized to non-centralized adjacent ethnic groups.

In columns (7)-(8) we compare centralized to non-centralized adjacent ethnic groups that are in the same country. In columns (6) and (8) we control for differences in the location and differences in geography between the two adjacent ethnic homelands. The set of control variables includes the differences of the two adjacent ethnic homelands in: distance of the centroid of each ethnicity from the capital city of each country, the distance from the sea coast, the distance from the national border, log (1 + area under water (lakes, rivers, and other streams)), 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.

Appendix Table 8: Partitioned Ethnicities

Ethnicity Name	% of Initial Homeland	Country	# of Partitions	Ethnicity Name	% of Initial Homeland	Country	# of Partitions
ABABDA	0.72	EGY	2	LAKA (ADAMAWA)	0.69	TCD	3
ABABDA	0.28	SDN	2	LAKA (ADAMAWA)	0.20	CMR	3
ADELE	0.48	GHA	2	LAKA (ADAMAWA)	0.11	CAF	3
ADELE	0.52	TGO	2	LAMBA	0.39	ZAR	2
AFAR	0.17	DJI	3	LAMBA	0.61	ZMB	2
AFAR	0.22	ERI	3	LAMBYA	0.17	MWI	3
AFAR	0.61	ETH	3	LAMBYA	0.33	TZA	3
ALUR	0.16	ZAR	2	LAMBYA	0.50	ZMB	3
ALUR	0.84	UGA	2	LIGBI, DEGHA (SE)	0.72	GHA	2
AMBA	0.87	ZAR	2	LIGBI, DEGHA (SE)	0.28	CIV	2
AMBA	0.13	UGA	2	LOBI	0.42	CIV	2
AMBO	0.41	AGO	2	LOBI	0.58	BFA	2
AMBO	0.59	NAM	2	LUGBARA	0.45	ZAR	3
AMER	0.56	ERI	2	LUGBARA	0.04	SDN	3
AMER	0.44	SDN	2	LUGBARA	0.51	UGA	3
ANA	0.33	BEN	2	LUNGU	0.31	TZA	2
ANA	0.67	TGO	2	LUNGU	0.69	ZMB	2
ANUAK	0.75	ETH	2	LUVALE	0.81	AGO	3
ANUAK	0.25	SDN	2	LUVALE	0.01	ZAR	3
ANYI	0.42	GHA	2	LUVALE	0.17	ZMB	3
ANYI	0.58	CIV	2	MADI	0.42	SDN	2
ASBEN	0.89	NER	2	MADI	0.58	UGA	2
ASBEN	0.11	DZA	2	MAKONDE	0.56	MOZ	2
ASSINI	0.51	GHA	2	MAKONDE	0.44	TZA	2
ASSINI	0.49	CIV	2	MALINKE	0.03	GMB	6
ATTA	0.51	MAR	2	MALINKE	0.13	CIV	6
ATTA	0.49	DZA	2	MALINKE	0.27	MLI	6
ATYUTI	0.13	GHA	2	MALINKE	0.04	GNB	6
ATYUTI	0.87	TGO	2	MALINKE	0.25	GIN	6
AULLIMINDEN	0.55	MLI	3	MALINKE	0.29	SEN	6
AULLIMINDEN	0.40	NER	3	MAMBILA	0.57	CMR	2
AULLIMINDEN	0.05	DZA	3	MAMBILA	0.43	NGA	2
AUSHI	0.27	ZAR	2	MANDARA	0.35	CMR	2
AUSHI	0.73	ZMB	2	MANDARA	0.65	NGA	2
AVATIME	0.51	GHA	2	MANGA	0.60	NER	2
AVATIME	0.49	TGO	2	MANGA	0.40	NGA	2
AZANDE	0.62	ZAR	3	MANYIKA	0.39	MOZ	2
AZANDE	0.15	CAF	3	MANYIKA	0.61	ZWE	2
AZANDE	0.23	SDN	3	MASAI	0.38	KEN	2
AZJER	0.24	LBY	3	MASAI	0.62	TZA	2
AZJER	0.00	NER	3	MASALIT	0.13	TCD	2
AZJER	0.75	DZA	3	MASALIT	0.87	SDN	2

BABUKUR	0.82	ZAR	2	MASHI	0.12	AGO	2
BABUKUR	0.18	SDN	2	MASHI	0.88	ZMB	2
BAJUN	0.37	KEN	2	MASINA	0.82	MLI	3
BAJUN	0.63	SOM	2	MASINA	0.09	BFA	3
BALANTE	0.73	GNB	2	MASINA	0.09	MRT	3
BALANTE	0.27	SEN	2	MATAKAM	0.70	CMR	2
BANYUN	0.48	GNB	2	MATAKAM	0.30	NGA	2
BANYUN	0.52	SEN	2	MBERE	0.02	TCD	3
BANZIRI	0.14	ZAR	2	MBERE	0.24	CMR	3
BANZIRI	0.86	CAF	2	MBERE	0.74	CAF	3
BARABRA	0.31	EGY	2	MBUKUSHU	0.74	AGO	3
BARABRA	0.69	SDN	2	MBUKUSHU	0.15	BWA	3
BARARETTA	0.18	ETH	3	MBUKUSHU	0.12	NAM	3
BARARETTA	0.44	KEN	3	MBUNDA	0.89	AGO	2
BARARETTA	0.38	SOM	3	MBUNDA	0.11	ZMB	2
BARGU	0.77	BEN	4	MENDE	0.18	LBR	3
BARGU	0.03	NER	4	MENDE	0.82	SLE	3
BARGU	0.19	NGA	4	MINIANKA	0.01	CIV	3
BARGU	0.02	BFA	4	MINIANKA	0.72	MLI	3
BASHI	0.09	BDI	3	MINIANKA	0.27	BFA	3
BASHI	0.83	ZAR	3	MOMBERA	0.72	MWI	2
BASHI	0.08	RWA	3	MOMBERA	0.28	ZMB	2
BATA	0.29	CMR	2	MPEZENI	0.11	MWI	2
BATA	0.71	NGA	2	MPEZENI	0.89	ZMB	2
BAYA	0.20	CMR	2	MUNDANG	0.80	TCD	2
BAYA	0.80	CAF	2	MUNDANG	0.20	CMR	2
BERABISH	0.80	MLI	2	MUNDU	0.30	ZAR	2
BERABISH	0.20	MRT	2	MUNDU	0.70	SDN	2
BERTA	0.75	ETH	2	MUSGU	0.76	TCD	2
BERTA	0.25	SDN	2	MUSGU	0.24	CMR	2
BIDEYAT	0.21	LBY	4	NAFANA	0.74	GHA	2
BIDEYAT	0.40	TCD	4	NAFANA	0.26	CIV	2
BIDEYAT	0.03	EGY	4	NALU	0.41	GNB	2
BIDEYAT	0.36	SDN	4	NALU	0.59	GIN	2
BIRIFON	0.52	GHA	3	NAMA	0.18	ZAF	2
BIRIFON	0.47	BFA	3	NAMA	0.82	NAM	2
BOBO	0.20	MLI	2	NAUDEBA	0.87	BEN	2
BOBO	0.80	BFA	2	NAUDEBA	0.13	TGO	2
BOKI	0.22	CMR	2	NDAU	0.86	MOZ	2
BOKI	0.78	NGA	2	NDAU	0.14	ZWE	2
BONDJO	0.14	ZAR	2	NDEMBU	0.26	AGO	3
BONDJO	0.86	COG	2	NDEMBU	0.39	ZAR	3
BONI	0.67	KEN	2	NDEMBU	0.35	ZMB	3
BONI	0.33	SOM	2	NDOGO	0.01	ZAR	3
BORAN	0.46	ETH	2	NDOGO	0.18	CAF	3
BORAN	0.54	KEN	2	NDOGO	0.81	SDN	3
BRONG	0.84	GHA	2	NDUKA	0.23	TCD	2
BRONG	0.16	CIV	2	NDUKA	0.77	CAF	2
BUEM	0.40	GHA	2	NGAMA	0.30	TCD	2

BUEM	0.60	TGO	2	NGAMA	0.70	CAF	2
BULOM	0.85	SLE	2	NGERE	0.65	CIV	3
BULOM	0.15	GIN	2	NGERE	0.29	LBR	3
BUSA	0.14	BEN	2	NGERE	0.06	GIN	3
BUSA	0.86	NGA	2	NGUMBA	0.65	CMR	2
BWAKA	0.81	ZAR	3	NGUMBA	0.35	GNQ	2
BWAKA	0.15	CAF	3	NGWAKETSE	0.86	BWA	2
BWAKA	0.04	COG	3	NGWAKETSE	0.14	ZAF	2
CHAGA	0.24	KEN	2	NSENGA	0.15	MOZ	3
CHAGA	0.76	TZA	2	NSENGA	0.78	ZMB	3
CHAKOSSI	0.27	GHA	2	NSENGA	0.06	ZWE	3
CHAKOSSI	0.73	TGO	2	NSUNGLI	0.78	CMR	2
CHEWA	0.34	MWI	3	NSUNGLI	0.22	NGA	2
CHEWA	0.50	MOZ	3	NUKWE	0.44	AGO	4
CHEWA	0.16	ZMB	3	NUKWE	0.24	BWA	4
CHIGA	0.12	RWA	3	NUKWE	0.05	ZMB	4
CHIGA	0.87	UGA	3	NUKWE	0.26	NAM	4
CHOKWE	0.81	AGO	2	NUSAN	0.30	BWA	3
CHOKWE	0.19	ZAR	2	NUSAN	0.37	ZAF	3
COMORIANS	0.82	COM	2	NUSAN	0.33	NAM	3
COMORIANS	0.18	MYT	2	NYAKYUSA	0.12	MWI	2
DAGARI	0.67	GHA	2	NYAKYUSA	0.88	TZA	2
DAGARI	0.33	BFA	2	NYANGIYA	0.17	SDN	2
DARI	0.78	TCD	2	NYANGIYA	0.83	UGA	2
DARI	0.22	CMR	2	NYANJA	0.64	MWI	2
DAZA	0.27	TCD	2	NYANJA	0.36	MOZ	2
DAZA	0.73	NER	2	NYASA	0.05	MWI	3
DELIM	0.55	ESH	2	NYASA	0.68	MOZ	3
DELIM	0.45	MRT	2	NYASA	0.27	TZA	3
DENDI	0.60	BEN	3	NZANKARA	0.14	ZAR	2
DENDI	0.39	NER	3	NZANKARA	0.86	CAF	2
DIALONKE	0.36	MLI	3	PANDE	0.38	CAF	2
DIALONKE	0.58	GIN	3	PANDE	0.62	COG	2
DIALONKE	0.06	SEN	3	POPO	0.72	BEN	2
DIDINGA	0.04	KEN	3	POPO	0.28	TGO	2
DIDINGA	0.89	SDN	3	PUKU	0.31	CMR	3
DIDINGA	0.07	UGA	3	PUKU	0.49	GNQ	3
DIGO	0.62	KEN	2	PUKU	0.19	GAB	3
DIGO	0.38	TZA	2	REGEIBAT	0.34	ESH	2
DIOLA	0.14	GMB	3	REGEIBAT	0.66	MRT	2
DIOLA	0.07	GNB	3	RESHIAT	0.83	ETH	3
DIOLA	0.78	SEN	3	RESHIAT	0.06	KEN	3
DUMA	0.63	GAB	2	RESHIAT	0.11	SDN	3
DUMA	0.37	COG	2	RONGA	0.60	MOZ	3
DZEM	0.74	CMR	3	RONGA	0.35	ZAF	3
DZEM	0.03	GAB	3	RONGA	0.05	SWZ	3
DZEM	0.24	COG	3	RUANDA	0.02	BDI	5
EGBA	0.41	BEN	3	RUANDA	0.06	ZAR	5
EGBA	0.52	NGA	3	RUANDA	0.89	RWA	5

EGBA	0.07	TGO	3	RUANDA	0.02	TZA	5
EKOI	0.38	CMR	2	RUANDA	0.02	UGA	5
EKOI	0.62	NGA	2	RUNDI	0.76	BDI	4
ESA	0.03	DJI	3	RUNDI	0.04	RWA	4
ESA	0.52	ETH	3	RUNDI	0.20	TZA	4
ESA	0.44	SOM	3	RUNGA	0.74	TCD	3
EWE	0.44	GHA	2	RUNGA	0.26	CAF	3
EWE	0.56	TGO	2	SABEI	0.56	KEN	2
FANG	0.37	CMR	4	SABEI	0.44	UGA	2
FANG	0.07	GNQ	4	SAHO	0.43	ERI	2
FANG	0.54	GAB	4	SAHO	0.57	ETH	2
FANG	0.02	COG	4	SAMO	0.12	MLI	2
FON	0.86	BEN	3	SAMO	0.88	BFA	2
FON	0.14	TGO	3	SANGA	0.26	CMR	3
FOUTADJALON	0.01	MLI	4	SANGA	0.19	CAF	3
FOUTADJALON	0.11	GNB	4	SANGA	0.55	COG	3
FOUTADJALON	0.88	GIN	4	SEKE	0.34	GNQ	2
FOUTADJALON	0.01	SEN	4	SEKE	0.66	GAB	2
FUNGON	0.81	CMR	2	SHAMBALA	0.10	KEN	2
FUNGON	0.19	NGA	2	SHAMBALA	0.90	TZA	2
GADAMES	0.25	LBY	3	SHEBELLE	0.58	ETH	2
GADAMES	0.27	TUN	3	SHEBELLE	0.42	SOM	2
GADAMES	0.48	DZA	3	SHUWA	0.62	TCD	3
GIL	0.80	MAR	2	SHUWA	0.17	CMR	3
GIL	0.20	DZA	2	SHUWA	0.21	NGA	3
GOMANI	0.86	MWI	2	SONGHAI	0.57	MLI	3
GOMANI	0.14	MOZ	2	SONGHAI	0.36	NER	3
GREBO	0.33	CIV	2	SONGHAI	0.07	BFA	3
GREBO	0.67	LBR	2	SONINKE	0.68	MLI	3
GRUNSHI	0.68	GHA	2	SONINKE	0.03	SEN	3
GRUNSHI	0.32	BFA	2	SONINKE	0.29	MRT	3
GUDE	0.83	CMR	2	SOTHO	0.24	LSO	2
GUDE	0.17	NGA	2	SOTHO	0.76	ZAF	2
GULA	0.61	TCD	2	SUBIA	0.11	BWA	4
GULA	0.39	CAF	2	SUBIA	0.53	ZMB	4
GUN	0.48	BEN	2	SUBIA	0.06	ZWE	4
GUN	0.52	NGA	2	SUBIA	0.30	NAM	4
GURENSI	0.74	GHA	3	SUNDI	0.37	ZAR	2
GURENSI	0.13	TGO	3	SUNDI	0.63	COG	2
GURENSI	0.13	BFA	3	SURI	0.71	ETH	2
GURMA	0.15	BEN	4	SURI	0.29	SDN	2
GURMA	0.12	NER	4	SWAZI	0.45	ZAF	2
GURMA	0.01	TGO	4	SWAZI	0.55	SWZ	2
GURMA	0.72	BFA	4	TABWA	0.57	ZAR	2
GUSII	0.53	KEN	2	TABWA	0.43	ZMB	2
GUSII	0.47	TZA	2	TAJAKANT	0.15	MAR	4
HAMAMA	0.80	TUN	2	TAJAKANT	0.14	ESH	4
HAMAMA	0.20	DZA	2	TAJAKANT	0.66	DZA	4
HAUSA	0.14	NER	2	TAJAKANT	0.05	MRT	4

HAUSA	0.86	NGA	2	TAMA	0.30	TCD	2
HIECHWARE	0.81	BWA	2	TAMA	0.70	SDN	2
HIECHWARE	0.19	ZWE	2	TAWARA	0.57	MOZ	2
HLENGWE	0.82	MOZ	3	TAWARA	0.43	ZWE	2
HLENGWE	0.00	ZAF	3	TEDA	0.34	LBY	3
HLENGWE	0.18	ZWE	3	TEDA	0.35	TCD	3
HOLO	0.84	AGO	2	TEDA	0.31	NER	3
HOLO	0.16	ZAR	2	TEKE	0.31	ZAR	3
IBIBIO	0.11	CMR	2	TEKE	0.03	GAB	3
IBIBIO	0.89	NGA	2	TEKE	0.66	COG	3
IFORA	0.30	MLI	2	TEKNA	0.53	MAR	2
IFORA	0.70	DZA	2	TEKNA	0.47	ESH	2
IMRAGEN	0.10	MAR	3	TEM	0.17	BEN	2
IMRAGEN	0.74	ESH	3	TEM	0.83	TGO	2
IMRAGEN	0.16	MRT	3	TENDA	0.57	GIN	2
ISHAAK	0.20	ETH	2	TENDA	0.43	SEN	2
ISHAAK	0.80	SOM	2	THONGA	0.58	MOZ	3
IWA	0.33	TZA	2	THONGA	0.42	ZAF	3
IWA	0.67	ZMB	2	TIENGA	0.22	NER	3
JERID	0.90	TUN	2	TIENGA	0.78	NGA	3
JERID	0.10	DZA	2	TIGON	0.32	CMR	2
JIE	0.24	KEN	2	TIGON	0.68	NGA	2
JIE	0.76	UGA	2	TIGRINYA	0.51	ERI	3
KABRE	0.39	BEN	2	TIGRINYA	0.44	ETH	3
KABRE	0.61	TGO	2	TIGRINYA	0.05	SDN	3
KANEMBU	0.73	TCD	3	TLOKWA	0.14	BWA	3
KANEMBU	0.25	NER	3	TLOKWA	0.77	ZAF	3
KANEMBU	0.02	NGA	3	TLOKWA	0.09	ZWE	3
KAONDE	0.21	ZAR	2	TOMA	0.29	LBR	2
KAONDE	0.79	ZMB	2	TOMA	0.71	GIN	2
KAPSIKI	0.65	CMR	2	TONGA	0.84	ZMB	2
KAPSIKI	0.35	NGA	2	TONGA	0.16	ZWE	2
KARA	0.85	CAF	2	TRIBU	0.25	GHA	2
KARA	0.15	SDN	2	TRIBU	0.75	TGO	2
KARAMOJONG	0.27	KEN	2	TRIPOLITANIANS	0.74	LBY	2
KARAMOJONG	0.73	UGA	2	TRIPOLITANIANS	0.26	TUN	2
KARE	0.75	ZAR	2	TUBURI	0.25	TCD	2
KARE	0.25	CAF	2	TUBURI	0.75	CMR	2
KGATLA	0.13	BWA	2	TUKULOR	0.39	SEN	2
KGATLA	0.87	ZAF	2	TUKULOR	0.61	MRT	2
KISSI	0.12	LBR	3	TUMBUKA	0.74	MWI	2
KISSI	0.02	SLE	3	TUMBUKA	0.26	ZMB	2
KISSI	0.86	GIN	3	TUNISIANS	0.87	TUN	2
KOBA	0.89	BWA	2	TUNISIANS	0.13	DZA	2
KOBA	0.11	NAM	2	UDALAN	0.82	MLI	3
KOMA	0.57	ETH	2	UDALAN	0.05	NER	3
KOMA	0.43	SDN	2	UDALAN	0.13	BFA	3
KOMONO	0.49	CIV	2	VAI	0.76	LBR	2
KOMONO	0.51	BFA	2	VAI	0.24	SLE	2

KONGO	0.77	AGO	3	VENDA	0.70	ZAF	2
KONGO	0.23	ZAR	3	VENDA	0.30	ZWE	2
KONJO	0.81	ZAR	2	VILI	0.20	AGO	4
KONJO	0.19	UGA	2	VILI	0.22	ZAR	4
KONKOMBA	0.24	GHA	2	VILI	0.11	GAB	4
KONKOMBA	0.76	TGO	2	VILI	0.47	COG	4
KONO	0.74	SLE	2	WAKURA	0.28	CMR	2
KONO	0.26	GIN	2	WAKURA	0.72	NGA	2
KONYANKE	0.30	CIV	2	WANGA	0.79	KEN	2
KONYANKE	0.70	GIN	2	WANGA	0.21	UGA	2
KORANKO	0.39	SLE	2	WUM	0.88	CMR	2
KORANKO	0.61	GIN	2	WUM	0.12	NGA	2
KOTA	0.41	GAB	2	YAKA	0.16	AGO	2
KOTA	0.59	COG	2	YAKA	0.84	ZAR	2
KOTOKO	0.67	TCD	2	YAKOMA	0.40	ZAR	2
KOTOKO	0.33	CMR	2	YAKOMA	0.60	CAF	2
KPELLE	0.48	LBR	3	YALUNKA	0.25	SLE	2
KPELLE	0.52	GIN	3	YALUNKA	0.75	GIN	2
KRAN	0.16	CIV	2	YAO	0.13	MWI	3
KRAN	0.84	LBR	2	YAO	0.65	MOZ	3
KREISH	0.10	CAF	2	YAO	0.22	TZA	3
KREISH	0.90	SDN	2	YOMBE	0.13	AGO	3
KUNDA	0.84	MOZ	3	YOMBE	0.48	ZAR	3
KUNDA	0.15	ZMB	3	YOMBE	0.39	COG	3
KUNG	0.10	BWA	2	ZAGHAWA	0.14	TCD	2
KUNG	0.90	NAM	2	ZAGHAWA	0.86	SDN	2
KUNTA	0.85	MLI	2	ZEKARA	0.83	MAR	2
KUNTA	0.15	DZA	2	ZEKARA	0.17	DZA	2
KWANGARE	0.84	AGO	2	ZIMBA	0.16	MWI	2
KWANGARE	0.16	NAM	2	ZIMBA	0.84	MOZ	2

Appendix Table 8 reports the name of partitioned ethnic groups (as coded by Murdock (1959)) and the percentage of the historical homeland of the split ethnic groups that fall into more than one country. Section 2 gives details on our approach in identifying partitioned ethnicities.