Sep 2013 No.154

National Institutions and Subnational Development in Africa

Stelios Michalopoulos
Brown University and NBER
Elias Papaioannou
London Business School, NBER and CEPR

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

National Institutions and Subnational Development in Africa*

Stelios Michalopoulos
Brown University and NBER

Elias Papaioannou
London Business School, NBER and CEPR

August 10, 2013

Abstract

We investigate the role of national institutions on subnational African development in a novel framework that accounts both for local geography and cultural-genetic traits. We exploit the fact that the political boundaries in the eve of African independence partitioned more than two hundred ethnic groups across adjacent countries subjecting similar cultures, residing in homogeneous geographic areas, to different formal institutions. Using both a matching-type and a spatial regression discontinuity approach we show that differences in countrywide institutional structures across the national border do not explain within-ethnicity differences in economic performance, as captured by satellite images of light density. The average non-effect of national institutions on ethnic development masks considerable heterogeneity partially driven by the diminishing role of national institutions in areas further from the capital cities.

Keywords: Africa, Borders, Ethnicities, Development, National Institutions, Regression Discontinuity
JEL classification Numbers: O10, O40, O43, N17, Z10.

*We thank seminar participants at Dartmouth, Tufts, Oxford, Vienna, Brown, Harvard, Princeton, Stanford, UC-Berkeley, UC-Davis, NYU, AUEB, the CEPR Development Economics Workshop,the CEPR-UPF Workshop on the Political Economy of Development and Conflict, the World Bank, the IMF, the NBER Political Economy Meetings, the NBER Summer Institute Meetings in Economic Growth and Income Distribution and the Macroeconomy for valuable comments. We also benefited from discussions with Yannis Ioannides, Rafael La Porta, Antonio Ciccone, Rob Johnson, Raphael Frank, Jim Feyrer, Ross Levine, Avner Greif, Jeremiah Dittmar, David Weil, Sandip Sukhtankar, Quamrul Ashraf, Oded Galor, Ed Kutsoati, Pauline Grosjean, Hans-Joachim Voth, Enrico Perotti, Pedro Dal Bo, Nathan Nunn, Raquel Fernandez, Jim Robinson, and Enrico Spolaore. We are particularly thankful to Andy Zeitlin, Melissa Dell, Andrei Shleifer, Nico Voightlander, Daron Acemoglu, and seven anonymous referees for detailed comments and useful suggestions. We also thank Nathan Nunn for providing the digitized version of Murdock's Tribal Map of Africa. This paper draws on material from Michalopoulos and Papaioannou ("Divide and Rule or the Rule of the Divided?" 2011). All errors are our sole responsibility.

1 Introduction

Few other issues have received more inquiry in the social sciences than "what are the fundamental determinants of comparative development?" The institutional view asserts that poorly performing institutional structures, such as lack of constraints on the executive, 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). Yet other works downplay the role of formal institutions, emphasizing instead the importance of geographical features, informal cultural norms, genetic, and epidemiological traits (see Spolaore and Wacziarg (2013) for a review).

This paper investigates the role of national institutions on comparative development in Africa, developing a methodology that exploits in a "quasi-experimental" setting the artificial drawing of African borders -that took place in the European capitals in the mid-late 19^{th} century, well before African independence and at a time when Europeans had hardly settled in the regions whose borders were designing. The drawing of colonial boundaries partitioned in the eve of African independence more than 200 ethnicities across two (or more) countries. Taking advantage of this historical accident, we compare economic performance in adjacent regions belonging to the historical homeland of the same ethnic group, but falling in different countries and are thus subject to different formal institutions. This identification strategy accounts both for differences in the natural environment (as the two neighboring regions have similar geography-ecology) and for ethnic-specific cultural and anthropological traits (as people from the same background reside at the two sides of the border).

Our identification strategy entails two major challenges. First, we need to identify in a systematic way ethnic homelands partitioned by national borders. The second challenge derives from the sparsity of economic performance indicators at the local (country-ethnicity) level. On the first issue we exploit anthropological data from the pioneering work of George Peter Murdock (1959), who has produced a map portraying the spatial distribution of African ethnicities in the mid/late 19th century (Figure 1a). Projecting Murdock's map on contemporary national boundaries (Figure 1b), allows us to identify in a systematic manner ethnic groups that have been split by the national borders. On the second issue, to overcome the paucity of economic data across African regions we build on the recent contribution of Henderson, Storeygard, and Weil (2012) and measure development at the ethnicity-country level using satellite images of

light density which are available at a fine grid.

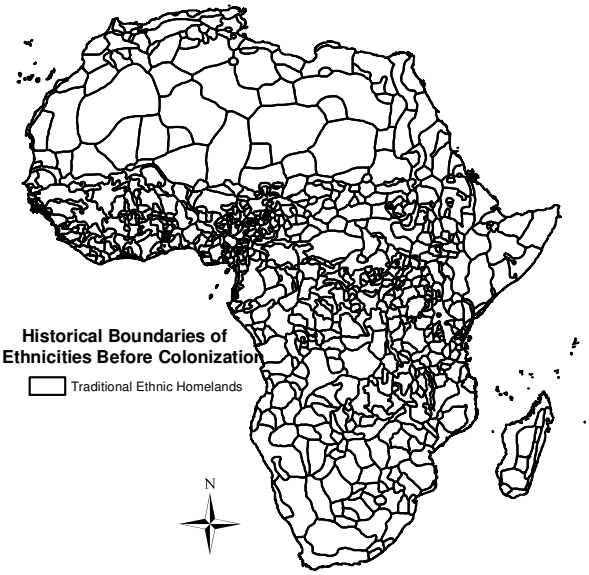

Figure **1a**: Ethnic Boundaries

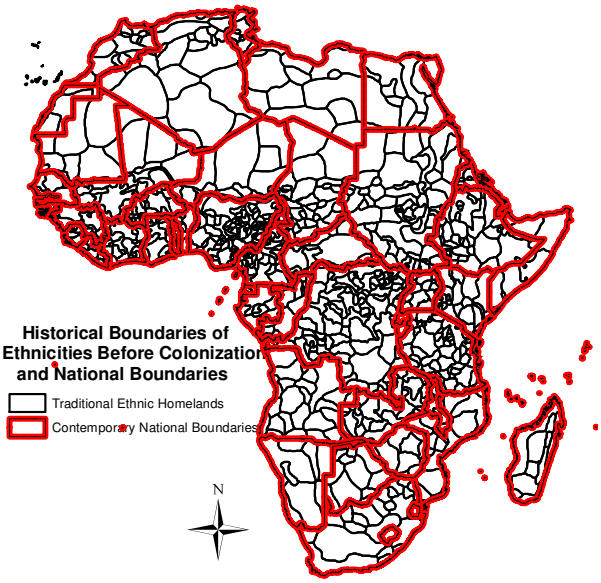

Figure **1b**: Ethnic and Country Boundaries

Figures $2a - 2b$ illustrate our approach using as example the Ambo, an ethnic group partitioned between Angola (a country scoring quite low in most proxies of national institutions) in the North and Namibia (an institutionally advanced African country) in the South. In our empirical analysis we investigate whether differences in national institutions between adjacent countries explain differences in economic performance (as proxied by luminosity) within the historical homeland of the same ethnic group. We perform our analysis at two levels of aggregation; across partitioned ethnic areas in each country (Figure $2a$) and across pixels falling in the homeland of split groups (Figure $2b$). For the Ambo luminosity is higher in Namibia, the more institutionally developed country.

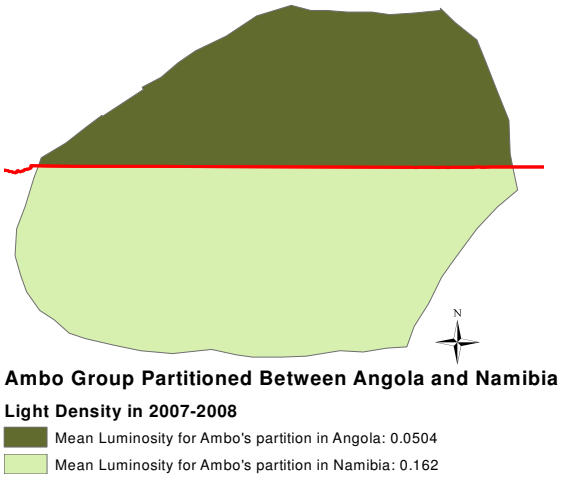

Figure **2a**

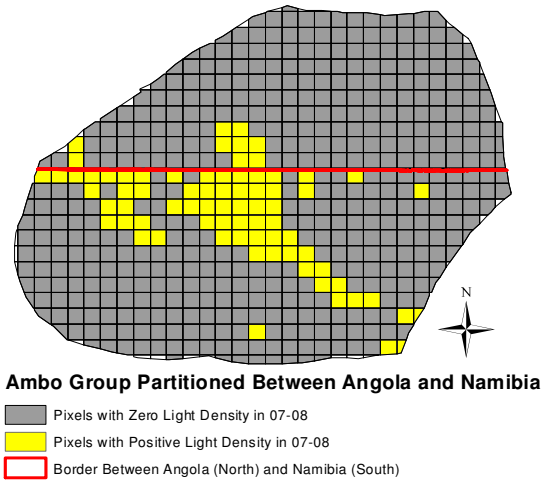

Figure **2b**

Yet, this is not the case for the Kaonde and the Anyi groups split by the Zambia - Democratic Republic of Congo border and the Ghana - Ivory Coast border, respectively. As Figures 2*c* and 2*d* show, a larger faction of the ethnic homeland of both groups is lit in the Democratic Republic of Congo and in Ivory Coast, although national institutional quality is significantly higher at the other side of the border (in Zambia and Ghana, respectively).

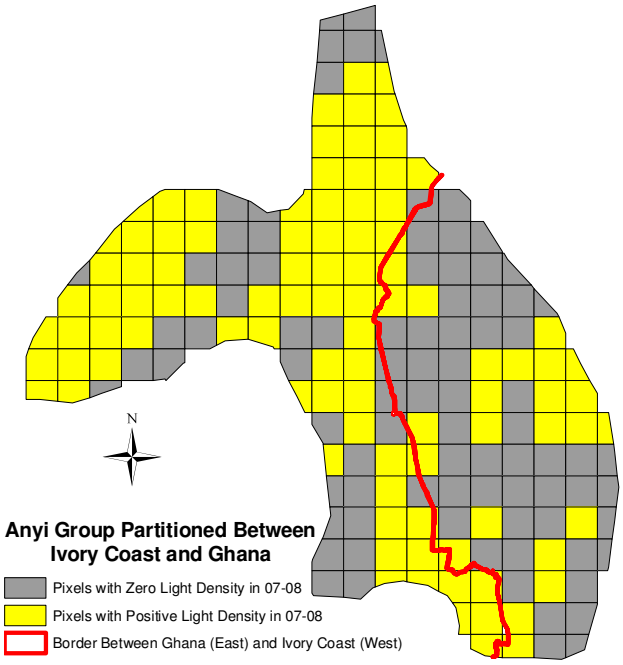

Figure **2c**

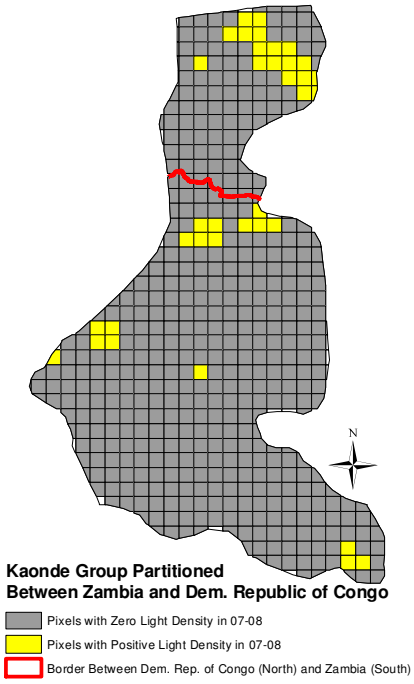

Figure **2d**

Results Summary Our analysis that spans all African partitioned ethnicities reveals new empirical regularities. First, we document that differences in national institutions across the border do not systematically translate into differences in economic performance within partitioned ethnicities. While there is a significant positive correlation between national institutions and development across ethnic homelands, once we properly account for geographic-ecological and ethnic-specific differences via the inclusion of ethnicity fixed effects, the correlation weakens considerably and becomes statistically indistinguishable from zero. This pattern obtains both when the unit of analysis is a partitioned ethnic region (see Figure 2*a*) and when we take advantage of the finer structure of the luminosity data to obtain multiple observations within each partition (see Figures 2*b*, *c*, *d*); moreover when we conduct a spatial regression discontinuity (RD) analysis that identifies the (local) effect of national institutions at the border we also find a close-to-zero effect. The lack of a systematic association between national

institutions and regional development within partitioned African homelands cautions against extrapolating from the positive cross-country correlations.

Our results go against the conventional wisdom in economics on the causal impact of national institutions on development, in Africa at least; yet they are consistent with the African historiography that de-emphasizes the importance of colonial and contemporary countrywide institutions in the hinterland (e.g., Herbst (2000), Davidson (1992)). Moreover, our findings are in line with recent works that emphasize traits besides formal institutions, such as cultural norms, family ties, or important historical episodes (see Algan and Cahuc (2013), Alesina and Giuliano (2013), and Nunn (2013) for recent reviews).

Second, we show that the insignificant correlation between national institutions and ethnic development masks considerable heterogeneity. We find that the average effect of institutions is economically negligible and statistically insignificant for approximately 60% of partitioned ethnicities. Yet, for some groups, consisting of approximately $20\% - 25\%$ of the sample, a significant positive association emerges, whereas for the remaining ones the within-ethnicity association turns negative. Overall, the uncovered heterogeneity provides a useful reminder that generalizing from the findings of case studies focusing on a single border discontinuity can be quite misleading. Meanwhile, it suggests that case studies may be useful for shedding light on the specific circumstances that allow countrywide policies to shape regional development.

Third, building on insights from the African historiography that stress the inability of states to broadcast power in regions far from the capital (e.g., Herbst (2000), Bates (1983)), we examine the spatial distribution of the uncovered heterogeneity. Contrary to the overall pattern, we find that national institutions do correlate with subnational development, but only when both partitions are close to the respective capital cities. We also present complementary evidence on the limited penetration of national institutions expanding our analysis to the universe of African groups; in particular, we show that the explanatory power of national institutions on regional development decays for ethnic homelands (and pixels within groups) further from the capital centers.

Fourth, we augment our analysis on the differential role of national institutions far from the capital using individual-level data from the Afrobarometer Surveys on law enforcement and self-identification (with the nation versus the ethnicity). Our within-country analysis shows that law enforcement weakens monotonically in areas further from the capitals. Moreover, national (as opposed to ethnic) identification is also inversely related to distance to the capital. These results thus further illustrate the limited penetration of country-wide institutions in the hinterland.

Related Literature Our research nests and advances over several strands of literature that examine the political economy of contemporary comparative development.

First, an influential body of research asserts that through persistence, the institutions that European powers established during colonization are the deep roots of contemporary economic performance (e.g., La Porta *et al.* (1997, 1998); Acemoglu *et al.* (2001, 2002)). Yet despite the implementation of ingenious instrumental-variable approaches employed in the cross-country literature, omitted variables and estimate stability remain major concerns (see Glaeser, La Porta, de Silanes, and Shleifer (2004), La Porta, Lopez-de-Silanes, and Shleifer (2008), and Nunn (2012), among others). By exploiting within-country variation, our study circumvents some inherent limitations of the cross-country framework;[1] moreover, we move beyond average effects and uncover the interplay between national institutions, state presence, and proximity to the capital.

Second, our identification scheme that exploits border discontinuities in institutional arrangements relates to works that study the role of national policies across a particular 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 efforts. Bubb (2012) investigates how differences in de jure property rights between Ghana and Ivory Coast affect development in border areas finding that in spite of large differences in formal institutions, there are no differences in the actual enforcement of property rights. Cogneau and Moradi (2011) examine differences in education and religiosity across the Ghana-Togo border and link them to different colonial investments; whereas Cogneau, Mesplé-Somps, and Spielvogel (2012) examine access to electricity, consumption, and health at the borders of Ivory Coast with Ghana, Mali, and Burkina Faso, finding mixed evidence on the role of economic and institutional development. Finally, Miles (1994) studies the development of the Hausa after their partitioning (at the Niger-Nigeria border), documenting that differences between French and British policies (mainly regarding the role of local chiefs) endured after independence and had long-lasting effects.

Our study, rather than focusing on the role of national features across a single border or within a single group, examines the role of national institutions on development across *all* partitioned ethnicities in Africa. Thus our results are less sensitive to the usual "external validity" arguments readily invoked in the context of case studies or works exploiting a single discontinuity. Moreover, our expansive focus sheds light on the underlying heterogeneity and uncovers the limited penetration of national institutions in areas far from the capital. The

[1] In this regard our work is mostly related to works that exploit within-country variation on institutions, such as Banerjee and Iyer (2005), Iyer (2010), Laeven and Woodruff (2007), De Long and Shleifer (1993), Dell (2010), Grosjean (2011), Tabellini (2010), and Gennaioli, La Porta, Lopez-de-Silanes, and Shleifer (2013).

latter finding is related to parallel works of Campante and Do (2013) and Campante, Do, and Guimaraes (2013), who link capital city isolation with misgovernance and economic efficiency.

Third, our results relate to the large literature in African political economy (see Bates (2008) and Dowden (2008) for eloquent narratives) that stresses features beyond national institutions; these include colonial investments in education, health and infrastructure (e.g., Huillery (2009) and Jedwab and Moradi (2012)), ethnic partitioning (e.g., Englebert, Tarango, and Carter (2002) and Michalopoulos and Papaioannou (2013a)), fractionalization (e.g., Easterly and Levine (1997) and Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003)), inequality across ethnic lines (e.g., Alesina, Michalopoulos, and Papaioannou (2013)), and geography-health (e.g., Nunn and Puga (2012) and Alsan (2012)).

Of most relevance from this body of research is our companion paper (Michalopoulos and Papaioannou (2013c)), where we explore the long-lasting importance of ethnic (as opposed to national) political/institutional traits for contemporary African development. The main finding is that among numerous ethnic-specific features (related, for example, to the type of subsistence economy, occupational specialization, slavery, etc.), the legacy of pre-colonial political centralization appears to be a robust correlate of contemporary regional economic performance (Gennaioli and Rainer (2006, 2007) provide similar cross-country evidence). In contrast to these works that highlight the role of historical persistence at the ethnic level, in this paper we examine the role of national institutions, which have been a dominant theme in the growth literature and in policy circles alike. In addition, our finding that national institutions wield significant explanatory power near the capitals, which rapidly diminishes for regions in the hinterland, highlights the importance of ethnic norms and reveals the coexistence of a dual institutional framework within African countries (Lewis (1954), Migdal (1988)). Moreover, a methodological innovation of the present study is that we identify the role of national institutions within partitioned ethnic groups taking advantage of the arbitrary drawing of African borders and thus accounting for any (un)observable ethnic-specific differences across the diverse African tribal landscape. This has significant advantages over works exploiting cross-cultural variation and is motivated by the standard-textbook introduction on the role of national institutions showing the divergence in economic performance within culturally and geographically homogeneous entities, like North and South Korea, or East and West Germany (see Weil (2008)).

In this regard our paper is related to subsequent work by Pinkovskiy (2013), who studies discontinuities in satellite-recorded light density at night across the globe and links them to country-level economic development -and its main correlates (e.g., human capital, institutional capacity, etc.). In line with our findings, Pinkovskiy (2013) shows that borders of African na-

tions tend to be relatively underdeveloped, with limited penetration of government activities and substantial economic activity concentrated around capital cities.[2] Moreover, our focus is on discontinuities *within* partitioned ethnic homelands; as such we effectively account for both geography-ecology and ethnic-specific features, related to history, social structure, and genetics. In fact, the uncovered weak and insignificant correlation between national institutions and subnational development within geographically and culturally homogeneous territories contributes to an emerging body of work that emphasizes the importance of cultural norms, historical persistence, and human and geographic traits for comparative development (see Diamond (1997), Landes (1998), Guiso, Sapienza, and Zingales (2006), Spolaore and Wacziarg (2009), Putterman and Weil (2010), Ashraf and Galor (2013) and Michalopoulos, Naghavi, and Prarolo (2012), among others).

Paper Structure In the next section we describe how we identify partitioned groups and discuss the luminosity data that we use to proxy subnational development. In section 3 we discuss the estimating framework and detail the identification design. In Section 4 we report the baseline estimates on the average effect of national institutions on economic development within partitioned ethnic homelands. We also present spatial regression discontinuity (RD) estimates that quantify the effect of national institutions at the border. In Section 5 we estimate the role of national institutions separately for each split group mapping the overall heterogeneity. In an effort to understand the latter, in Section 6 we explore how proximity to the capital shapes the explanatory power of national institutions on regional development. In section 7 we summarize discussing avenues for future research.

2 Data

2.1 Identifying Partitioned Ethnicities

Murdock's map (Figure 1*a*) portrays the spatial distribution of ethnicities across Africa in the mid/late 19^{th}; it depicts 826 ethnic areas inhabited upon colonization (in total there are 834 polygons, but 8 regions are classified as "uninhabited"). Intersecting the pre-colonial map with the 2000 Digital Chart of the World (Figure 1*b*) that portrays contemporary national boundaries we obtain $1,247$ country-ethnicity observations (e.g., Ambo in Angola, Ewe in Togo, Zulu in South Africa). We classify an ethnicity as partitioned when at least 10% of the historical homeland belongs to more than one contemporary state. We drop partitioned areas of less than 100 square kilometers as tiny partitions are most likely due to the lack

[2]Note that focusing on Africa is desirable since, in contrast to African borders, in Europe and most other parts of the world national borders have been contested by numerous wars and are the outcome of active state formation and (in)voluntary people's movements.

of precision and projection error. Our procedure identifies 526 partitions that belong to the historical homeland of 227 ethnic groups. We report estimates for 220 partitioned ethnicities (507 country-ethnicity observations), because for six groups (the Ifora, the Kunta, the Berabish, the Atta, the Teda, and the Asben) the local population estimates indicate zero population (in one of the two partitioned areas); moreover, we lose information from the Imragen -which are partitioned between Western Sahara and Mauritania- because we do not have information on national institutions for Western Sahara.

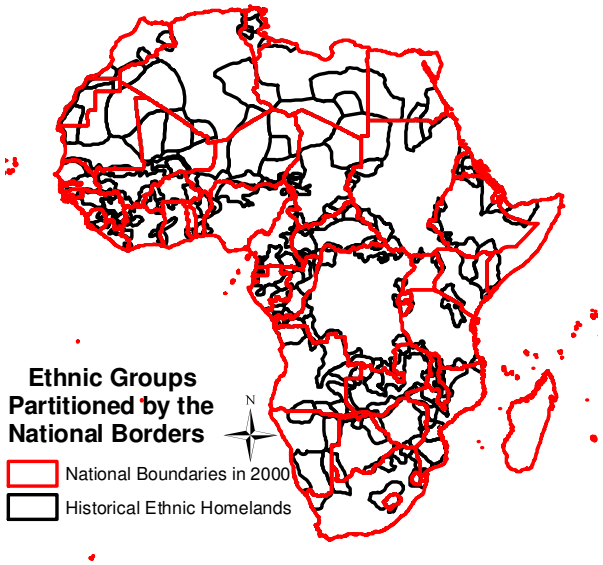

Figure **3a**

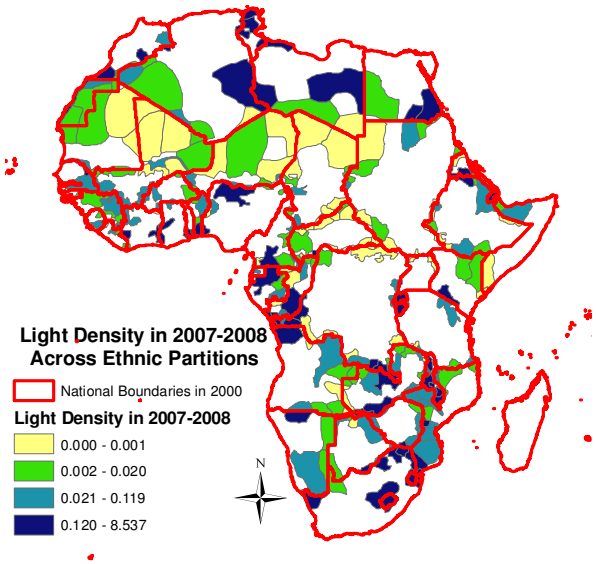

Figure **3b**

Appendix Table *A* lists partitioned ethnicities. Examples include the Ewe that have been partitioned between Ghana and Togo (shares 44% and 56%, respectively), the Esa between Ethiopia and Somalia (shares 52% and 44%, respectively, while a small fraction falls in Djibouti), and the Yao (Wayao) between Mozambique (65%), Malawi (13%), and Tanzania (22%). Our procedure reveals that the median country in Africa has 43% of its population belonging to partitioned ethnicities. This estimate is similar to that of Asiwaju (1985) and Alesina, Easterly, and Matuszeski (2011), who using alternative sources and techniques estimate that on average 40% of the African population belongs to partitioned ethnic groups. Thus our analysis, while focusing on partitioned homelands, captures a significant fraction of the African population (Figure 3*a*).

Case study and anecdotal evidence suggest that in spite of population movements ethnic populations tend to reside in their respective historical homelands. Nunn and Wantchekon (2011) document that close to 55% of respondents in the Afrobarometer Surveys currently live in their ethnic group's ancestral homeland. In the same vein, Glennerster, Miguel, and Rothenberg (2013) 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.

2.2 Satellite Light Density at Night

The nature of our study requires data on economic development at the country-ethnic homeland level. Since there are limited geocoded measures of economic development in Africa, we build on the recent contribution of Henderson, Storeygard, and Weil (2012) and use satellite images on light density to proxy local economic activity.

Data come from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) that reports time-stable images of the earth at night captured between $20:00$ and $21:30$. The measure ranges from 0 to 63 and is available for every 30-second area pixel (approximately 1 square kilometer). The satellite detects lights from human settlements, fires, gas flares, lightning, and the aurora. The annual composite measure is created by overlaying all daily images captured in a year, dropping images where lights are shrouded by cloud cover or overpowered by the aurora or solar glare (near the poles), and after removing ephemeral lights (like fires, lightning and other noise).[3] Using these data we construct light density per square kilometer for 2007 and 2008 averaging across pixels that fall within the historical homeland of each ethnic group in each country.

Besides its availability at a very fine geographic level, luminosity is well suited to spatial analyses of development in Africa for some additional reasons. First, most African countries have low quality income statistics, both at the national and the sub-national level (Jerven (2013)). Second, we lack data on regional incomes for most African countries; and while there are some local proxies of poverty and health, these do not map to our ethnicity level unit of analysis.

Despite a series of works that establish a strong within-country correlation (both across time and across regions) between light density at night and GDP in the on-line Supplementary Appendix (Section 1) we provide further cross-validation checks of the luminosity data in our context.[4] Specifically, using individual-level data from the Afrobarometer Surveys on access to electrification, presence of a sewage system, access to piped water, and education we show that light density correlates strongly with these proxies of public goods and development both

[3] See Henderson, Storeygard, and Weil (2012) and Chen and Nordhaus (2011) for technical details on the lights data. Satellite data on light density at night are subject to overglow/blooming; this happens because lights tend to appear larger than they are over water and snow. While this issue is not particularly important in our application as we do not have observations near the poles and covered with snow, in many specifications we control for water area and distance to the coast. Another issue with the light data comes from top coding - that usually happens in the capitals of developed countries. Yet, in Africa instances of top-coding are rare.

[4] See Henderson, Storeygard, and Weil (2012), Elvidge, Baugh, Kihn, Kroehl, and Davis (1997), Doll, Muller, and Morley (2006), Michalopoulos and Papaioannou (2013c), and Pinkovskiy (2013), among others.

across Afrobarometer enumeration areas within countries and within ethnic homelands as well as across the border within partitioned ethnicities.

2.3 National Institutions

We measure national institutions using data from World Bank's Governance Matters Database (Kaufmann, Kraay, and Mastruzzi (2008)). 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 principal components aiming to minimize measurement error. We mainly 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 range theoretically between -2.5 and $+2.5$ with higher values indicating better functioning institutions and less corruption. To account for reverse causality and measurement error, we use predetermined values of the institutional quality measures, taking the mean over the period $1996-2006$ (the results are robust to the choice of year). In our sample the countries with the lowest rule of law are Somalia (-2.195), the Democratic Republic of Congo (-1.88), and Liberia (-1.73) while South Africa (0.156), Namibia (0.1675), and Botswana (0.615) are the most institutionally developed countries in Africa.

2.4 Data Patterns

There is significant variation both in national institutions and luminosity across African borders (Figure 3*b*). 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 and other Somali tribes reside); and between Gabon and Congo (where the Duma live). Likewise, there are evident changes in luminosity across the border within the homeland of the same ethnic group. On the one hand, despite differences in national institutions, in around 25% of the sample there are negligible differences in light density across the border. On the other hand, in about 40% there are more than one log point differences in luminosity. For example, sizable jumps in luminosity appear in the Egypt-Sudan border (where the Ababda and the Barabra are split), in the border between Ghana and the Ivory Coast (where the Assini reside), and between Benin and Togo (where the Popo are split).

3 Identification

3.1 Estimating Framework

Specification at the Country-Ethnic Homeland Level Our analysis on the relationship between national institutions and regional development is based on variants of the following specification:

$$y_{i,c} = a_0 + \gamma IQL_c + \lambda_1 PD_{i,c} + \lambda_2 AREA_{i,c} + X'_{i,c}\Phi + a_i + \varepsilon_{i,c}. \quad (1)$$

The dependent variable, $y_{i,c}$, reflects the level of economic activity in the historical homeland of ethnic group i in country c, as proxied by light density. Since a significant fraction (around 30%) of the (country-ethnicity) observations takes on the value of zero, we use as dependent variable the log of light density adding a small number ($(y_{i,c} \equiv \ln(0.01+LightDensity_{i,c})$).[5] The logarithmic transformation is useful because we use all observations and because we account for some extreme values in luminosity (outliers).

Specification at the Pixel Level We also exploit the fine structure of the luminosity data to obtain multiple observations (pixels) within a partitioned ethnic area in each country; we do so running the following regression equation (variant of (1)).

$$y_{p,i,c} = a_0 + \gamma IQL_c + \lambda_1 PD_{p,i,c} + \lambda_2 AREA_{p,i,c} + X'_{p,i,c}\Phi + a_i + \zeta_{p,i,c}. \quad (2)$$

The dependent variable, $y_{p,i,c}$, is a dummy that takes on the value one if pixel p is lit and zero otherwise. Each pixel, p, falls in the historical homeland of partitioned ethnicity i located in country c (see Figures $2b-d$ for examples). In our analysis we use pixels of 0.125 x 0.125 decimal degrees (approximately $12.5km$ x $12.5km$). Since there are several unpopulated pixels (in the Sahara or in the rainforests) and to make sure that we examine the role of national institutions on development properly, we exclude pixels with zero population.[6] This results to more than $120,000$ observations across all African ethnic homelands, with $42,710$ of those belonging to partitioned ethnicities.

While conceptually the unit of our analysis is a partition of a group (Figure $2a$), the pixel-level specifications (Figure $2b$) are useful for several reasons. First, by looking at the extensive margin of light density we account for the non-linear nature of luminosity. Second, the linear probability models facilitate the interpretation of the estimates. Third, we are able

[5] 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. In the previous draft of the paper we added one (rather than 0.01) to the luminosity data before taking the logarithm finding similar results.

[6] The coefficient estimates are similar when we also consider unpopulated pixels and/or if we construct pixels of alternative size.

to control for various -relevant for development- characteristics at a fine level. Fourth, it is straightforward to augment the pixel-level specification with RD polynomials on distance to the national border, so as to estimate the effect of national institutions at the border. Fifth, we can estimate the pixel-level specification for each partitioned ethnicity separately and explore potential heterogeneity on the impact of institutions on development.

Independent Variables IQL_c denotes institutional quality of country c, as reflected in the rule of law and the control of corruption measures. For split groups, each partition is assigned to the corresponding country c. For example, regional light density in the part of the Egba in Benin is matched to the institutional quality of Benin, while the adjacent area of the Egba in Nigeria is assigned the value of Nigeria.

In most specifications we control for log population density, $PD_{(p),i,c}$, because the association between luminosity and economic development strengthens and because by doing so the estimates capture the role of institutions beyond population density. We also control for the log of land area, $AREA_{(p),i,c}$; although most pixels are of the same size, those intersected by the coast line, national and ethnic boundaries are smaller (see Figures $2b, c, d$).

Vector $X_{(p),i,c}$ includes additional controls at the ethnic-homeland level (in (1)) or at the pixel level (in (2)).[7] The Data Appendix provides variable definitions and data sources. Table 1 reports summary statistics in the sample of partitioned ethnicities (at the country-ethnic homeland level in Panel A and at the pixel level in Panel B).[8]

Inference In all specifications we employ the approach of Cameron, Gelbach, and Miller (2011) and cluster standard errors along the country and the ethnic-family dimension. Murdock assigns the 826 groups into 96 ethnolinguistic clusters. This double-clustering parameterization accounts for two main concerns related to non-adjusted standard errors. First, within each country we have several ethnic homelands (or pixels) where the country-level rule of law and the control-of-corruption measures take the same value and thus clustering at the

[7] Specifically, we control for land suitability for agriculture and average elevation. The latter has affected African development both via goods and slave trades (see Nunn and Puga (2012)). We also control for surface under water to account for blooming in the luminosity data and for the potential positive effect of water streams on development via trade. Moreover, we control for malaria suitability since many studies document its detrimental effect on development (see Gallup and Sachs (2001)). We also include indicators for the existence of oil and diamonds fields to account for the "natural resource" curse (see Ross (2012)). Finally, we add a set of controls measuring the distance from the capital city, the national border, and the coast, respectively. 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.

[8] In Supplementary Appendix Table 2 we report summary statistics for the full sample of African ethnic homelands, $1,209$ country-ethnic observations and $120,501$ pixels.

country-level is required (Moulton (1990)). Likewise, clustering at the ethnic family accounts for intra-ethnic-family correlation in (un)observed features. Second, the multi-way clustering method allows for arbitrary residual correlation within both dimensions and thus accounts for spatial correlation; Cameron, Gelbach, and Miller (2011) explicitly cite spatial correlation as an application of the multi-clustering approach (see Nunn and Wantchekon (2011) and Michalopoulos and Papaioannou (2013c) for applications of the multi-way clustering method in a similar context). We also estimated standard errors accounting for spatial correlation of an unknown form using Conley's (1999) GMM method. The two approaches yield similar standard errors.

3.2 Validity of Identification Design

For our identification strategy that compares economic development at the homelands of the same ethnicity in adjacent countries to be valid, one needs that border drawing should not have been influenced by local circumstances and factors that themselves shape or reflect economic well-being (see Angrist and Pischke (2008)). Moreover the ethnic areas across the border need to be similar across all relevant for development dimensions. We have thus investigated these issues in detail.

First, all the anecdotal evidence suggests that colonizers drew African borders in an arbitrary manner (see Wesseling (1996), Asiwaju (1985), Herbst (2000), and Englebert (2009) for eloquent reviews); the delineation of African borders took place in a period (late 19th century), when Europeans had not yet settled (most of) Africa and had limited knowledge of its political and economic geography; moreover, at that time colonizers were mostly assigning spheres of influence via the establishment of protectorates, colonies and free-trade-areas, let alone the borders of future African states. Consistent with the historical narratives on the artificiality of African borders, Alesina, Easterly, and Matuszeski (2011) document that close to 80% of African borders follow latitudinal or longitudinal lines, the highest percentage across all continents. Moreover, Michalopoulos and Papaioannou (2013a) show that besides land mass and water area, there are no systematic differences between non-split and partitioned ethnicities, across dozens of potentially relevant for development factors, related to geography, natural resources, the disease environment, and historical traits (capturing the intensity of pre-colonial conflict, political centralization, the type of subsistence economy, the intensity of the slave trades, etc.). Yet, since some borders may have been drawn taking into account local conditions, in the Supplementary Appendix we report estimates restricting attention to ethnicities partitioned by borders that follow relatively straight lines (as identified by the "box-count" method of Alesina, Easterly, and Matuszeski (2011)). The results are virtually

unchanged (see the associated discussion in the Supplementary Appendix and Appendix Tables $9A$ and $9B$).

Second, although research shows that ethnic groups tend to occupy geographically homogenous territories (Michalopoulos (2012)), we investigated whether differences in institutional quality across the border correlate with differences in various characteristics. Table 2 - Panels A and B report ethnicity fixed effects estimates ("similarity" regressions) that associate geographic, ecological, and natural resource measures with the rule of law and the control of corruption index, respectively. In Panels C and D we focus on the two largest partitions of the 220 split groups (440 observations) and examine whether there are systematic differences between regions falling in the relatively high and the relatively low institutional quality country. The results are supportive of our identification strategy, as they point out that the two (or more) partitions of partitioned ethnicities are appropriate counterfactuals. Differences in geography-ecology, location, and natural resources across the border within partitioned ethnic homelands are small and not systematically linked to differences in national institutions.[9]

4 Baseline Results

4.1 Preliminary Evidence. Cross-sectional Analysis.

Before we report the baseline ethnicity fixed effects estimates, we examine the cross-sectional patterns running simple LS specifications that associate regional development with national institutions. Due to reverse causation and omitted variables, these specifications do not identify causal effects; yet they are useful for understanding the raw data patterns. Table 3 - Panel A reports estimates at the country-ethnicity level ($1,209$ observations), while Panel B reports similar estimates at the pixel level ($120,501$ observations). Specification (1) shows that there is a positive and significant correlation between the rule of law index and regional development. In column (2) we control for log population density and log land area, whereas in column (3) we control for distance of each ethnic area (in Panel A) or each pixel (in Panel B) to the capital city, the national border, and the coast ("location controls"). While all distance terms enter with significant coefficients, the estimate on the rule of law index retains its economic and statistical significance. In column (4) we augment the model with a rich set of geographic-ecological and natural-resource controls. Conditioning on geography reduces the coefficient on national institutions (approximately by 35%) though the estimate retains its statistical significance.[10]

[9] In Supplementary Appendix Section 2 we report similar specifications at the pixel level and present regression discontinuity graphs showing that there are no jumps at the border within ethnicities across numerous geographic and location feature (Appendix Figures $4a - 4g$).

[10] Land suitability for agriculture, which reflects climatic (temperature and precipitation) conditions, enters most models with a positive and significant estimate. The malaria stability index enters in all specifications with

The results are similar when we use the control of corruption index (in (5)-(8)).

The coefficient in column (4) of Panel A implies that a one-point-increase in the rule of law index (roughly 2 standard deviations), which corresponds to moving approximately from the institutional quality of Angola to that of Gabon, is associated with a 0.70 log points increase in regional luminosity (approximately half a standard deviation) and an 8% higher likelihood that a pixel is lit.

The correlations in Table 3 echo the findings of cross-country works; although the association between institutional quality and development weakens when one accounts for geography, it remains significant. Yet the fact that a small set of observable geographic features explains a moderate fraction of the cross-sectional correlation between institutional quality and ethnic development suggests that other unobserved or hard-to-measure aspects of geography-ecology and culture may further weaken the association.

4.2 Ethnicity-Fixed-Effects Estimates

Having established that regions across the national border that are part of the historical homeland of the same ethnicity are reasonable counterfactuals (in Table 2) and having explored the cross-sectional patterns (in Table 3), we now examine the correlation between national institutions and luminosity within partitioned groups. Table 4 - Panel A gives the results at the country-ethnic homeland level (equation (1)), while Panel B reports the results at the pixel level (equation (2)). For comparability, in odd-numbered columns we report cross-sectional estimates, while in even-numbered specifications we add ethnicity fixed effects. The cross-sectional estimates in columns (1) and (5) suggest that border areas that belong to countries with higher institutional quality display higher levels of development. For example, the estimate of the linear probability model in column (1) implies that a one-point increase in the rule of law index increases the likelihood that a populated pixel is lit by approximately 11%.

Yet when we solely exploit within-ethnicity variation, the coefficients on rule of law and control of corruption drop sizably and become statistically insignificant. This is a uniform pattern across all permutations. At the country-ethnic homeland level the estimates fall from 0.65 to 0.19 and from 0.79 to 0.26. The coefficient drop in the linear probability models is starker; the estimate drops from 0.11 to 0.025 in the case of rule of law and from 0.14 to 0.037 in the case of the control of corruption index.[11] We also investigated formally whether the

a statistically negative estimate. The coefficient on land area under water is positive and in many specifications significant. Elevation enters with a negative estimate, which is significant in some models. The petroleum dummy enters always with a significantly positive coefficient. The diamond dummy enters with a negative coefficient that is significant in some permutations.

[11] The results are similar when we use the level of luminosity at the country-ethnic homeland level as the dependent variable. In the analogous to columns (1) and (5) cross-sectional models the estimate (standard

estimates on national institutions are statistically different between the cross-sectional and the within-ethnicity models performing a seemingly unrelated regression (SUR) Hausman-like test; the latter suggests that the difference in the coefficients is statistically significant at standard confidence levels. The insignificance of the ethnicity fixed-effects estimates is not driven by a decrease in the precision of the coefficients, since the standard errors remain largely unchanged (in Panel *A*) or even fall (in Panel *B*); it is the drop in the coefficient magnitude that is substantial.

Figures 4*a*-4*b* below illustrate the lack of a systematic within-ethnicity correlation between light density and institutional quality at the national level. The insignificance is not driven by a few (influential) observations; the within R^2 of the unconditional model is tiny, less than 0.02, implying a negligible economic effect.

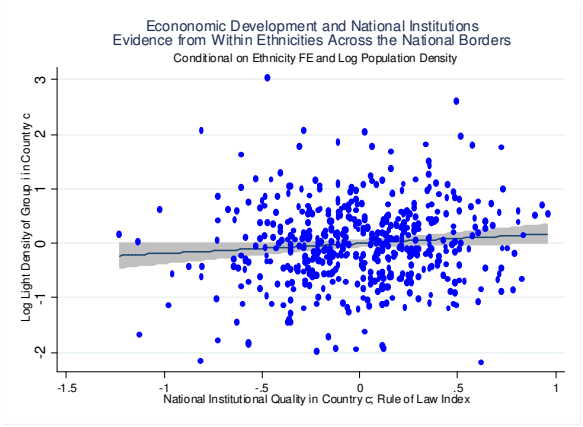

Figure **4a**

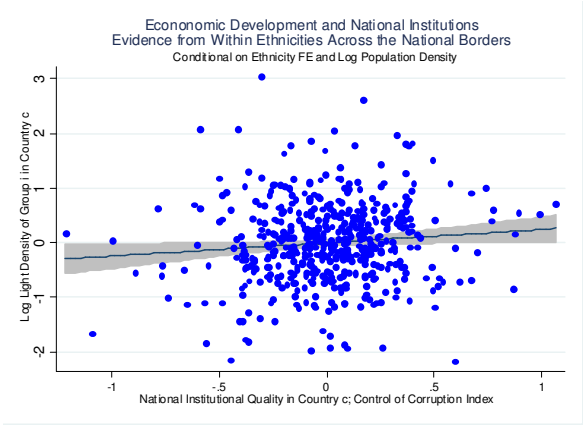

Figure **4b**

In Table 5 we restrict estimation to areas close to the national border focusing in columns (1)-(2) and (5)-(6) on areas within 50 kilometers from each side of the border (total 100 kilometers), while in columns (3), (4), (7), and (8) we focus on regions within 25 kilometers of each side of the border (total 50 kilometers). The idea of this approach (which is similar in spirit to local regressions in regression discontinuity designs) is that by focusing on areas very close to the border, we further neutralize the role of unobservable factors. While in the cross section there is a strong positive association between national institutions and regional development, the within-ethnicity coefficients of rule of law and control of corruption are small and statistically indistinguishable from zero. This applies both at the country-ethnicity sample (in Panel *A*) and at the pixel level (in Panel *B*). The lack of significance is not driven by an increase in the standard errors; in all permutations two-standard-error bands in the within-ethnicity estimates exclude the analogous cross-sectional ones. The cross-sectional estimates fall by more

error) on the national institutions measures is 0.18 (0.09) and 0.215 (0.1065). Once we include ethnicity fixed effects the estimates (standard errors) fall considerably to 0.05 (0.09) and 0.05 (0.13). See Appendix Table 6.

than two thirds once we add ethnicity fixed effects (from 0.68 to 0.083 in (1) and from 0.86 to 0.09 in (4)). Similarly, the linear probability estimates drop from $0.10 - 0.14$ to 0.02, implying tiny economic effects. The statistical and economic insignificance of the coefficient is therefore casting doubt on the causal interpretation of the simple cross-sectional association between national institutional quality and subnational development.[12]

A Note on Ethnicity Fixed Effects Across all model permutations in Tables $4 - 5$, the ethnicity fixed effects are jointly significant at the 99% confidence level. Moreover, most of the ethnicity constants are individually highly significant. This suggests that African subnational development has a strong local geographic-ecological or/and ethnic-specific component. Since our focus is on isolating the effect of national institutions from local geographic and cultural traits, we do not examine in detail which aspects of the local environment the ethnicity fixed effects capture. Yet, given the importance of this issue, in the Supplementary Appendix (Section 3) we conduct a preliminary analysis using anthropological data from George Peter Murdock's Ethnographic Atlas (1967) that records various quantitative measures of cultural, economic, and political traits for a large sample of African ethnicities. The cross-ethnicity specifications show that simply conditioning on the degree of pre-colonial political centralization (a key correlate of both economic and institutional development in Africa; see Gennaioli and Rainer (2007, 2006) and Michalopoulos and Papaioannou (2013c)) or/and the type of the subsistence economy (dependence on agriculture and pastoralism, in particular) weakens considerably the cross-sectional coefficient on the national institutions proxy that becomes statistically indistinguishable from zero. These preliminary findings are in accord with recent works on the deep origins of African development that emphasize the key role of ethnic traits, as opposed to national features (see Michalopoulos and Papaioannou (2013b) for a review of recent studies).

4.3 Regression Discontinuity (RD) Estimates

We now report the results from the regression discontinuity (RD) approach that identifies the (average) effect of national institutions at the border.

Our RD specification within the two major partitions of each ethnic group takes the following form:

$$y_{p,i,c} = a_0 + \gamma IQL_c^{HIGH} + f(BD_{p,i,c}) + \lambda_1 PD_{p,i,c} + \lambda_2 AREA_{p,i,c} + X'_{p,i,c}\Phi + a_i + \zeta_{p,i,c}. \quad (3)$$

[12] The insignificance of the coefficient on national institutions is not driven by the double clustering of the standard errors. In many specifications clustering only at the country level suffices to make the estimate statistically indistinguishable from zero at standard confidence levels.

The difference with specification (2) is that since we aim at identifying the effect of national institutions on regional development at the border, we add RD-polynomials of the distance from the centroid of each pixel to the national border ($f(BD_{p,i,c})$), allowing the coefficients on the polynomial terms to be different for each side of the boundary. IQL_c^{HIGH} is a dummy variable that takes on the value one for pixels falling in the country with the relatively better institutions, as reflected either in the rule of law or in the control of corruption index.[13] 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.

Previous research has employed variants of the above regression equation using different control functions of the running variable (distance to the border in our application) and limiting estimation close to the discontinuity using different bandwidths. Some researchers control for high-order polynomials in the forcing/running variable using 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)). Others estimate local regressions limiting their analysis to an area close to the discontinuity either including or excluding the polynomial control function (e.g., Dell (2010), Angrist and Lavy (2002)). For completeness, we employ all approaches and experiment with different bandwidths.

Baseline Results Table 6 - Panel A reports the RD results that identify the effect of national institutions on development exactly at the border. In columns (1)-(2) we use all pixels across the two major partitions of split-by-the-border ethnic homelands, controlling flexibly for the distance to the discontinuity using a cubic RD polynomial and a fourth-order RD polynomial, respectively. On average the likelihood that a pixel is lit is just one percentage point higher in the country with the relatively higher institutional quality. The coefficient on the dummy that switches to one for ethnic regions falling in the relatively less corrupt country is even smaller (in (7)-(12)). Specifications (3)-(6) and (9)-(12) report local-linear regression results narrowing estimation to pixels within 50 or 25 kilometers from each side of the border (total bandwidth $100km$ and $50km$, respectively), while including the RD polynomial terms. The coefficient on the high institutional quality country is close to zero (in some specifications is negative) and statistically insignificant. In Panel B rather than including a common to all partitioned ethnicities RD polynomial, we include ethnic-specific polynomial terms on distance to the border (i.e., allowing for the effect of the running variable to be different for each

[13] In previous drafts of the paper, we used the value of the rule of law or the control of corruption measures, rather than the high/low institutional quality transformation. The coefficients on the two proxies of national institutions measures in these "difference-in-discontinuities" specifications were small and statistically indistinguishable from zero.

group). The RD estimates on the national institutions measures are small and statistically indistinguishable from zero across all perturbations.

Graphical Illustration Figures 5*a* and 5*b* provide a visual illustration of the RD results. The Figures plot the average likelihood that a pixel is lit for bins of 5 kilometer width, conditioning on log pixel area and log population density in areas close to the national border (the vertical line marks the national border). We report results using a 100 kilometer and a 50 kilometer bandwidth, respectively. The Figures also plot predicted pixel-level luminosity from a regression that includes a third-order RD polynomial on distance to the border fitted separately for pixels in the partitions that fall in the relatively high institutional quality country (where distance takes on positive values) and pixels falling in the relatively low institutional quality country (where distance takes on negative values). In line with the RD estimates in Table 6 there are no discernible differences in luminosity crossing the national boundary towards the more institutionally developed country.

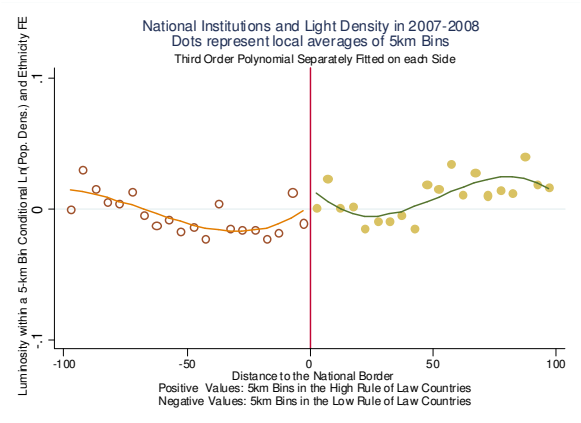

Figure **5a**

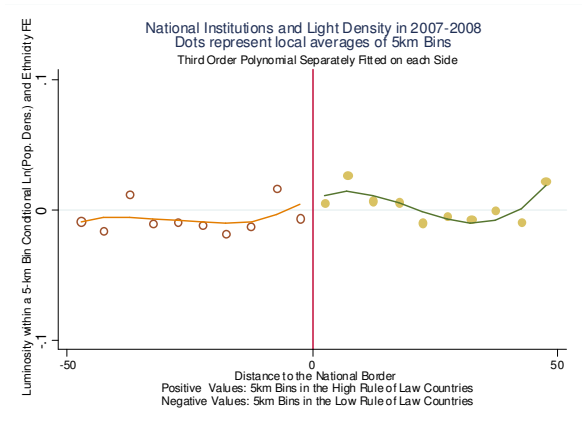

Figure **5b**

Large Discontinuities In Table 7 we report RD specifications (with a third and fourth-order polynomial in distance to the national border) as well as simple ethnicity fixed effects estimates for the two-major partitions of groups split between countries with large, defined as higher than the median or larger than the $75th$ percentile, differences in rule of law or in graft. By focusing on adjacent countries exhibiting substantial differences in national institutions (e.g., Botswana-Zimbabwe, Kenya-Somalia, Ghana-Ivory Coast) we minimize concerns that the insignificance of the estimates is due to measurement error or/and is driven by small differences in national institutions across the border. The estimated coefficients on the two measures of national institutions are quantitatively small (very close to zero) and statistically insignificant across all permutations.

4.4 Sensitivity Checks

We have performed numerous perturbations to investigate the robustness of the lack of systematic association between national institutions and regional development within partitioned African ethnicities. We report and discuss in detail these robustness checks in the on-line Supplementary Appendix (in Section 4). Specifically, we show that the results are similar when: (1) we express luminosity in levels and estimate the specifications with either OLS or Poisson pseudo-maximum likelihood estimator that accounts for the non-linear nature of the luminosity (see Silva and Tenreyro (2006)); (2) we use the alternative institutional quality measures from World Bank's Governance Database; (3) we use the 2007 values of the rule of law or the control of corruption measures; (4) we restrict estimation to ethnic homelands partitioned by straight-line-like borders; (5) we proxy ethnic development with log population density; (6) we account for population clustering near the border, temporary migration, and bleeding of the lights data by excluding areas very close to the border; (7) we drop observations iteratively from the Northern, Southern, Central, Western, or Eastern Africa, respectively; (8) we replace ethnicity fixed effects with border fixed effects; (9) we experiment with a two-dimensional RD polynomial on latitude and longitude (see Dell (2010)); (10) we perform our analysis in a smaller sample of 32 partitioned ethnicities using data from the 2005 Afrobarometer household surveys on public goods provision and education as the dependent variable.

5 Heterogeneity

Our estimates so far summarize the average effect of national institutions on development across all split ethnicities; conducting the analysis for the universe of split groups is useful as it accounts for idiosyncratic features both across specific borders and across ethnicities. Moreover, studying the link between national institutions and ethnic development across all partitioned homelands makes the estimates less prone to "external validity" concerns that characterize case-study evidence. Nevertheless, the richness of the quasi-natural experiment of ethnic partitioning in Africa allows us to estimate the role of national institutions separately for each split group. Doing so will help reveal whether the weak correlation between national institutions and subnational development within groups reflects a generalized phenomenon or conceals considerable heterogeneity. Finding the latter will set stage for exploring the factors that give rise to the observed pattern.

As a first step of moving beyond average effects, we repeated our estimation for groups split *within* each of the main African regions (using Nunn's (2008) classification). The results (reproduced in Appendix Table 16) suggest that the relationship between national institutions

and ethnic development varies across African regions. On average for (28) groups split within countries in Southern Africa those partitions belonging to the more institutionally developed nations are more economically developed. However, the opposite is true when we focus within North Africa (though the number of split groups is quite small). In Eastern and Western Africa differences in national institutions have no bearing on the within-ethnicity economic performance, while in Central Africa there is some evidence of a significantly positive association.

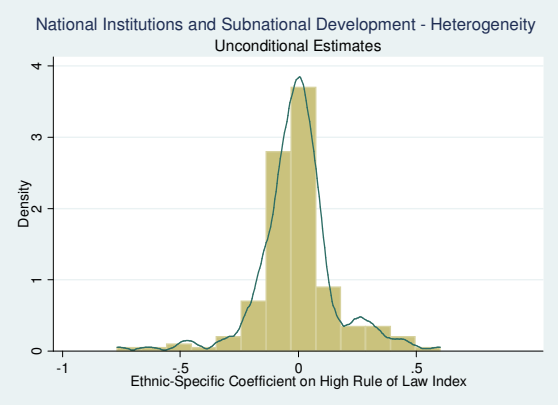

Figure **6a**

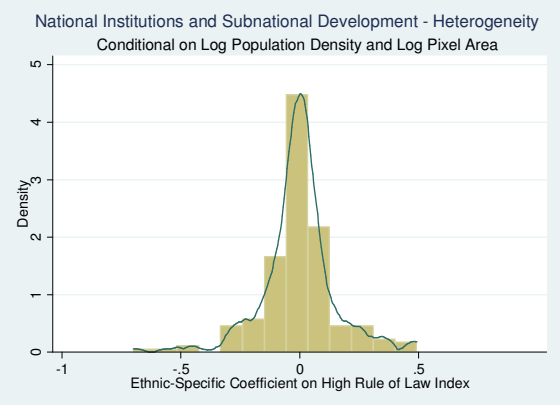

Figure **6b**

Perhaps more importantly, we estimate the correlation between national institutions and pixel-level development separately for each split group. Figure 6*a* reports the distribution of the unconditional estimates on the high rule of law binary index; this is similar to a simple test of means of the likelihood that a pixel is lit on the two sides of the border. Figure 6*b* plots the distribution of the estimate when we condition on log population density and log pixel area.[14] (The results are similar when we include additional controls or/and RD polynomials.)

A couple of interesting patterns emerge. First, both the average and the median value of the ethnic-specific estimates are zero; this applies both to the unconditional and to the conditional coefficients. Moreover, the mass of the distribution is centered on zero. Not only the coefficient is small, but for about 120 groups (62.5%) the absolute value of the t-statistic in the conditional estimates is less than 1.60, implying statistically insignificant patterns.

Second, better national institutions translate into a significantly higher regional development for approximately 38 cases (around 20% of the sample). The Ambo, depicted in Figure 2*b*, is one such example. Other examples where differences in rule of law translate into increased luminosity include the Thonga (split by the border of South Africa and Mozambique),

[14] The histogram plots 190 ethnic-specific coefficients. This is because for 30 ethnicities, all pixels on both sides of the border are unlit (hence the role of national institutions cannot be estimated). Examples of ethnicities with zero luminosity on both sides of the border include the Mundu (split by the Sudan-Zaire border), the Koma and the Suro (split by the Ethiopian-Sudan border), and the Karamojong (split between Kenya and Uganda).

the Seke (split between Gabon and Equatorial Guinea), and the Tukulor (split by the Senegal-Mauritania border).

Third, in approximately $30 - 34$ cases the coefficient for the high institutional quality country is significant, but with the "wrong" sign; so in roughly 17% of the sample of partitioned groups, we obtain the counter-intuitive pattern of within-ethnicity development being higher in the country with the relatively worse national institutions. For example, in the homeland of the Anyi in Ivory Coast (rule of law index -1.25) 60% of populated pixels (132) are lit, while in the more institutionally advanced Ghana (rule of law index -0.23), only 30% of all (89) populated pixels are lit (see Figure 2*c*).[15]

Overall, the uncovered heterogeneity provides a useful reminder that extrapolating from the findings of an analysis that focuses on a single border discontinuity can be quite misleading. At the same time it raises the question about the determinants of the observed variation.

6 Penetration of National Institutions

In this section we link the heterogeneous relationship between national institutions and subnational development with the limited reach of African governments in areas far from the capital. First, we discuss the argument of the African historiography on the limited presence of the colonial and national governments in the hinterland. Second, we estimate the baseline empirical model (equation (2)) linking national institutions to development separately for pixels close to and far from the respective capitals. Third, we explore the interaction between capital distance and national institutions in shaping regional development for *all* African ethnic homelands. Fourth, we test the conjecture regarding the limited reach of the state in remote areas using direct proxies of law enforcement at the regional level.

6.1 The Argument

Europeans' presence in Africa was (with some exceptions) limited to the coastline and the capital cities (see Herbst (2000) for a summary). Hence, colonial institutional arrangements, reflected through persistence on today's national institutions, would have limited reach far from the capital cities. Similarly, several African scholars have argued that due to the lack of infrastructure (roads, transportation systems) and the unfavorable geography (desert areas, rugged terrains, rainforests), colonial and post-colonial institutions have minimal impact on the hinterland (Mamdani (1996)). Herbst (2000), for example, reproduces a quote of a senior official

[15] One possible interpretation for this finding may be related to the observation that law enforcement declines further from the capital city (see Section 6 below). This implies that in order to benefit from well-functioning institutions economic activity may relocate closer to the capital producing, in some instances the counter-intuitive negative relationship between national institutions and economic development at the border.

in Central African Republic saying that "*the State stops at PK 12, twelve kilometers from the capital, Bangui.*" In his influential work Bates (1983) argues that African national leaders did not even try to extend state's power in the countryside, focusing instead on consolidating power in the capital and the few large urban centers.

Moreover, the Scramble for Africa resulted in many countries having peculiar shapes, further inhibiting the presence of national governments in remote areas (Herbst (2000), Englebert (2009)). For example, the Casamance region in Southern Senegal (where the partitioned Diola-Jola reside) is isolated from Dakar, as Gambia effectively cuts Senegal into two parts. Likewise, the rainforest of Central Africa limits the presence of the government of the Democratic Republic of Congo in the Eastern provinces of North and South Kivu, located hundreds of miles away from Kinshasa. The latter are quite often ruled by local militias and rebel groups.

Furthermore, Europeans mostly ruled from the capitals that were located close to the coast, as colonizers had little interest in settling the hinterland that was controlled with the assistance of local chiefs. 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*". He lists many examples where colonizers decided to ignore local needs and established capital cities outside preexisting polities. As extreme examples he lists Mauritania and Bechuanaland (Botswana) that were ruled during colonization by capitals outside their nominal territories (Saint-Louis and Mafeking, respectively).[16]

6.2 National Institutions and Proximity to the Capital. Evidence from Partitioned Ethnicities

Baseline Estimates The key hypothesis that emerges from this narrative is that colonial and contemporary national institutions are less likely to be influential in areas far from the capital centers. Yet in the vicinity of the latter, national institutions are likely to be important. In this section we formally explore this conjecture. In Table 8 in odd-numbered columns we estimate the baseline ethnicity-fixed-effects specification for pixels of split groups that are close to the capitals; in even-numbered columns we examine the within-ethnicity association between national institutions and development across pixels that are far from the capital city.

Panel *A* reports results using all pixels, while Panel *B* gives local linear regression estimates restricting estimation to pixels within 50 kilometers from each side of the border. In columns (1)-(4) we use as a cutoff the median distance to the capital across all pixels (367.4 kilometers). The estimates on the rule of law and the control of corruption measures in columns (1) and (3) are positive (0.07 and 0.12), considerably larger than the baseline estimates in the

[16] 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 them.

full sample of split ethnicities (where the estimate was around 0.03 for both measures; see table 4). The coefficient on the control of corruption index is significant at standard confidence levels, while the estimate on the rule of law index is weakly insignificant (p-value 0.15). In contrast, when we focus on pixels that are far from the respective capital city on both sides of the national border, the coefficients on the rule of law and the control of corruption index in columns (2) and (4) become much smaller (0.02 - 0.03); both estimates are statistically indistinguishable from zero.

In columns (5)-(8) we use the median relative distance to the capital within each country to distinguish between pixels that are relatively close and relatively far from the capital.[17] The estimates in columns (5) and (7) on the two proxies of national institutions are positive (0.105 and 0.146) and statistically significant, suggesting that better functioning national institutions translate into a higher development within partitioned ethnic homelands, but only when we look at pixels that are relatively close to the respective capital cities. In contrast, the small $(0.01 - 0.02)$ and statistically insignificant estimates on national institutions in columns (6) and (8) suggest that within partitioned groups for pixels that are relatively far from the capital on both sides of the border there seems to be no association between institutional quality and subnational development. The coefficient in column (7) implies that for pixels of partitioned ethnic homelands that are relatively close to the respective capitals, a one-standard-deviation increase in the control of corruption index (approximately 0.5) increases the probability that a pixel is lit by approximately 7% (more than half of the average likelihood that a pixel is lit; see Table 1).

A couple of examples help to illustrate these results. Both partitions of the Fon, a group split between Benin and Togo, are close to the corresponding capitals (on average pixels are 100 kilometers away from Porto-Novo and Lome, respectively). Luminosity in the Fon-land in Benin, where the rule of law index equals -0.43, is on average 53.8%, while luminosity is 24% in the less institutionally developed Togo, where the rule of law index equals -0.94. Another example is the Chaga, split by the Tanzanian-Kenyan border; all pixels in both sides of the border are close to Nairobi and Dodoma. 37.5% of populated pixels of the Chaga in Tanzania are lit (rule of law index equals -0.45) while luminosity at the Kenyan side of the border is significantly lower with only 22.8% of pixels being lit (rule of law index in Kenya is -1.02). Another interesting case is the Ronga, a Shona group split between Mozambique (rule

[17] To derive the relative distance within-country we divide pixel-level capital distance by the maximum distance to the capital in each country. This standardization is motivated by the fact that African states are of different land mass (ranging from small ones like Burundi and Benin to very large ones, such as Sudan, Angola, and the Democratic Republic of Congo). Another benefit of using relative (as compared to absolute) distance is that we mitigate concerns that our estimates are driven by large countries which naturally have larger values in absolute capital distance.

of law equals −0.745) and South Africa (rule of law index equals 0.156). Ronga's homeland in Mozambique's has 41.5% of its pixels lit. Incidentally the capital Maputo falls within its boundaries hence all pixels are classified as close. Yet, in the more institutionally advanced South Africa 64% of the pixels in the Ronga partition that are close to Pretoria are lit.

It is instructive to put the magnitudes of the coefficients in perspective and obtain a welfare assessment in terms of the affected population. Across African countries (as of 2000) there were roughly 200 million people that reside in the homelands of partitioned ethnicities. Of those 200 million individuals, 45% live relatively far from the respective capitals with the remaining 55% reside relatively close to the capitals, where national institutions seem to have a bearing on regional economic performance.

RD Visualization Figures $7a - 7b$ provide a graphical illustration of the differential association between national institutions and regional development within partitioned groups for pixels that are relatively close (on both sides of the border) and pixels that are relatively far from the capital city (on both sides of the border), respectively.

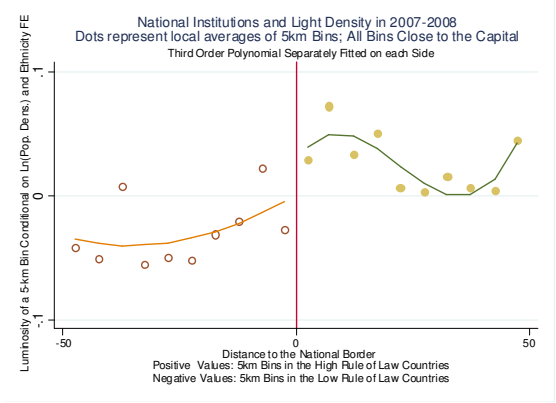

Figure **7a**

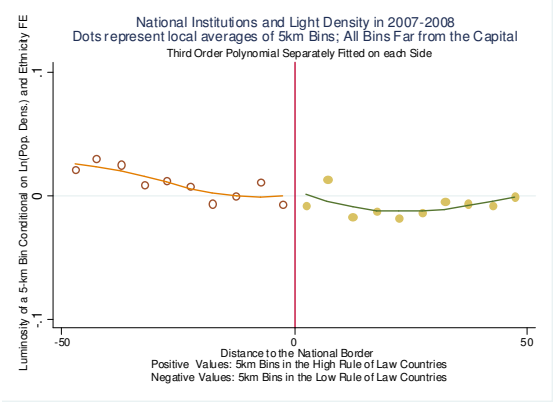

Figure **7b**

The Figures plot the average likelihood that pixels within 5 kilometer bins are lit, conditional on log population density and log land area; the Figures also plot predicted luminosity using cubic RD polynomials on distance to the border. When we focus on pixels of partitioned ethnicities that are relatively close to the respective capitals on both sides of the border (Figure $7a$), pixel-level luminosity at the side of the border with the better-functioning national institutions is significantly higher. In contrast, when we zoom in on pixels that are relatively far from the respective capital cities (Figure $7b$), there are no differences in luminosity, in spite of differences in national institutions.

Robustness In the Supplementary Appendix (in Section 5) we report several sensitivity checks showing that regional development is significantly higher in the part of split

homelands falling in the country with better functioning institutions, but *only* for pixels that are (relatively) close the capital. Specifically, this pattern obtains when: (1) we use information from all split groups by pooling all pixels (40, 209 observations) and interact the national institutions measures (as well as all controls) with a dummy variable indicating pixels that are close (or relatively close) to the respective capital; (2) we drop iteratively each of the five main African regions; (3) we proxy national institutions with the alternative composite measures of World Bank's Governance Database; (4) following Acemoglu and Johnson (2005) we proxy "property rights" institutions with Polity's Project executive constraints measure or we proxy "contractual" institutions with the legal formalism indicators of Djankov, La Porta, Lopez-de-Silanes, and Shleifer (2003); and (5) we use proxies of "early" political institutions (at African independence).

6.3 Evidence from All Groups

In principle, the conjecture of the African historiography that the role of national institutions dissipates as one moves further from the capital regards the universe of ethnic homelands. In Table 9 we augment the baseline model (equation (2)) with the (absolute or relative) distance of each pixel to the capital city and an interaction term between distance to the capital and national institutions.

In columns (1) and (4) we report cross-sectional specifications. Due to omitted variables, reverse causation, and other forms of endogeneity, these specifications do not identify casual relationships; yet these models are useful to quantify the correlation between institutional and economic development *at the capital*; the estimate in column (4) suggests that a one-point-increase in the rule of law index increases the likelihood that a pixel at the capital city is lit by approximately 18%; for comparability the analogous cross-sectional estimate in table 3 was approximately half, 0.095. Moreover, the negative and significant coefficients on both absolute and relative distance to the capital indicate that subnational development declines as one moves further from the capital. Yet the most interesting finding is the significantly negative coefficient on the interaction term between distance to the capital and rule of law; this implies that the correlation between national institutions and regional development declines for regions further from the capital centers.

In columns (2) and (5) we include a vector of country-specific constants as this allows us to account for time-invariant country-level characteristics. The coefficient on distance to the capital and most importantly the coefficient on the interaction term between distance to the capital and the rule of law index remain significant at conventional levels.

Figure **8a** Figure **8b**

Column (3) and (6) report our preferred country-ethnicity fixed effect specifications; estimating these quite restrictive models -that include a vector of 1209 ethnic constants- is feasible as proximity to the capital changes not only across but also *within* ethnic homelands. As in our baseline estimates in the sample of partitioned ethnic groups (in Tables $4-8$), by exploiting solely within ethnic homeland variation, we account for ethnic-specific features, related to either geography or culture; moreover the inclusion of ethnicity fixed effects also account for any relation between pre-colonial ethnic features and the location of the capital by colonizers.[18] The coefficient on the interaction term between national institutions and regional development continues to enter with a negative sign; moreover, the coefficient is statistically significant at standard confidence levels. Thus, the within-ethnicity results in the full sample of African ethnic homelands are quite similar to the pattern obtained within split ethnicities (shown in Table 8). The statistically negative estimate on the interaction in columns (3) and (6) implies that the decaying role of countrywide institutional structures far from the capitals is a general phenomenon and not an artefact of country-specific and more importantly ethnic-specific differences within countries. Figures $8a-b$ provide a graphical illustration of how (relative) distance to the capital mediates the within-ethnicity correlation between rule of law and subnational development in the African continent.[19]

6.4 The Reach of the Law beyond the Capital Cities

The pattern uncovered in Tables $8-9$ points out that the influence of national institutions on subnational development diminishes in areas further from the capital cities. These findings are consistent with the view of the African historiography that both the colonial and the national state is becoming increasingly absent in remote areas with a concomitant decline in law enforcement. In an attempt to provide direct evidence of the differential state presence in areas far and close to the capital, we used individual-level data from the third round of the Afrobarometer Surveys, conducted in 2005 across a representative sample of either $1,200$ or

[18] Most African capitals were imposed by the colonial powers with little worry on the post-independence needs of African states (Herbst (2000)). Nevertheless, it is plausible that colonizers or early state-builders may have decided on the location of the capital taking into account pre-existing differences across ethnic groups, placing, for example, the capital closer to groups that cooperated or were more developed. The inclusion of ethnicity fixed-effects, thus, also accounts for this potentiality.

[19] Our finding that the positive correlation between national institutions and regional development weakens as one moves further from the capital cities is also related to the literature on optimal country size (Alesina and Spolaore (2003)) and recent works studying the implications of the location of the capitals on institutional and economic development (Campante and Do (2013) and Campante, Do, and Guimaraes (2013)).

2,400 respondents in 17 Sub-Saharan countries.[20]

The Afrobarometer Surveys include two questions closely related to law enforcement. The first question regards law enforcement of a serious crime and reads: "*How likely do you think it would be that the authorities could enforce the law if a person like you committed a serious crime?*" The second question regards tax evasion and reads: "*How likely do you think it would be that the authorities could enforce the law if a person like you did not pay a tax on some of the income you earned?*" Both measures range from 1 to 4, with higher scores indicating that law enforcement is more likely.

To examine whether perceived law enforcement decays in remote areas, we regressed these two proxies on (absolute and relative) distance to the capital. In columns (1)-(8) in Table 10 we report the results. In all specifications we exploit within-country variation including a vector of country fixed effects. Odd-numbered columns report unconditional estimates, while in even-numbered columns we include a rich set of individual controls.[21] The coefficient on absolute distance to the capital (in (1), (2), (5), and (6)) is negative and significant, suggesting that the likelihood of law enforcement declines for respondents residing further from the capital city. The same pattern applies when we use relative distance to the capital (in columns (3), (4), (7), and (8)). Individuals located in enumeration areas (villages/towns/cities) relatively further from the respective capitals answer that it is less likely for the law to be enforced. These results clearly indicate a declining presence of authorities in the hinterland.

The specifications reported in columns (9)-(12) provide auxiliary evidence on the declining presence of the state in areas far from the capital using survey data on individual's self-identification. In these linear probability models the dependent variable is a dummy equal to one if the respondent identifies either exclusively or predominantly with the nation rather than his/her ethnic group. Across all specifications both absolute and relative distance enter negatively implying that in areas far from the capital people are more likely to identify with their ethnicity, as compared to the state.[22]

It should be acknowledged that systematic measurement error in the institutional quality proxies 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

[20] These countries are: Benin, Botswana, Ghana, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Namibia, Nigeria, Senegal, South Africa, Tanzania, Uganda, Zambia and Zimbabwe. The question on national versus ethnic identification was not asked in Zimbabwe. The 3rd round of the Afrobarometer Surveys were also conducted in Cape Verde; yet we do not have answers to these questions.

[21] Following Nunn and Wantchekon (2011) in even numbered columns we control for the respondent's age and age squared, a gender indicator, a dummy variable for urban households, 22 religion fixed effects, 25 occupational fixed effects, 5 living conditions fixed effects, and 9 education constants.

[22] In Supplementary Appendix 23 we show that a similar pattern emerges when we use as outcomes an index capturing how easy it is to obtain identity documents or basic household services (like electrification, piped water, and sewage system) from the government.

capital cities capturing the rules governing activities of the formal economy, they might not reflect accurately the institutional features in rural areas (Pande and Udry (2006)). Yet the evidence uncovered in Table 10 clearly indicates that law enforcement and state presence more generally decline with distance to the capital centers.

7 Conclusion

We study the role of national institutions in shaping regional development captured by satellite images on light density at night by employing a novel approach that accounts both for (un)observed geographical-ecological features and ethnic-specific (cultural, historical) factors. We take advantage of the fact that the arbitrarily drawn colonial borders that endured after African independence partitioned several ethnic groups in different countries, subjecting identical cultures residing in geographically homogeneous territories to different country-level institutions.

We present four sets of results. First, our analysis documents that -in contrast to previous cross-sectional works- differences in economic performance within ethnicities partitioned between neighboring countries cannot be explained by differences in national institutions. The lack of within-ethnicity association between formal institutions and subnational development casts doubt on the causal interpretation of the cross-sectional strong correlation, for Africa at least. This finding, nevertheless, is in line with studies of the African historiography that downplays the role of national institutional structures stressing instead the key role of ethnic-specific traits, related to the role of chiefs, culture, and pre-colonial organization (see Michalopoulos and Papaioannou (2013b) for a literature review). Moreover, our results are in accord with a growing body of research that emphasizes the persistent effect of geographical features, cultural and ethnic norms, epidemiological traits as well as history on long-run comparative development (see Spolaore and Wacziarg (2013) and Nunn (2013) for reviews).

Second, we move beyond average effects by examining the association between national institutions and regional development for each partitioned group in Africa. The analysis reveals considerable heterogeneity. Third, in an effort to shed light on the factors shaping heterogeneity and motivated by studies in the African historiography that emphasize the inability of many African governments to broadcast power outside the capitals (Herbst (2000)), we find that unlike split groups in the hinterland, differences in national institutions map into within-ethnicity differences in luminosity for those located close to the respective capitals. This phenomenon is not specific to split groups. In fact, it reflects a generalized pattern. Looking at the universe of African groups reveals that the explanatory power of national institutions monotonically declines with distance to the capital city. Fourth, we use individual-level data from the

Afrobarometer Surveys and show that law enforcement decays with distance to the capital. Moreover, ethnic (as opposed to national) identification becomes stronger for individuals in the hinterland.

These correlations vividly illustrate the limited penetration of the state in remote areas, suggesting -that at least within Africa- treating countries as homogeneous entities where nationwide characteristics, in our case the quality of national institutions, exert a uniform influence across regions may be quite misleading. In fact, the uncovered patterns are supportive to an old idea among development scholars (e.g., Lewis (1954)) on the coexistence in Africa (and other parts of the developing world) of a "dual" economic-institutional framework with customary rules being dominant in the countryside and colonial-national institutions becoming relevant for regions closer to the capitals (see also Herbst (2000)).

Our paper calls for future research. One could employ our methodology that compares well-being in border areas using high resolution proxies of development (such as satellite light density) to shed light on the debate regarding the determinants (correlates) of comparative economic development; examining, for example, the effect of financial development, public policies, foreign aid, fragmentation, etc. Moreover, subsequent works could build on our analysis regarding the declining role of national institutions in areas far from the capital and explore the impact of other geographic, cultural, national or ethnic-specific features that may attenuate or strengthen the importance of national policies in promoting subnational development. Finally, our evidence suggests that research on the role of institutions on development should move beyond average effects and examine the delicate interplay between formal institutions, state capacity, and ethnic traits.

8 Data Appendix

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 at the ethnic homeland we use Log (0.01 + Average Luminosity). In the Supplementary Appendix we also use the level of luminosity as the dependent variable. In the regressions at the pixel level we use as the dependent variable a dummy variable that takes on the value of one if the area is lit and zero otherwise.
Source: Available at http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html.

Population Density: Log (population density per sq. km. in 2000). *Source: Nelson, Andy,* 2004. *African Population Database Documentation, UNEP GRID Sioux Falls. Available at:* http://na.unep.net/siouxfalls/datasets/datalist.php

Area: Log (land area) at the country-ethnic homeland level or at the pixel level. Please note that while most pixels are of the same surface area ($2.5x2.5$ degrees), pixels intersected by the coastline, borders, and lakes are smaller.

Water Area: In the regressions at the country-ethnicity level we use the Log (1 + total area covered by rivers or lakes in sq. km.). In the regressions at the pixel level we use a dummy variable that takes on the value of one when some water body (river, lake, or other stream) falls in each pixel and zero otherwise. *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, CO, USA. Global Ministry Mapping System.*

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

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

Malaria Stability Index: Average prevalence of malaria within each ethnicity-country or within each pixel. 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. *Source: Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004)*

Distance to the Capital City: The geodesic distance from the centroid of each ethnicity-country/pixel to the capital city of the country it belongs to, measured in 1000*s*

of km's. *Source: Calculated using the Haversine formula.*

Relative Distance to the Capital City: Distance to the Capital City divided by the maximum distance to the capital in the same country.

Distance to the Sea Coast: The geodesic distance to the nearest coastline from the centroid of each ethnicity-country/pixel, 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 National Border: The geodesic distance to the nearest national border from the centroid of each ethnicity-country/pixel, measured in 1000*s* of km's. *Source: Calculated using ArcGis.*

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

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

Rule of Law: The index is "capturing perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence." The standardized index ranges from -2.5 to $+2.5$ with higher values indicating better functioning institutions. We use the average value over the period $1996-2006$. In the regression discontinuity (RD) analysis we also define a dummy (indicator) variable that takes on the value one pixels falling in the country with the higher level of the rule of law index. *Source: World Bank Governance Matters Indicators Database (Kaufmann, Kraay, and Mastruzzi (2008)). 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. We use the average value over the period $1996-2006$. In the regression discontinuity (RD) analysis we also define a dummy (indicator) variable that takes on the value one for pixels falling in the country with the higher level of the control of corruption index. *Source: World*

Bank Governance Matters Indicators Database (Kaufmann, Kraay, and Mastruzzi (2008)). available at:
http://info.worldbank.org/governance/wgi/index.asp.

Law Enforcement 1: Ordered (range $1-4$) variable capturing law enforcement regarding committing a serious crime based on individuals' responses to the following question: "*How likely do you think it would be that the authorities could enforce the law if a person like you committed a serious crime?*" A score of 1 indicates "not at all likely"; a score of 2 suggests "not very likely"; a score of 3 indicates "likely"; and a score of 4 indicates "very likely". *Source: Afrobarometer Surveys. 3rd Round. Question Number: Q70B. available at: http://www.afrobarometer.org/data/round-3-merged*

Law Enforcement 2: Ordered (range $1-4$) variable capturing law enforcement regarding tax evasion based on individuals' responses to the following question: *"How likely do you think it would be that the authorities could enforce the law if a person like you did not pay a tax on some of the income you earned?*" A score of 1 indicates "not at all likely"; a score of 2 suggests "not very likely"; a score of 3 indicates "likely"; and a score of 4 indicates "very likely". *Source: Afrobarometer Surveys. 3rd Round. Question Number: Q70D. available at: http://www.afrobarometer.org/data/round-3-merged*

National Identification: Dummy variable that takes on the value of one if the respondent replies that he or she identifies more strongly with the nation rather than his/her ethnicity. The question reads: "*Let us suppose that you had to choose between being a [Ghanaian/Kenyan/etc.] and being a _ _ _ _ _ _ _ _ _ [respondent's identity group]. Which of these two groups do you feel most strongly attached to?*" The variable equals one if the respondent replies that he identifies solely or predominantly with his nation and zero otherwise. *Source: Afrobarometer Surveys. 3rd Round. Question Q82. available at:*
http://www.afrobarometer.org/data/round-3-merged

References

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

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

——— (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 107(4), 1231–1294.

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

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003): "Fractionalization," *Journal of Economic Growth*, 8(2), 155–194.

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

Alesina, A., and P. Giuliano (2013): "Family Ties," in *Handbook of Economic Growth, Volume 2*, ed. by P. Aghion, and S. Durlauf, vol. 2. Elsevier, Amsterdam, Netherlands.

Alesina, A., S. Michalopoulos, and E. Papaioannou (2013): "Ethnic Inequality," Working Paper, NBER No. 18512.

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

Algan, Y., and P. Cahuc (2013): "Trust, Institutions and Economic Development," in *Handbook of Economic Growth, Volume 2*, ed. by P. Aghion, and S. Durlauf, vol. 2. Elsevier, Amsterdam, Netherlands.

Alsan, M. M. (2012): "The Effect of the TseTse Fly on African Developmen," manuscript, Harvard University.

Angrist, J., and V. C. Lavy (2002): "New Evidence on Classroom Computers and Pupil Learning," *Economic Journal*, 112(482), 735–765.

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

Ashraf, Q., and O. Galor (2013): "The Out of Africa Hypothesis, Human Genetic Diversity, and Comparative Economic Development," *American Economic Review*, 103(1), 1–46.

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. (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.

——— (2008): *When Things Fell Apart. State Failure in Late-Century Africa*. Cambridge University Press, New York, NY.

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

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

Campante, F., and Q.-A. Do (2013): "Isolated Capital Cities, Accountability, and Corruption: Evidence from US States," mimeo, Harvard Kennedy School.

Campante, F., Q.-A. Do, and B. Guimaraes (2013): "Isolated Capital Cities and Misgovernance: Theory and Evidence," mimeo, Harvard Kennedy School.

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

Cogneau, D., S. Mesplé-Somps, and G. Spielvogel (2012): "Development at the Border: A Study of National Integration in Post-colonial West Africa," mimeo, Paris Scool of Economics.

Cogneau, D., and A. Moradi (2011): "Borders that Divide: Education and Religion in Ghana and Togo since Colonial Times," Department of Economics, University of Sussex Working Paper 2911.

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

Davidson, B. (1992): *The Black Man's Burden. Africa and the Curse of the Nation-State*. James Currey, Oxford, United Kingdom.

De Long, B., and A. Shleifer (1993): "Princes or Merchants? City Growth before the Industrial Revolution," *Journal of Law and Economics*, 36(2), 671–702.

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

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

Djankov, S., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer (2003): "Courts," *Quarterly Journal of Economics*, 118(2), 453–517.

Doll, C. N., J.-P. Muller, and J. G. Morley (2006): "Mapping Regional Economic Activity from Night-Time Light Satellite Imagery," *Ecological Economics*, 57(1), 75–92.

Dowden, R. (2008): *Africa: Altered States, Ordinary Miracles*. Portobello Books Ltd, London, UK.

Easterly, W., and R. Levine (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 112(4), 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. (2009): *Africa, Unity, Sovereignty and Sorrow*. Lynne Rienner Publishers, Inc., Boulder, Colorado.

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

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

Gennaioli, N., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer (2013): "Human Capital and Regional Development," *Quarterly Journal of Economics*, 128(1), 105–164.

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. La 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 (2013): "Collective Action in Diverse Sierra Leone Communities," *Economic Journal*, 123(568), 285–316.

Grosjean, P. (2011): "Long Term Institutional Persistence: Ottoman Rule and Financial Development in the Regions of Europe," *Journal of Comparative Economics*, 39(2), 1–16.

Guiso, L., P. Sapienza, and L. Zingales (2006): "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 20(2), 23–48.

Henderson, V. J., A. Storeygard, and D. N. Weil (2012): "Measuring Economic Growth from Outer Space," *American Economic Review*, 102(2), 994–1028.

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

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

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.

Jedwab, R., and A. Moradi (2012): "Colonial Investments and Long-Term Development in Africa: Evidence from Ghanaian Railways," mimeo George Washington University and Sussex University.

Jerven, M. (2013): *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*. Cornell University Press, Ithaca, NY.

Kaufmann, D., A. Kraay, and M. Mastruzzi (2008): "Governance Matters VIII: Aggregate and Individual Governance Indicators, 1996-2007," *Policy Research Working Paper Series 4654, 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., F. Lopez-de-Silanes, and A. Shleifer (2008): "The Economic Consequences of Legal Origins," *Journal of Economic Literature*, 46(2), 285–332.

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

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

Laeven, L., and C. Woodruff (2007): "The Quality of the Legal System, Firm Ownership, and Firm Size," *Review of Economics and Statistics*, 89(4), 601–614.

Landes, D. S. (1998): *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. W.W. Norton, New York, NY.

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(3), 807–859.

Lewis, A. W. (1954): "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 22(1), 139–191.

Mamdani, M. (1996): *Citizen and Subject*. James Currey, London, UK.

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

Michalopoulos, S., A. Naghavi, and G. Prarolo (2012): "Trade and Geography in the Origins and Spread of Islam," NBER Working Paper No. 18438.

Michalopoulos, S., and E. Papaioannou (2011): "Divide and Rule or the Rule of the Divided? Evidence from Africa," NBER Working Paper No. 17184.

——— (2013a): "The Long-Run Effects of the Scramble for Africa," NBER Working Paper No. 17620.

——— (2013b): "On the Ethnic Origins of African Development: Traditional Chiefs and Precolonial Political Centralization," mimeo Brown University and London Business School.

——— (2013c): "Pre-colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 81(1), 113–152.

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

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

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

Moulton, B. R. (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables in Micro Units," *Review of Economics and Statistics*, 72(2), 334–338.

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

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

Nunn, N. (2012): "Culture and the Historical Process„" *Economic History of Developing Regions*, 27(S1), 108–126.

——— (2013): "Historical Development," in *Handbook of Economic Growth, Volume 2*, ed. by P. Aghion, and S. Durlauf, vol. 2. Elsevier, Amsterdam, Netherlands.

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

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

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. (2013): "Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night," mimeo MIT.

Putterman, L., and D. N. Weil (2010): "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality," *Quarterly Journal of Economics*, 125(4), 1627–1682.

Ross, M. L. (2012): *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press, Princeton, NJ.

Silva, J. M. S., and S. Tenreyro (2006): "The Log of Gravity," *Review o Economics and Statistics*, 88(4), 641–658.

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

——— (2013): "How Deep Are the Roots of Economic Development?," *Journal of Economic Literature*, 51(2).

Tabellini, G. (2010): "Culture and Institutions: Economic Development in the Regions of Europe," *Journal of the European Economic Association*, 8(4), 677–716.

Weil, D. N. (2008): *Economic Growth*. Prentice Hall.

Wesseling, H. L. (1996): *Divide and Rule: The Partition of Africa, 1880-1914*. Praeger, Amsterdam, Netherlands.

Table 1: Summary Statistics

variable	Obs.	mean	st. dev.	p25	median	p75	min	max
			Panel A: Partitioned Ethnic Homelands Sample					
Light Density	507	0.228	0.768	0.000	0.015	0.095	0.000	8.561
Ln (0.01 + Light Density)	507	-3.168	1.586	-4.605	-3.696	-2.258	-4.605	2.148
Ln (0.01 + Population Density)	507	2.650	1.929	1.647	2.936	3.945	-5.069	6.386
Ln (1 + Water Area)	507	0.242	0.407	0.000	0.075	0.264	0.000	2.464
Ln (Area)	507	1.876	1.490	0.798	1.906	2.963	-2.170	5.665
Mean Elevation	507	0.605	0.434	0.281	0.469	0.938	0.000	2.181
Land Suitability For Agriculture	507	0.415	0.224	0.273	0.439	0.573	0.002	0.959
Malaria Stability Index	507	0.747	0.323	0.548	0.905	1.000	0.000	1.000
Oil Deposit Indicator	507	0.063	0.243	0.000	0.000	0.000	0.000	1.000
Diamond Mine Indicator	507	0.101	0.301	0.000	0.000	0.000	0.000	1.000
Distance to the Capital City	507	0.524	0.371	0.280	0.429	0.685	0.013	1.882
Distance to the Sea Coast	507	0.591	0.425	0.225	0.552	0.931	0.000	1.739
Distance to the Border	507	0.039	0.039	0.012	0.029	0.052	0.000	0.250
Rule of Law	507	-0.914	0.573	-1.355	-1.021	-0.543	-2.198	0.615
Control of Corruption	507	-0.797	0.501	-1.091	-0.864	-0.651	-1.664	0.814
			Panel B: Pixel-level. Partitioned Ethnic Homelands Sample					
Light Density	42710	0.312	2.335	0.000	0.000	0.000	0.000	62.857
Ln (0.01 + Light Density)	42710	-4.111	1.465	-4.605	-4.605	-4.605	-4.605	4.141
Light Density 0-1	42710	0.124	0.329	0.000	0.000	0.000	0.000	1.000
Ln (0.01 + Population Density)	42710	2.074	1.771	0.734	2.143	3.328	-5.094	9.121
Water Area Indicator	42710	0.129	0.335	0.000	0.000	0.000	0.000	1.000
Ln (Area)	42710	4.885	0.650	4.857	5.211	5.243	2.303	5.259
Mean Elevation	42566	641.612	470.408	292.111	526.048	969.556	-406.000	4623.750
Land Suitability For Agriculture	41149	0.341	0.259	0.103	0.325	0.511	0.001	0.999
Malaria Stability Index	42710	0.710	0.395	0.365	0.989	1.000	0.000	1.000
Oil Deposit Indicator	42710	0.018	0.135	0.000	0.000	0.000	0.000	1.000
Diamond Mine Indicator	42710	0.004	0.062	0.000	0.000	0.000	0.000	1.000
Distance to the Capital City	42710	0.532	0.369	0.271	0.439	0.707	0.001	1.933
Distance to the Sea Coast	42710	581.363	414.433	233.430	507.473	896.488	0.286	1735.680
Distance to the Border	42710	65.938	68.091	15.754	44.255	93.799	0.000	448.539
Rule of Law	42710	-0.909	0.606	-1.355	-0.820	-0.506	-2.198	0.615
Control of Corruption	42710	-0.779	0.534	-1.141	-0.864	-0.651	-1.664	0.814

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

Table 2: Validity of Identification Design at the Country-Ethnicity Homeland Level

	Dependent Variable is:									
	Ln (Land Area)	Ln (Area under Water)	Mean Elevation	Land Suitability	Malaria Stability	Oil Indicator	Diamond Indicator	Distance to the Capital	Distance to the Sea Coast	Distance to the Border
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Rule of Law										
Rule of Law	-0.0474	0.0480	0.0041	0.0091	-0.015	0.0365*	-0.0161	-0.1887	0.0001	-0.0037
Double-clustered s.e.	(0.2058)	(0.0453)	(0.0217)	(0.0184)	(0.0230)	(0.0199)	(0.0433)	(0.1233)	(0.0178)	(0.0049)
Adjusted R-squared	0.682	0.796	0.964	0.935	0.941	0.753	0.547	0.582	0.986	0.639
Observations	507	507	507	507	507	507	507	507	507	507
Panel B: Control of Corruption										
Control of Corruption	-0.0686	0.037	0.0119	-0.0052	-0.0324	0.0431	-0.0199	-0.1753	0.0037	-0.007
Double-clustered s.e.	(0.2229)	(0.0610)	(0.0222)	(0.0243)	(0.0233)	(0.0294)	(0.0612)	(0.1490)	(0.0176)	(0.0077)
Adjusted R-squared	0.683	0.795	0.964	0.935	0.941	0.753	0.547	0.565	0.986	0.551
Observations	507	507	507	507	507	507	507	507	507	507
Panel C: High Rule of Law - Two Major Partitioned Ethnic Areas										
High Rule of Law	0.0696	0.0475	-0.0013	0.0129	-0.0103	0.0409	-0.0273	-0.1584	0.0036	0.0031
Double-clustered s.e.	(0.1111)	(0.0346)	(0.0171)	(0.0132)	(0.0195)	(0.0271)	(0.0414)	(0.1029)	(0.0111)	(0.0049)
Adjusted R-squared	0.818	0.866	0.972	0.952	0.953	0.804	0.629	0.612	0.991	0.639
Observations	440	440	440	440	440	440	440	440	440	
Panel C: High Control for Corruption - Two Major Partitioned Ethnic Areas										
High Control of Corruption	0.0262	0.0547	-0.0074	0.0105	-0.0025	0.0318	-0.0091	-0.0304	-0.0032	-0.0002
Double-clustered s.e.	(0.1259)	(0.0351)	(0.0185)	(0.0160)	(0.0217)	(0.0262)	(0.0455)	(0.1182)	(0.0103)	(0.0058)
Adjusted R-squared	0.817	0.867	0.972	0.952	0.953	0.802	0.627	0.569	0.991	0.638
Observations	440	440	440	440	440	440	440	440	440	440
Ethnicities	220	220	220	220	220	220	220	220	220	220
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports within-ethnicity OLS estimates associating various geographical, ecological, and other characteristics with contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index (in Panel A) and control of corruption index (in Panel B) in areas of partitioned ethnicities. In Panels C and D we focus on the two major partitions of an ethnic group (in total 440 ethnic areas of 220 partitioned ethnicities). The high institutional quality indicator takes on the value one for the areas falling in the country with the higher value in the rule of law index (in Panel C) or in the control for corruption index (in Panel D).
The dependent variable in column (1) is the log (land area); in column (2) is the log (1 + land area under water (lakes, rivers, and other streams)); in column (3) is mean elevation; in column (4) is an index of land suitability for agriculture; in column (5) is the average of a malaria stability index; in column (6) is an indicator for country-ethnic areas with an oil field; in column (7) is an indicator for country-ethnic areas with a diamond mine. In columns (8)-(10) the dependent variable is the distance of the centroid of each country-ethnic area from the capital city, the sea coast, and the national border, respectively. 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 3 - Cross-Sectional Analysis
National Institutions and Regional Development

	Rule of Law				Control of Corruption			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Country-Ethnic Homeland Level								
Institutional Quality	0.8229***	0.8017***	0.7646***	0.6513***	1.0328***	1.0351***	0.9945***	0.8218***
Double-clustered s.e.	(0.2839)	(0.2451)	(0.2507)	(0.1962)	(0.3421)	(0.2517)	(0.2334)	(0.2158)
Adjusted R-squared	0.080	0.314	0.379	0.468	0.0950	0.3340	0.3950	0.475
Observations	1209	1209	1209	1209	1209	1209	1209	1209
Panel B: Pixel Level								
Institutional Quality	0.1248***	0.1199***	0.0988***	0.0765**	0.1470***	0.1467***	0.1223***	0.0968**
Double-clustered s.e.	(0.0443)	(0.0415)	(0.0380)	(0.0334)	(0.0478)	(0.0431)	(0.0404)	(0.0409)
Adjusted R-squared	0.049	0.175	0.192	0.256	0.056	0.185	0.198	0.260
Observations	120501	120501	120501	116412	120501	120501	120501	116412
Pop. Dens. & Area	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Location Controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes

The table reports 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 of corruption index (in columns (5)-(8)) at the country-ethnicity level (in Panel A) and at the pixel level (in Panel B). In Panel A the dependent variable is log (0.01 + light density at night from satellite) at the ethnicity-country level. In Panel B the dependent variable is an indicator that takes on the value of one if the pixel (of 0.125 x 0.125 decimal degrees) is lit and zero otherwise.
In the specifications in columns (2)-(4) and (6)-(8) we control for log (population density in 2000) and log land area at the country-ethnicity level (in Panel A) or at the pixel-level (in Panel B). In columns (3), (4), (7), and (8) we control for location characteristics augmenting the specification with distance of the centroid of each ethnicity-country area or each pixel 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) and (8) includes log (1 + area under water (lakes, rivers, and other streams)) for the models in Panel A and an indicator that takes on the value one if there is some water body in the pixel for the models in Panel B, 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 at the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

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

	Rule of Law				Control of Corruption			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Country-Ethnic Homeland Level								
Institutional Quality	0.6510***	0.1943	0.5150**	0.2159	0.7904***	0.2566	0.6019***	0.2675
Double-clustered s.e.	(0.1951)	(0.1898)	(0.2024)	(0.2135)	(0.2268)	(0.2197)	(0.2329)	(0.2439)
Adjusted R-squared	0.292	0.792	0.392	0.798	0.298	0.792	0.393	0.798
Within R-squared	__	0.061	__	0.067	__	0.062	__	0.067
Observations	507	507	507	507	507	507	507	507
Panel B: Pixel Level								
Institutional Quality	0.1072***	0.0246	0.0834**	0.0278	0.1371***	0.0370	0.1097***	0.0403
Double-clustered s.e.	(0.0400)	(0.0165)	(0.0324)	(0.0181)	(0.0464)	(0.0273)	(0.0415)	(0.0290)
Adjusted R-squared	0.149	0.331	0.202	0.340	0.160	0.331	0.209	0.340
Within R-squared	__	0.059	__	0.066	__	0.059	__	0.066
Observations	42710	42710	41025	41025	42710	42710	41025	41025
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Pop. Dens. & Area	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

The table 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 of corruption index (in columns (5)-(8)) in areas of partitioned ethnicities at the country-ethnicity level (in Panel A) and at the pixel level (in Panel B). In Panel A the dependent variable is log (0.01 + light density at night from satellite) at the ethnicity-country level. In Panel B the dependent variable is an indicator that takes on the value of one if the pixel (of 0.125 x 0.125 decimal degrees) is lit 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 (population density in 2000) and log land area at the country-ethnicity level (in Panel A) or at the pixel level (in Panel B). In columns (3), (4), (7), and (8) we control for location and geography, augmenting the specification with distance of the centroid of each ethnicity-country (in Panel A) or each pixel (in Panel B) from the capital city of each country, distance from the closest sea coast, and distance from the national border. The set of geographic controls includes log (1 + area under water (lakes, rivers, and other streams)) for the models in Panel A and an indicator that takes on the value one if there is some water body in the pixel for the models in Panel B, land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. Besides the adjusted R-squared, the table also reports the within R-squared (defined as the difference of the overall R-squared minus the overall R-squared of a model simply with ethnicity fixed effects).

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 5: National Institutions and Regional Development across and within Partitioned Ethnic Groups in Areas Close to the Border. Local Regressions

	Rule of Law				Control of Corruption			
Bandwidth	100 km		50 km		100 km		50 km	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Country-Ethnic Homeland Level							
Institutional Quality	0.6788***	0.0827	0.5975***	0.1503	0.8585***	0.0866	0.7458***	0.1765
Double-clustered s.e.	(0.2116)	(0.2318)	(0.2098)	(0.2480)	(0.2465)	(0.2602)	(0.2591)	(0.2951)
Adjusted R-squared	0.311	0.842	0.285	0.848	0.324	0.842	0.293	0.848
Within R-squared	__	0.027	__	0.033	__	0.027	__	0.033
Observations	432	432	432	432	432	432	432	432
	Panel B: Pixel Level							
Institutional Quality	0.0994***	0.0168	0.1031**	0.0159	0.1292***	0.0269	0.1346***	0.0197
Double-clustered s.e.	(0.0384)	(0.0181)	(0.0407)	(0.0163)	(0.0471)	(0.0259)	(0.0492)	(0.0210)
Adjusted R-squared	0.134	0.319	0.137	0.345	0.144	0.319	0.149	0.345
Within R-squared	__	0.051	__	0.046	__	0.051	__	0.046
Observations	21289	21289	13408	13408	21289	21289	13408	13408
Ethnicity Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Pop. Dens. & Area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table 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 of corruption index (in columns (5)-(8)) in areas of partitioned ethnicities at the country-ethnicity level (in Panel A) and at the pixel level (in Panel B). In Panel A the dependent variable is log (0.01 + light density at night from satellite) at the ethnicity-country level. In Panel B the dependent variable is an indicator that takes on the value of one if the pixel (of 0.125 x 0.125 decimal degrees) is lit and zero otherwise.
In columns (1), (2), (5), and (6) we focus on 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). 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 (population density in 2000) and log land area at the country-ethnic homeland level (in Panel A) and in the pixel-level (in Panel B). Besides the adjusted R-squared, the table also reports the within R-squared (defined as the difference of the overall R-squared minus the overall R-squared of a model simply with ethnicity fixed effects).
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: National Institutions and Regional Development at the Border
Regression Discontinuity (RD) Estimates

	Rule of Law						Control of Corruption					
Bandwidth	All Pixels		100 km		50 km		All Observations		100 km		50 km	
RD Polynomial	3rd-order	4th-order	3rd-order	4th-order	3rd-order	4th-order	3rd-order	4th-order	3rd-order	4th-order	3rd-order	4th-order
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Global Regression Discontinuity (RD) Polynomial												
Institutional Quality	0.0166	0.0058	0.0123	0.0050	-0.0010	0.0122	0.0038	-0.0086	0.0116	0.0123	-0.0030	0.0079
Double-clustered s.e.	(0.0138)	(0.0135)	(0.0165)	(0.0185)	(0.0240)	(0.0272)	(0.0154)	(0.0137)	(0.0166)	(0.0176)	(0.0223)	(0.0276)
Adjusted R-squared	0.342	0.342	0.320	0.320	0.347	0.347	0.342	0.343	0.320	0.320	0.347	0.347
within R-squared	0.059	0.059	0.037	0.037	0.064	0.064	0.059	0.060	0.037	0.037	0.064	0.064
Panel B: Ethnic-Specific Regression Discontinuity (RD) Polynomial												
Institutional Quality	0.0175	0.0203	0.0037	0.0135	0.0068	0.0185	0.0020	0.0114	0.0011	0.0094	0.0124	0.0128
Double-clustered s.e.	(0.0125)	(0.0154)	(0.0124)	(0.0183)	(0.0176)	(0.0318)	(0.0119)	(0.0138)	(0.0136)	(0.0181)	(0.0189)	(0.0283)
Adjusted R-squared	0.416	0.427	0.412	0.423	0.441	0.459	0.416	0.427	0.412	0.423	0.441	0.459
within R-squared	0.133	0.144	0.129	0.140	0.158	0.176	0.133	0.144	0.129	0.140	0.158	0.176
Observations	40209	40209	21289	21289	13408	13408	40209	40209	21289	21289	13408	13408
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pixel Area & Pop. Dens.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports regression discontinuity (RD) estimates. Estimation is performed across the two major partitions of split ethnic group. In Panel A we include a global (common to all partitioned ethnicities) RD polynomial in distance of the centroid of each pixel to the national border, allowing the polynomial terms to differ on the two sides of the border. In Panel B we include ethnic-specific RD polynomials in distance of the centroid of each pixel to the national border, allowing the polynomial terms to differ on the two sides of the border. In odd-numbered columns we include a third-order RD polynomial, while in even-numbered columns we include a fourth-order RD polynomial. In columns (3)-(6) and (9)-(12) we restrict estimation to pixels close to the national border using two different bandwidths. In columns (3), (4), (9), and (10) we focus on pixels within 50 kilometers on each side of the national border (total 100 kilometers). In columns (5), (6), (11), and (12) we focus on pixels within 25 kilometers of each side of the national border (total 50 kilometers). The high institutional quality indicator takes on the value of one for pixels falling in the country with the higher value in the rule of law index (in columns (1)-(6)) or in the control of corruption index (in columns (7)-(12)). All specifications include ethnicity fixed effects (constants not reported), the log of pixel-level population density and the log of pixel surface area. Besides the adjusted R-squared, the table also reports the within R-squared (defined as the difference of the overall R-squared minus the overall R-squared of a model simply with ethnicity fixed effects).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 at the ethnolinguistic level dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 7: National Institutions and Regional Development at the Border
Regression Discontinuity (RD) Estimates. Large Discontinuities

	Rule of Law						Control of Corruption					
Cutt-off	median			75th percentile			median			75th percentile		
Global RD Polynomial	No	3rd-order	4th-order	No	3rd-order	4th-order	No	3rd-order	4th-order	No	3rd-order	4th-order
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Institutional Quality	0.0159	0.0116	-0.0027	0.0134	0.0049	-0.0207	0.009	0.0133	-0.003	0.041	0.0367	0.0024
	(0.0186)	(0.0199)	(0.0212)	(0.0311)	(0.0376)	(0.0330)	(0.0204)	(0.0215)	(0.0201)	(0.0346)	(0.0336)	(0.0251)
Adjusted R-squared	0.398	0.400	0.400	0.458	0.461	0.462	0.399	0.399	0.400	0.457	0.460	0.461
within R-squared	0.072	0.074	0.075	0.075	0.078	0.080	0.064	0.065	0.065	0.078	0.080	0.082
Observations	19349	19349	19349	10031	10031	10031	19814	19814	19814	10353	10353	10353
Ethnicities	110	110	110	51	51	51	108	108	108	46	46	46
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pixel Area & Pop. Dens.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports simple fixed effects as well as regression discontinuity (RD) estimates focusing on large discontinuities in national institutions across the border. Estimation is performed across the two major partitions of each ethnic group. In columns (2), (5), (8), and (11) we include a third-order RD polynomial on distance to the national border. In columns (3), (6), (9), and (12) we include a fourth-order RD polynomial on distance to the national border. In columns (1)-(3) and columns (7)-(9) we use the median value of differences in the rule of law index across adjacent countries to identify large discontinuities. In columns (4)-(6) and columns (10)-(12) we use the 75th percentile of differences in the rule of law index across adjacent countries to identify large discontinuities. The high institutional quality indicator takes on the value of one for pixels falling in the country with the higher value in the rule of law index (in columns (1)-(6)) or in the control of corruption index (in columns (7)-(12)). All specifications include ethnicity fixed effects (constants not reported), the log of pixel-level population density and the log of pixel surface area. Besides the adjusted R-squared, the table also reports the within R-squared (defined as the difference of the overall R-squared minus the overall R-squared of a model simply with ethnicity fixed effects).
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 at the ethnolinguistic level dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 8 - Heterogeneity: National Institutions and Regional Development within Partitioned Ethnicities Close and Far from the Capital Cities

	Rule of Law		Control of Corruption		Rule of Law		Control of Corruption	
	Absolute Distance to the Capital				Relative Distance to the Capital			
	Close	Far	Close	Far	Close	Far	Close	Far
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Observations/Pixels								
Institutional Quality	0.0675	0.0206	0.1220**	0.0330	0.1064*	0.0127	0.1455**	0.0218
Double-clustered s.e.	(0.0438)	(0.0233)	(0.0581)	(0.0258)	(0.0585)	(0.0225)	(0.0675)	(0.0237)
Adjusted R-squared	0.421	0.210	0.423	0.211	0.451	0.245	0.452	0.246
within R-squared	0.089	0.062	0.091	0.063	0.084	0.046	0.085	0.047
Observations	12546	15225	12546	15225	10612	19385	10612	19385
Number of Ethnicities	85	95	85	95	66	139	66	139
Panel B: Local Regressions (100km bandwidth)								
Institutional Quality	0.0450	0.0040	0.1111*	0.0153	0.1140*	-0.0005	0.1564**	0.0112
Double-clustered s.e.	(0.0618)	(0.0239)	(0.0651)	(0.0269)	(0.0641)	(0.0235)	(0.0620)	(0.0258)
Adjusted R-squared	0.407	0.195	0.410	0.196	0.445	0.225	0.447	0.225
within R-squared	0.085	0.044	0.087	0.044	0.090	0.036	0.092	0.036
Observations	6129	8120	6129	8120	4655	10987	4655	10987
Number of Ethnicities	81	94	81	94	60	137	60	137
Ethnicity Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pop. Dens. & Area	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location & Geography	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports 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), (2), (5), and (6)) and control of corruption index (in columns (3), (4), (7), and (8)) in areas of partitioned ethnicities. The unit of analysis is a pixel of 0.125 x 0.125 decimal degrees (around 12 x 12 kilometers) within partitioned ethnicities. The dependent variable is a dummy variable that takes the value one if the pixel is lit and zero otherwise. In odd-numbered columns we restrict estimation to pixels of partitioned ethnic areas that are close to the capital city. In even-numbered columns we restrict estimation to pixels of partitioned ethnic areas that are far from the capital city. In columns (1)-(4) we use as a cut-off the median value of absolute distance to the capital across all African pixels, while in columns (5)-(8) we use as a cut-off the median value of the within-country distance to the respective capital city.

In Panel A we use all pixels of partitioned ethnicities. In Panel B we report local linear regressions restricting estimation to pixels of partitioned ethnicities within 50 kilometers from each side of the national border (total 100 kilometers). All specifications include a vector of ethnicity fixed effects (constants not reported). In all specifications we control for log population density, log land area, distance of the centroid of each pixel from the capital city of each country (in columns (5)-(8) we use relative distance to the capital city of each country), distance from the closest sea coast, distance from the national border, an indicator for pixels with a water body (lakes, rivers, and other streams)), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. Besides the adjusted R-squared, the table also reports the within R-squared (defined as the difference of the overall R-squared minus the overall R-squared of a model simply with ethnicity fixed effects). 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 9: National Institutions, Distance to the Capital, and Regional Development
Pixel-Level Estimates in the Full Sample

	Absolute Distance to the Capital			Relative Distance to the Capital		
	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	0.1434***	__	__	0.1790***	__	__
Double-clustered s.e.	(0.0426)			(0.0472)		
Distance to the Capital	-0.1352***	-0.2103***	-0.1097**	-0.3098***	-0.2530***	-0.1457***
Double-clustered s.e.	(0.0511)	(0.0404)	(0.0534)	(0.0846)	(0.0502)	(0.0499)
Rule of Law X	-0.1243***	-0.1513***	-0.0850**	-0.2297***	-0.1850***	-0.1042**
Distance to the Capital	(0.0363)	(0.0315)	(0.0352)	(0.0624)	(0.0461)	(0.0439)
Adjusted R-squared	0.261	0.362	0.429	0.267	0.362	0.429
Observations	116412	116412	116412	116412	116412	116412
Pixel Area & Pop. Dens.	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	No	No	Yes	No
Country-Ethnicity FE	No	No	Yes	No	No	Yes

The table reports OLS estimates associating regional development with distance to the capital city and contemporary national institutions, as reflected in World Bank's Governance Matters rule of law index across the full sample of country-ethnic homelands in Africa. The unit of analysis is pixels of 0.125 x 0.125 decimal degrees (around 12 x 12 kilometers). The dependent variable is a dummy variable that takes the value one if the pixel is lit and zero otherwise.
Columns (1) and (5) give unconditional estimates. In columns (2)-(4) and (6)-(8) we control for log (population density), log land area, distance of the centroid of each pixel from the capital city of each country, distance from the closest sea coast, distance from the national border, an indicator for pixels with a water body (lakes, rivers, and other streams)), land suitability for agriculture, elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The specifications in columns (3) and (7) include a vector of country fixed effects (constants not reported). The specifications in columns (4) and (8) include a vector of country-ethnicity fixed effects (constants not reported). 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 10: Law Enforcement and Ethnic Identification as Function of Distance to the Capital
Afrobarometer Sample. Country-Fixed-Effects Estimates

	Law Enforcement 1: Serious Crime				Law Enforcement 2: Tax Evasion				National versus Ethnic Identification			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Absolute Distance	-0.1238***	-0.1027**			-0.1118***	-0.1006***			-0.0850**	-0.0740*		
Double-clustered s.e.	(0.0446)	(0.0408)			(0.0410)	(0.0386)			(0.0414)	(0.0382)		
Relative Distance			-0.0860**	-0.0665*			-0.0655*	-0.0563*			-0.0949**	-0.0819**
Double-clustered s.e.			(0.0375)	(0.0354)			(0.0341)	(0.0326)			(0.0411)	(0.0379)
Adjusted R-squared	0.069	0.075	0.068	0.075	0.053	0.062	0.052	0.062	0.096	0.108	0.097	0.109
Observations	20466	19862	20466	19862	20842	20230	20842	20230	20125	19503	20125	19503
Individual Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports country fixed effects OLS estimates, associating legal enforcement (in columns (1)-(8)) and ethnic identification (in columns (9)-(12)) as reflected by individual responses in the Afrobarometer Surveys (3rd round) at the individual level with distance to the capital.
The dependent variable is columns (1)-(4) is an ordered index (range from 1 to 4) based on individual responses on how likely it is that authorities will enforce the law if a person like the respondent committed a serious crime. The dependent variable is columns (5)-(8) is an ordered index (range from 1 to 4) based on individual responses on how likely it is that authorities will enforce the law if a person like the respondent did not pay taxes on income earned. For both measures a score of 1 indicates "not at all likely", a score of 2 indicates "not very likely", a score of 3 indicates "likely" and a score of 4 indicates "very likely". The dependent variable in columns (9)-(12) is a dummy variable that takes on the value of one if the respondent replies that she identifies more strongly with her nation than the ethnicity she belongs to.
Even-numbered columns include a rich set of individual controls. The individual-level conditioning set includes age, age squared, a gender indicator variable, an urban indicator, 22 religion fixed effects, 25 occupation fixed effects, 5 living conditions fixed effects, and 9 education fixed effects. All specifications include a set of country fixed effects (constants not reported).
The Data Appendix gives detailed variable definitions and data sources. Below the estimates we report in parentheses double-clustered standard errors at the enumeration area and at the ethnicity level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.