DISCUSSION PAPER SERIES

No. 8088

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

Stelios Michalopoulos and Elias Papaioannou

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8088.asp**

ISSN 0265-8003

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

Stelios Michalopoulos, Tufts University
Elias Papaioannou, Dartmouth College and CEPR

Discussion Paper No. 8088
November 2010

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8088

November 2010

ABSTRACT

Divide and Rule or the Rule of the Divided? Evidence from Africa*

We investigate jointly the importance of contemporary country-level institutional structures and local ethnicity-specific pre-colonial institutions in shaping comparative regional development in Africa. We utilize information on the spatial distribution of African ethnicities before colonization and exploit within ethnicity (across countries) and within-country (across ethnicities) regional variation in economic performance, as proxied by satellite light density at night. The fact that political boundaries across the African landscape partitioned ethnic groups in different countries, thus subjecting identical cultures to different country-level institutions, offers a regression discontinuity framework. After identifying the partitioned ethnicities we document a positive cross-sectional association between national institutions and regional economic development. However, our ethnicity fixed-effects specifications show that differences in countrywide institutional arrangements do not explain differences in regional economic performance within ethnic groups. In contrast, we document that local ethnic traits proxied by tribal pre-colonial political institutions and class stratification exert even today a significant effect on regional development. The positive within country effect of pre-colonial institutions also obtains in regions of partitioned ethnicities along the national boundaries.

JEL Classification: N17, O10, O40, O43 and Z10
Keywords: Africa, borders, development, ethnicities, institutions

Stelios Michalopoulos
Department of Economics
Tufts University
8 Upper Campus Rd, Braker Hall 302
Medford, MA 02155
USA

Email:
Stelios.Michalopoulos@tufts.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=165801

Elias Papaioannou
Dartmouth College
Economics Department
6106 Rockefeller Hall
Hanover, NH 03755
USA

Email:
Elias.Papaioannou@dartmouth.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=156542

* We thank seminar participants at Dartmouth, Tufts, Oxford, Vienna, Brown, Harvard, NYU, the CEPR Development Economics Workshop, 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, Jeremiah Dittmar, David Weil, Sandip Sukhtankar, Quamrul Ashraf, Oded Galor, Ed Kutsoati, Pauline Grosjean, Enrico Perotti, Pedro Dal Bó, Nathan Nunn, Raquel Fernandez, and Enrico Spolaore. We are particularly thankful to Andy Zeitlin, Melissa Dell, and Daron Acemoglu for detailed comments and useful suggestions. We also thank Nathan Nunn for providing the digitized version of Murdock's Tribal Map of Africa. All errors are our sole responsibility..

Submitted 23 October 2010

1 Introduction

In recent years there has been a surge of empirical research on the determinants of African and more generally global under-development and state failures. The predominant institutional view suggests that poorly performing national institutional structures, such as lack of constraints on the executive and poor property rights protection, as well as inefficient legal and court systems are the ultimate causes of under-development (see Acemoglu *et al.* (2001, 2002, 2005) on the former and La Porta *et al.* (1998, 1999) on the latter). This body of research puts an emphasis on the impact of colonization on contemporary country-level institutions and in turn on economic development. Yet many downplay the importance of colonial and contemporary institutional structures in Africa. Recent works on weak and strong states emphasize the limited state capacity of most African states and their inability to provide public goods, collect taxes, and enforce contracts (Acemoglu (2005); Besley and Persson (2008, 2010)). The inability of African governments to broadcast power outside the capital cities has led many influential African scholars to highlight the role of pre-colonial ethnic-specific institutional and cultural traits (Herbst (2000); Boone (2003)). They argue that the presence of the Europeans in Africa was (with some exceptions) quite limited both regarding timing and location. As a result of the negligible penetration of Europeans in the mainland and the poor network infrastructure that has endured after independence, it is local tribal level, rather than national institutional structures, that shape African development.

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

In contrast to most previous works that have relied on cross-country data and methods, we tackle these questions exploiting within-country and within-ethnicity regional variation across approximately $1,200$ African ethnic regions. We utilize data from the pioneering work of Murdock (1959, 1967), who through extensive field work has produced a map portraying the spatial distribution of ethnicities (Figure $1a$) as well as quantitative information on the economy, institutions, and cultural traits for many ethnic groups before the European colonization of Africa. To overcome the paucity of regional data across African ethnicities, we follow Henderson, Storeygard, and Weil (2009) and measure regional economic development at the ethnicity-country level using satellite data on light density at night, which are available at

a very fine disaggregation.

Figure 1*a*: Ethnic Boundaries

Figure 1*b*: Ethnic and Country Boundaries

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

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

[1]There is no ambiguity among African scholars and historians that almost all of African borders were artificially drawn. For example, the borders of Congo and of the other Central African countries were designed before Europeans even sent missionaries to explore the area. Huillery (2009) shows that the French did not respect ethnicities' homelands when designing the administrative areas of French West Africa.

contemporary national institutions. The regression discontinuity (RD) approach allows us to account for geography, the disease environment, and other ecological features at a very fine level (see Dell (2009)). By comparing development in adjacent border regions which belong to the historical homeland of the same ethnic group (see Figures $2a - 2d$ for examples), allows us to control effectively for culture and other ethnic-specific traits. Our results show that there is no systematic relationship between country-wide differences in institutions and regional economic performance within partitioned ethnicities Africa.

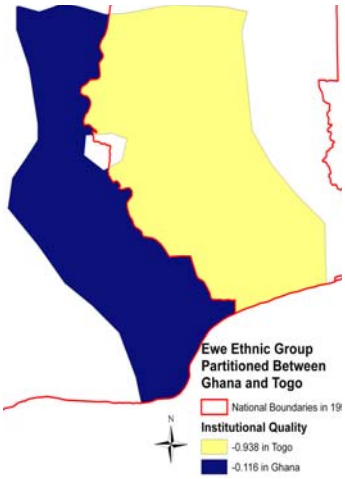

Figure $2a$

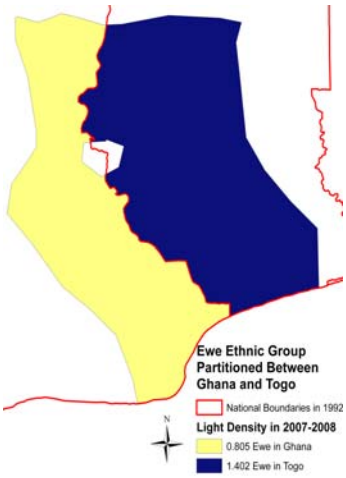

Figure $2b$

Figure $2c$

Figure $2d$

We then turn our focus on the economic impact of pre-colonial tribal institutions and societal arrangements. Our analysis establishes that political centralization and statehood experience before the advent of European colonizers correlates significantly with contemporary development. The same pattern applies to class stratification, a societal trait that has been linked to property rights protection (e.g. Rudmin (1995); Acemoglu, Bautista, Querubin, and Robinson (2008)). The strong positive correlation between ethnic pre-colonial institutions

(and class stratification) and regional development prevails when we account at a very fine level for the disease environment, land suitability for agriculture and elevation among others. Pre-colonial ethnic institutions correlate significantly with regional development even when we solely utilize the within country variation, so as to account for unobserved country-level attributes. We also find that the positive effect of local institutions on regional development persists across ethnic groups partitioned by the national boundaries. This suggests that along these territories characterized by the negligible penetration of countrywide policies it is the tribal institutions, determined well before the colonial era, that shape contemporary development.

Related Literature Our research nests and advances over many strands of literature that examine the historical roots of economic development in Africa and elsewhere. First, an influential body of research asserts that through persistence, the institutions that European powers established in the eve of colonization are the deep roots of contemporary development. While there is ambiguity on the exact mechanisms via which colonization affected African (and more generally non-European) development, there is an agreement that the type of colonization and the identity of the colonizing power had long-lasting effects on institutional development (e.g. Acemoglu *et al.* (2001, 2002); La Porta *et al.* (1997, 1998, 1999); Alcalá and Ciccone (2004); Glaeser, Porta, de Silanes, and Shleifer (2004); see Acemoglu, Johnson, and Robinson (2005) for a review). Yet in spite of the ingenious instrumental variables identification schemes employed in the cross-country literature, it is always hard to pin down the exact mechanism. This is because colonization might affect development via channels other than the institutional legacy and because omitted variables and unobserved heterogeneity are always major concerns in cross-country approaches.[2]

Our regional focus adds to a vibrant body of research that takes a micro approach examining the within-country impact of institutions (e.g. Banerjee and Iyer (2005); Iyer (2010); Dell (2009); Huillery (2009); see Acemoglu and Dell (2008) and La-Porta, de Silanes, Shleifer, and Vishny (2010) for recent evidence across administrative regions within countries). Our identification scheme on the impact of the national institutions explores discontinuities across the border within partitioned ethnicities. A defining characteristic of the empirical design is that it keeps all ethnicity-specific factors constant. This is key as recent works have shown that traits related to culture, social capital and genetic distance/diversity correlate significantly with economic and institutional development even within countries (e.g. Guiso, Sapienza, and

[2]For example, La Porta *et al.* (2008) argue, one should be cautious when using legal origin indicators to instrument legal institutions, because the legal tradition has also shaped regulation, state involvement in the economy, and the polity (see also Nunn (2009) and Pande and Udry (2006) on this point). Huillery (2009) for example, using administrative data from former French West Africa shows a significant effect of colonization on contemporary development working through early colonial investments.

Zingales (2008); Tabellini (2010); Spolaore and Wacziarg (2009); Ashraf and Galor (2008)).[3] Our results suggest that it may be premature to conclude from cross-sectional studies that a higher degree of countrywide rule of law and a lower level of corruption translate to higher levels of regional development in Africa.

Second, our findings advance the literature on the importance of pre-colonial institutional and cultural features in African development (Herbst (2000); Bates (1981, 1983); Boone (2003); see Robinson (2002) for a critical review). Anthropologists have shown that there were marked differences in institutional and social traits across African regions at the time of colonization (Murdock (1959, 1967)). There were noteworthy differences on political centralization, land rights, and the power of local chiefs, among others. As colonizers did not expand their power in remote areas far from the capital cities and the coastline, such local institutions were preserved and were instrumental even during the half century period of colonial rule (roughly $1890 - 1940$). Moreover, many African case studies stress the ongoing crucial role of ethnic institutions and traditions (e.g. Dowden (2008)). Gennaioli and Rainer (2007) use ethnicity-level data from Murdock's Ethnolinguistic Atlas (1967) on how centralized ethnic groups were before colonization, and construct a country-level measure of political centralization which is shown to be positively correlated with various proxy measures of economic development (see also Englebert (2000)).

The African historiography has proposed many channels on why ethnic institutions still matter. Gennaioli and Rainer (2006) and Boone (2003) argue that in centralized societies there is a high degree of political accountability of local chiefs. Others have argued that centralized societies were quicker in adopting growth enhancing Western technologies and habits, because the colonizers collaborated more strongly with politically and socially complex ethnic groups (Schapera (1956, 1970)). Herbst (2000) and Migdal (1988) stress the role of ethnic class stratification and political centralization in establishing well-defined and secure land rights (see also Golstein and Udry (2008)). Furthermore, complex tribal societies with strong political institutions seem to have been more successful in getting concessions both from colonial powers and from national governments after independence.

We improve upon this body of research by showing that pre-colonial institutions exert a positive effect on regional development even when we control for local geography and most importantly when we condition on country fixed-effects. Accounting for common to all ethnicities country factors is central, as Gennaioli and Rainer (2006) show a positive cross-

[3] There is a large literature in education, urban and trade economics that concentrates on bordering regions to investigate the effect of various policies (e.g. Holmes (1998), McCallum (1995), Gopinath, Gourinchas, Hsieh, and Li (2010) among others). The rationale is that focusing on adjacent regions, while accounting for observable characteristics, should neutralize any local unobservable differences that would otherwise contaminate inference.

country correlation between ethnic-specific institutional indicators and current measures of institutional development; thus is not clear from the cross-country evidence whether pre-colonial local structures and/or nationwide institutions affect current economic performance. Moreover, controlling for geography at a fine level is essential as studies on African institutional development argue that pre-colonial political centralization was driven by the suitability of land for agriculture, population density, and other geographic features (e.g. Bates (1981); Fenske (2009)). Our results offer strong support to those emphasizing the importance of pre-colonial ethnicity-specific institutions in current times. In this regard they are in line with recent empirical studies showing that historically determined socioeconomic and political factors have persistent effects on comparative development (examples include the forced labor practices of Spanish colonizers in Peru (Dell (2009)); the formation of city-states in Italy during the late period of the Middle Ages (Guiso, Sapienza, and Zingales (2008)); 19th century inequality in Colombia (Acemoglu, Bautista, Querubin, and Robinson (2008)); the type of colonization and early inequality in Brazil (Naritomi, Soares, and Assunção (2009)); the type of the colonial tax system in Nigeria (Berger (2009); see Nunn (2009) for a review).

The uncovered evidence on the limited effect of national institutions on regional development relates to works on state capacity (e.g. Tilly (1985); Migdal (1988); Acemoglu (2005); Acemoglu, Ticchi, and Vindigni (2009); Besley and Persson (2008, 2010)). Likewise, the finding that the positive correlation between national institutions and regional development weakens in border areas has implications for the literature on the optimal country formation (e.g. Alesina and Spolaore (2003); Spolaore and Wacziarg (2005)).

Finally, this study contributes to a large body of work on the roots of contemporary African development. Nunn (2008) and Nunn and Puga (2010) stress the importance of the slave trade, while Alesina, Easterly, and Matuszeski (2010) and Englebert, Tarango, and Carter (2002) show a significant negative impact on development of improper border design and state artificiality. A natural message coming out from our analysis is that research which aims to understand regional development in Africa needs to focus on the ethnicity level rather than the country level. In this regard our work is related to recent country-specific studies that stress the role of local chiefs in the provision of public goods in Africa (Miguel and Gugerty (2005); Franck and Rainer (2009); Glennerster, Miguel, and Rothenberg (2010)).

Paper Structure In the next section we discuss the luminosity data that we use to proxy regional development and present the pre-colonial ethnic institutional measures. In section 3 we detail our econometric methodology. We present the regression discontinuity design and discuss estimation and inference. Section 4 reports our main results on the effect

of contemporary national institutions and pre-colonial ethnic institutional features on regional development. Section 5 provides further evidence. Section 6 summarizes and concludes.

2 Data

2.1 Data on partitioning

The starting point in compiling our dataset is George Peter Murdock's (1959) Ethnolinguistic map that reports the spatial distribution of ethnicities across Africa. Using extensive field work and other resources, Murdock explicitly tried to map ethnicities before (or around) European colonization in the mid/late 19th century. Murdock's map (reproduced in Figure 1a) includes 843 tribal areas (the mapped ethnicities correspond roughly to levels $7-8$ of the Ethnologue's language family tree); yet 8 areas are classified as uninhabited upon colonization and are therefore not considered in our analysis.[4] One may wonder how much the spatial distribution of ethnicities across the continent has changed over the past $100-150$ years. Reassuringly, using individual data from the Afrobarometer Nunn and Wantchekon (2009) show a strong correlation (around 0.62) between the location of the respondents in 2005 and the historical homeland of their ethnicity as identified in Murdock's (1959) map. In the same vein, Glennerster, Miguel, and Rothenberg (2010) document that following the massive displacement that took place during the $1991-2002$ civil war in Sierra Leone there has been a systematic movement of individuals towards the areas of their ethnic group's historical homeland. In fact, in rural areas the historical ethnic diversity strongly predicts current diversity, with a coefficient estimate of 0.80.

We project on top of Murdock's ethnolinguistic map the Digital Chart of the World (Figure 1b) that portrays contemporary national boundaries. This allows to identify in a systematic way partitioned ethnicities across Africa. Appendix Table 1 reports split groups, defined as groups where at least 10% of their historical homeland belongs to more than one contemporary states. Our procedure identifies most major ethnic groups that have been separated by African borders. For example, the Maasai were partitioned between Kenya and Tanzania (shares 62% and 38% respectively), the Anyi between Ghana and the Ivory Coast (shares 58% and 42%), and the Chewa between Mozambique (50%), Malawi (34%), and Zimbabwe (16%). We also checked whether our codification of partitioned ethnicities is in line with Asiwaju (1985), who provides the only (to our knowledge) codification of split ethnicities in Africa. Our strategy identifies almost all ethnic groups that Asiwaju (1985) lists as partitioned.

[4] In the empirical analysis we also eliminate the Guanche, a small group in the Madeira islands that is currently part of Spain.

2.2 Satellite Light Density at Night

The nature of our study requires detailed data on economic development at the grid level. To the best of our knowledge, geocoded high resolution measures of economic development spanning Africa are not readily available. To overcome this issue we follow Henderson, Storeygard, and Weil (2009) and use satellite data on light density at night to proxy for local economic activity.

Data come from the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) that reports images of the earth at night captured from 20:00 to 21:30 local time from an altitude of 830 km. The satellite detects lights from human settlements, fires, gas flares, heavily lit fishing boats, lightning, and the aurora. The measure is a six-bit $(0-63)$ digital number calculated for every 30-second area pixel (approximately 1 square kilometer). The resulting data are a series of annual composite images. These are created by overlaying all images captured during a calendar year, dropping images where lights are shrouded by cloud cover or overpowered by the aurora or solar glare (near the poles), and removing ephemeral lights (like fires, lightning and other noise). The result is a series of global images of time stable night lights.[5] Using these data we construct average light density per square kilometer for 2007 and 2008 at the desired level of aggregation (ethnicity-country).[6] We do so by averaging across pixels that fall within the historical homeland of each ethnic group in each country (using the median value yields similar results).

This high resolution dataset makes the data uniquely suited to spatial analyses of economic development in Africa for many reasons. First, most African countries have low quality income statistics, even at the national level.[7] Second, we lack data on regional income or value added for many African countries. And while there are some regional proxies of poverty and health (see Young (2009)), these data do not map to our ethnicity level unit of analysis. Third, by using light density we also capture the economic activities of the underground economy, which are not reflected in the aggregate statistics. As the share of the shadow economy is high in Africa (e.g. La-Porta and Shleifer (2008)), the usage of luminosity data is particularly

[5] See Henderson, Storeygard, and Weil (2009), Min (2008), Chen and Nordhaus (2010) and the references therein for technical details on measurement error of the light data. Satellite data on lights are subject to overglow or blooming, which means that lights tend to appear larger than they actually are, especially for bright lights and over water and snow. Thus, water area is a standard control. Many of these issues, however are less pressing within Africa since there are few instances of top-coding, no long summer nights and no snow.

[6] We use the mean value of the last two available observations (in 2007 and 2008) to average out any effect from variation in cloudiness and other weather conditions. The correlation between light intensity in 2007 and 2008 is greater than 0.95 and thus all our results go through if we use either the 2007 or the 2008 values.

[7] For example, the codebook of the Penn World Tables assigns the lowest two scores (out of four possible ratings) on data quality for all African countries. Actually, one may reasonably argue that light density is a superior measure of economic activity, due to problems in measuring output and prices across African countries (see Deaton and Heston (2010); Ciccone and Jarocinski (2010); Johnson, Larson, Papageorgiou, and Subramanian (2009)).

desirable in our setting.

Henderson, Storeygard, and Weil (2009) show that light density at night provides a good proxy of economic activity. They also establish a strong within-country correlation (both across time and across regions) between light density at night and GDP and consumption. Moreover, there is a strong association between luminosity and access to electricity and public goods provision, especially among low income countries (see Elvidge, Baugh, Kihn, Kroehl, and Davis (1997) and Min (2008)). As one of the main effects of institutions on development is via the provision of public goods, such as electrification, the use of the light density data allows us to directly assess this channel.

In Figures 3*a* and 3*b* we examine whether luminosity correlates with development across African countries. Figure 3*a* illustrates the unconditional correlation between log light density and log GDP per capita in 2000. There is a clear positive relationship. The R^2 is 0.35 and the point estimate is more than 6 standard errors larger than zero. Besides economic performance, light density also reflects urbanization. Figure 3*b* shows the relationship between log GDP per capita and log light density after partialling out the effect of log population density. The relationship between log light density and log GDP per capita becomes now even stronger (the regression coefficient increases from 0.31 to 0.47 and the t-stat jumps to 10).

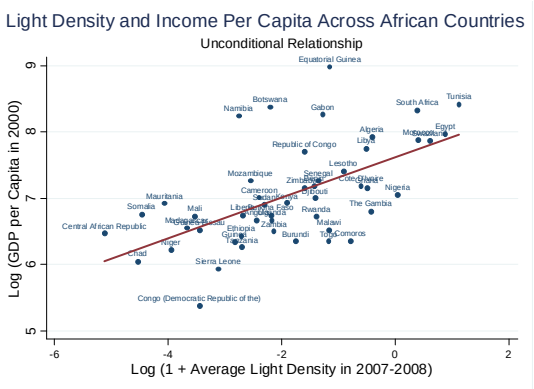

Figure 3*a*

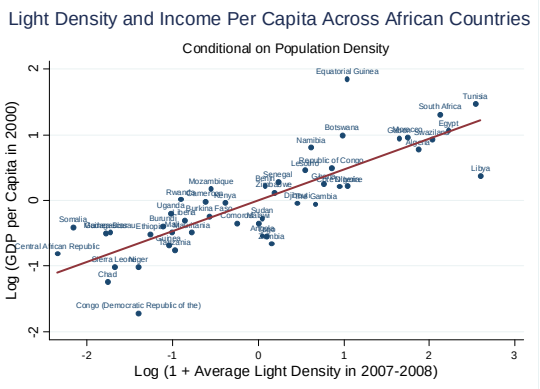

Figure 3*b*

We also examined the correlation between satellite light density and infant mortality, as an alternative proxy measure of development, across 264 African regions.[8] Figures 4*a* and 4*b* illustrate the negative correlation between log light density and infant mortality. The estimate is -9.44 with a t-stat of 9; when we condition on log population density, the estimate increases in absolute value (-14.89) retaining its significance at the 99% confidence level.

[8] The data on regional infant mortality come from the Center for International Earth Science Information Network (CIESIN) of Columbia University's Earth Institute available at: http://sedac.ciesin.columbia.edu/povmap/

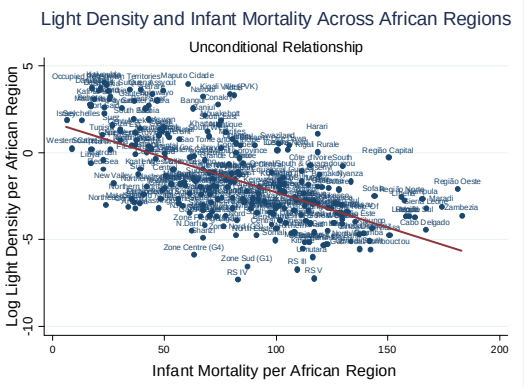

Figure 4*a*

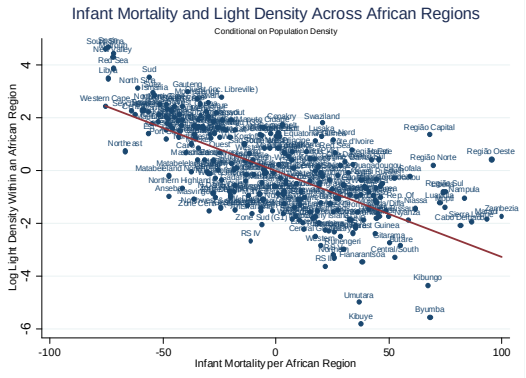

Figure 4*b*

The uncovered robust association between luminosity and development across African countries and regions suggests that in absence of direct measures of regional economic performance, light density is an informative proxy of local economic conditions.[9]

2.3 Ethnic Institutional Traits

In work following the mapping of the spatial distribution of African ethnicities, Murdock (1967) produced an Ethnographic Atlas (published in the anthropological journal *Ethnology*) that coded around 70 variables, capturing cultural, geographical, and economic characteristics of 1,270 ethnicities around the world. We assigned the 835 African ethnicities of Murdock's Map of 1959 to the ethnic groups in his Ethnolinguistic Atlas of 1967. As the two sources do not always use the same name for identifying ethnic groups we employed several sources and the updated version of Murdock's Atlas produced by J. Patrick Gray to match a total of 534 ethnicities.

We extract from this database two indicators proxying pre-colonial ethnicity-specific traits. First, following Gennaioli and Rainer (2006, 2007), we proxy pre-colonial political institutions using Murdock's (1967) index of "Jurisdictional Hierarchy beyond the Local Community Level". This is an ordered variable, ranging from 0 to 4, that describes the number of political jurisdictional hierarchies above the local community level for each society. A zero score indicates stateless societies "*lacking any form of centralized political organization.*" Scores 1 and 2 designate petty and larger-paramount chiefdoms, while 3 and 4 indicate groups that are part of large organized states. Murdock (1967) explicitly excludes colonial regimes (such as protectorates) and attempts to capture political centralization/complexity before Europeans started the settlement of Africa. Figure 5*a* illustrates the significant heterogeneity in political

[9] Even Chen and Nordhaus (2010) who are skeptical on the use of luminosity data to proxy GDP, argue that satellite light density is quite useful for countries with low quality statistics.

centralization across African groups. Examples of large ethnicities without any level of political organization above the local (village) level include the Bura and the Lango in Uganda. Examples of tribes belonging to small chiefdoms are the Mende in Sierra Leone and the Ibo of Nigeria. The Mbundu in Angola and the Zerma in Niger are part of large paramount chiefdoms, while the Yoruba in Nigeria and the Mossi in Burkina Faso are societies that were parts of large states before colonization. Since there are only 4 groups classified as being part of large complex states, i.e. have a score of 4, the Bubi in Equatorial Guinea, the Kafa in Ethiopia, and the Beduin Arabs in Morocco and Tunisia, these are merged with those tribes that are classified as being part of states i.e. get a score of 3.[10]

Second, we use Murdock's (1967) class stratification index. This is an ordered index that captures "*the degree of class differentiation, excluding purely political and religious statuses*". A score of 0 indicates "*absence of significant class distinctions among freemen*" A score of 1 suggests the presence of "*wealth distinctions, which however have not crystallized into distinct and hereditary social classes.*" Scores 2 and 3 indicate elite stratification and a dual/aristocratic stratification respectively. The index takes the maximum value of 4 when "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses.*"[11] Figure 5*b* plots class stratification. As with the jurisdictional hierarchy measure, there was significant heterogeneity in societal structure across Africa at the time of colonization. Highly stratified societies include the Yoruba and the Nupe in Nigeria, while examples of ethnicities without wealth or elite distinctions are the Maasai in Kenya and Tanzania, the Chewa in Malawi, and the Songo in Angola.

[10] This has no effect in the results. Following Gennaioli and Rainer (2006, 2007) we also construct and use a binary political centralization index that equals one when the jurisdictional hierarchy beyond the local level takes on the value of 2, 3, or 4 and zero when the jurisdictional hierarchy index equals 0 or 1.

[11] Results are similar if as Gennaioli and Rainer (2006, 2007), we classify societies in two groups: those that have no class distinctions, i.e. the class stratification takes on the value of 0 and societies characterized by some some type of class differentiation.

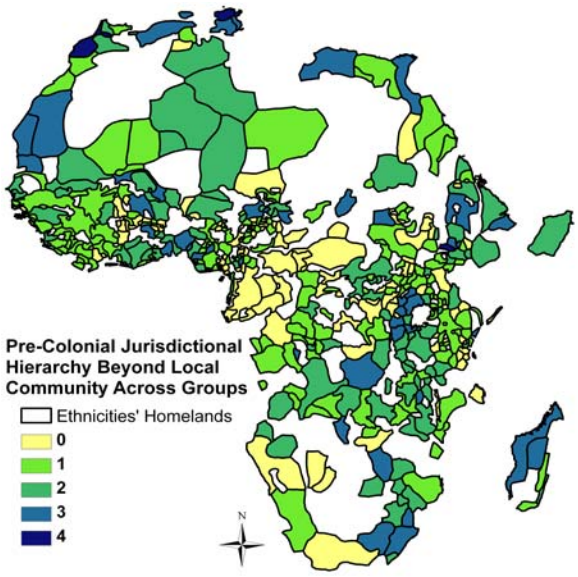

Figure 5*a*

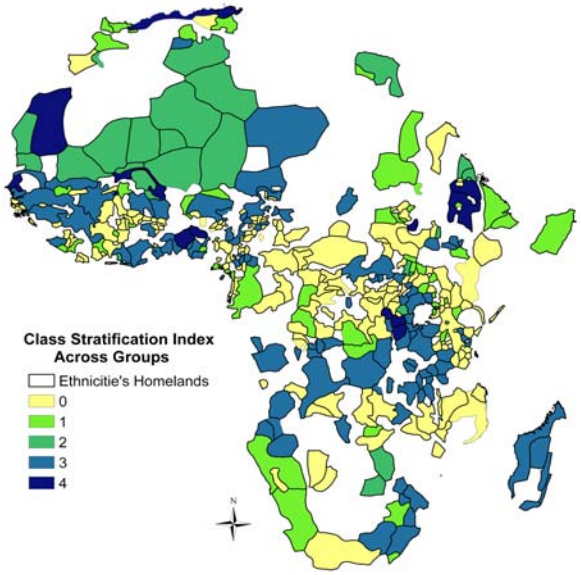

Figure 5*b*

2.4 National Institutions

For national institutions we rely on World Bank's Governance Matters Database (Kaufmann, Kraay, and Mastruzzi (2005)). The World Bank assembles numerous de facto institutional quality measures (originally compiled by various non-governmental organizations and risk assessment agencies) and aggregates them into six main categories via a principal component analysis, so as to minimize measurement error. For our benchmark estimates we use the rule of law index that reflects the effectiveness of the judiciary and the quality of property rights protection. As many studies on African development focus on graft, we also report results using the control of corruption index. Both variables have a theoretical minimum and maximum of -2.5 and $+2.5$, with higher values indicating better functioning institutions and less corruption.[12] In our sample the countries with the minimum values in rule of law are Somalia (-1.91) and the Democratic Republic of Congo (-1.84), while Namibia (0.64) and Botswana (0.71) are the most institutionally developed countries. Table 1-Panel *A* reports descriptive statistics for all variables employed in the empirical analysis.

[12] Results are qualitatively similar if we use alternative measures of national institutions, like ICRG risk of expropriation, Polity's executive constraints index, or measures from World Bank's Doing Business Around the World Project.

3 Empirical Design

3.1 Specification

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

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

The dependent variable, $y_{i,c}$, is the level of local economic activity in the historical homeland of ethnic group i in country c, as proxied by light density at night. For ethnicities that fall into more than one country, each area of the partitioned group is assigned to the corresponding country c. For example, regional light density in the part of the Ewe in Ghana is matched to the rule of law value of Ghana, while the adjacent region of the Ewe in Togo is assigned the value of the rule of law in Togo. $LOCINST_i$ denotes local ethnic institutions, as reflected in the degree of jurisdictional hierarchy beyond the local level and class stratification.

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

A potential merit of our regional focus is that we can account properly for local geography. This is non-trivial as there is a fierce debate in the literature on the institutional origins of development on whether the strong correlation between institutional and economic development is driven by geographical features (such as distance from the sea, land fertility), the disease environment, or endowments (see for example Gallup, Sachs, and Mellinger (1999), Easterly and Levine (2003)). Our cross-regional analysis allows us to account for geographic differences at a fine level (as compared to cross-country studies). In many specifications we include a rich set of geographic controls, $X_{i,c}$, reflecting land endowments (elevation, total surface area, area under water) and ecological features (malaria stability index, suitability for agriculture). The Data Appendix gives detailed variable definitions and sources.

Crucially, we also introduce a cubic geographic polynomial ($f(GEO_{i,c})$) in distance of the centroid of each ethnic group i in country c from the capital city and distance from the closest sea coast (see also Dell (2009)).[13] As most contemporary capital cities in Africa were

[13] Letting x denote distance from the capital city and y denote distance from the closest sea coast, the polynomial becomes $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$. In the previous version of the paper we simply controlled for distance from the capital and distance from the sea finding similar results. We also experimented

established by Europeans during the colonial period, distance from the capital captures the impact of colonization and the limited penetration of national institutions due to the poor infrastructure (we formally explore this possibility below).[14] In the same spirit, and besides capturing the effect of trade, distance to the sea coast reflects contact with Europeans and therefore also the penetration of colonization. This is because during the colonial era (and the slave trades) Europeans mainly settled in coastal areas. The joint inclusion of the geographic polynomial ($f(GEO_{i,c})$) and the country fixed-effects (a_c) or ethnicity fixed-effects (a_i) makes the above specification a regression discontinuity (RD) type of analysis at the spatial dimension (we discuss the validity of the RD design in the next section). As the areas of the same ethnic group across the two (or more) countries share a common border, one could think of the within ethnicity approach as a "matching estimator" with regions matched to neighboring regions.

3.2 Technical Remarks

Before presenting the results, we discuss some technical issues regarding estimation and inference. First, a significant fraction of the observations on regional development takes on the value of zero. 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.[15] Out of the $1,218$ observations there are 356 country-tribe areas with zero light density. In order not to lose these observations in the LS estimation we use as the dependent variable the log of light density adding one ($y_{i,c} \equiv \ln(1 + LightDensity_{i,c})$). Since this transformation is not ideal, in our sensitivity analysis we estimate specifications ignoring unlit areas ($y_{i,c} \equiv \ln(LightDensity_{i,c})$) which is normally distributed. Looking at the "intensive margin" also guarantees that we investigate the role of institutions in explaining variation in economic performance across densely populated regions displaying non-trivial economic activity.[16]

Second, the dependent variable is highly skewed as besides the zeros, we have many observations close to zero and a few extreme observations in the right tail of the distribution. While the mean of satellite light density is 0.198 the median is more than ten times smaller, 0.018. This occurs because there are a few areas where light density is extremely high. For

with a second order and a fourth order polynomial finding similar results. We preferred the cubic polynomial because in almost all specifications all terms enter with significant coefficients.

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

[15] Note that we exclude from the analysis tribe-country observations with zero population density in 2000.

[16] Dimly lit areas have an average population density of 23.37 people per square kilometer whereas regions with positive light density have an average of 51.18.

example, we have 65 observations (5%) where log light density exceeds 1 and 13 observations (1%) where log light density exceeds 2. Thus in all tables we report trimmed specifications ignoring the upper 2% of the distribution of the dependent variable ($y_{i,c} < p(98)$).[17] Moreover, since linear models might not be appropriate for "corner-solution" specifications where a significant mass of the non-negative observations is (close to) zero, we also report specifications with Poisson ML (Wooldridge (2002)). The Poisson estimator is appealing, because it does not require log-linearizing the dependent variable and thus preserves the higher moments of the distribution, (see Silva and Tenreyro (2006) and Silva, Tenreyro, and Windmeijer (2010)).[18] In the previous version of the paper we also reported Tobit estimates that also account for censoring in the dependent variable. The Tobit specifications (not shown for brevity) deliver quite similar and if anything stronger results.

Third, in all specifications we report standard errors based on two-way clustering. We use the approach of Miller, A.Cameron, and Gelbach (2006) and cluster standard errors both at the country-level and at the ethnic-family level.[19] This accounts for two main concerns related to non-adjusted (or heteroskedasticity-adjusted) standard errors. First, within each country we have many ethnicities where the country-level rule of law and the corruption measures take the same value and thus clustering at the country-level is required (Moulton (1986)). Likewise, partitioned ethnicities appear more than one time in our sample and thus clustering at the ethnic family accounts for unobserved features within each ethnolinguistic cluster.[20] Moreover, as we report specifications using on the RHS tribal indicators that exhibit within ethnic family correlation, it is appropriate to also cluster standard errors at the ethnic-family level. Second, the multi-way clustering method accounts for arbitrary residual correlation within both dimensions and thus accounts for spatial correlation (Miller, A.Cameron, and Gelbach (2006) explicitly cite spatial correlation as an application of the multi-clustering approach). We also estimated models with one-way clustering of standard errors (at either dimension) as well as standard errors accounting for spatial correlation of an unknown form using Conley's (1999) method. The two-way clustering produces the largest in absolute value standard errors and thus yields the most conservative inference. Moreover, as in many specifications we include country or ethnicity fixed-effects this soaks up further the spatial correlation at each dimension.

[17] The results are similar if we trim the dependent variable at the 5% or 1%. We also estimated specifications winsorizing the dependent variable at the top 1%, 2%, and 5% finding similar results. To account for outliers we also estimated median regressions, finding similar (and if anything stronger) results.

[18] Since there is no evidence of overdispersion in the dependent variable, i.e. mean and variance of light density are 0.20 and 0.27 respectively, the Poisson estimator is preferred over negative binomial (though using the latter yields similar results).

[19] Murdock assigns the 835 groups into 96 ethnolinguistic clusters/families.

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

4 Main Results

4.1 Preliminary Evidence

Table 2 reports cross-sectional specifications that associate regional development with contemporary national and pre-colonial ethnic institutions. Panel A presents LS results, while Panel B reports Poisson ML estimates. Column (1) shows that there is a positive and significant correlation between the rule of law index and regional development. In column (2) we add the RD polynomial on distance of the centroid of each ethnicity from the capital city and distance from the closest sea coast ($f(GEO_{i,c})$). While all polynomial terms enter with significant coefficients, the estimate on rule of law retains its economic and statistical significance. In columns (3) and (4) we control for population density and geography. Conditioning on geography reduces the coefficient by 20% in the LS specifications and by 25% in the Poisson.[21] Yet the estimate retains significance at standard confidence levels. Overall, the correlations in (1)-(4) echo the findings of cross-country works; although the association between institutional quality and development weakens somewhat when one accounts for geography, it remains highly significant.

In columns (5) to (8) we associate regional development with ethnic pre-colonial institutions, as reflected in the jurisdictional hierarchy measure (the results are similar with the class stratification index). Column (5) reports unconditional estimates. The coefficient on the proxy of early local institutions is positive and significant at the 99% confidence level (in both LS and Poisson). Including the RD polynomial, population density, and the set of geographic controls (in columns (6)-(8)) has a moderate effect on the results. Although, the estimate drops when we control for population density and geography (elevation, malaria, land's suitability for agriculture, surface area, and area under water), it remains at least two standard errors above zero in all permutations.

In columns (9)-(12) we regress regional light density on both national and ethnic institutions. Given the positive correlation (0.16) between rule of law and jurisdictional hierarchy, this is useful so as to investigate the stability of the results in columns (1)-(8). Column (9) introduces both the rule of law index and the jurisdictional hierarchy measure. The unconditional LS estimate of rule of law in the sample of 669 ethnicity-country observations is 0.1395 (specification not shown). Once we control for the degree of jurisdictional hierarchy the estimate on rule of law retains its significance but falls by around $10\% - 15\%$. Likewise the

[21] Most of the geographic controls enter with significant estimates. The malaria stability index enters in most specifications with a statistically significant negative estimate. The coefficients on land area under water and suitability for agriculture are positive and in many specifications statistically significant. Elevation enters with a negative estimate, which is significant in some models. Yet when we include the regression discontinuity (RD) polynomial the explanatory power of the geography controls diminishes substantially.

coefficient on jurisdictional hierarchy is positive and highly significant, though its magnitude is somewhat smaller compared to the analogous specification in (5). A similar pattern obtains when we include in the specification the RD polynomial, population density, and the set of geographic-ecological controls (in (10)-(12)).

The LS coefficient in column (12) implies that a one point increase in the rule of law index (moving approximately from the institutional quality level of Angola to that of Gabon) is associated with a 12% increase in regional development. The Poisson coefficient in column (12) implies that a one point increase in rule of law (two standard deviations) is associated with higher light density of 0.76 points (approximately 1.5 standard deviations; see Table 1-Panel *A*). Turning now to the effect of pre-colonial institutions, the most conservative LS estimate (0.0406) implies that regional development increases by approximately 12% as one moves from areas where stateless societies reside to regions with ethnic groups featuring strong pre-colonial institutions (i.e. have a jurisdictional hierarchy index equal to 3). The Poisson estimates suggest that a one point increase in the jurisdictional hierarchy beyond the local community index (approximately one standard deviation) is associated with an increase in light density of 0.24 points (half a standard deviation). The preliminary results in Table 2 are informative about the broad data patterns. Yet these estimates do not identify the one way effect of neither contemporary national institutions nor ethnic historical institutional traits on regional development.

4.2 National Institutions and Regional Development

Identifying the causal impact of contemporary institutions on regional comparative development is a particularly demanding task; this is because, among other challenges, there are rarely otherwise identical cultures exposed to different institutional settings. Yet the arbitrary design of borders in Africa offers an ideal setting to isolate the effect of nationwide institutions from cultural traits and ethnic institutional norms.

There is significant variation in both the rule of law across African countries (Figure 6*a*) as well as in light density within partitioned ethnicities. Figure 6*b* shows ethnic groups where at least 10% of their territory has been partitioned by borders. Sharp border discontinuities in rule of law appear in several parts of Africa. For example, in the Botswana and Zimbabwe border (where the Hiechware, the Subia, and the Tlokwa are partitioned); across the Namibia and Angola border (where the Ambo are split); in the Egypt and Sudan border (where the Ababda reside); between Kenya and Somalia (where the Bararetta group resides); or between Gabon and Congo (where the Duma reside).

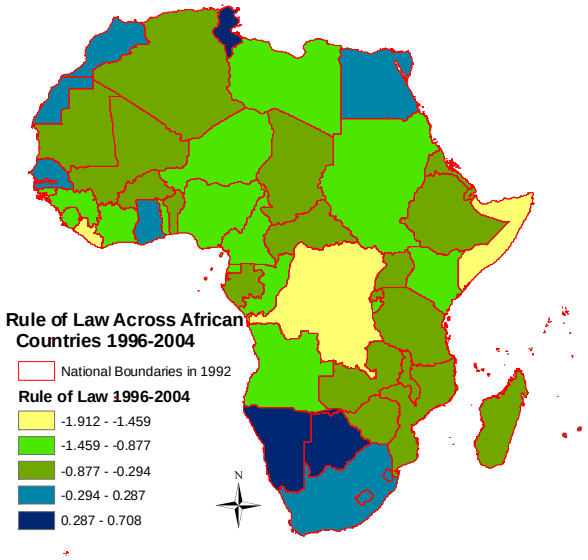

Figure 6*a*

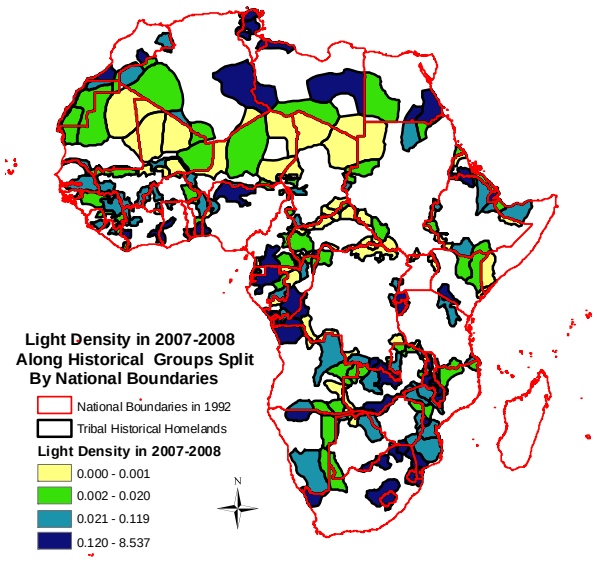

Figure 6*b*

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

Validity of the Regression Discontinuity Design Before investigating whether there are systematic differences in comparative development within tribal regions partitioned by national boundaries it is necessary to check the validity of the regression discontinuity design. The RD design requires that all relevant factors besides the treatment -national institutions in our application- vary smoothly at the border. Our focus on partitions of the same ethnic group (with the inclusion of ethnicity fixed-effects) allows us to account for cultural, religious, and other hard-to-measure tribe-specific traits. Yet a concern is that the geography or historical development of partitioned ethnicities in the relatively low institutional quality countries is systematically different from the partitions falling into the relatively high institutional quality ones. In this case the two (or more) areas of each partitioned ethnicity might not be appropriate counterfactuals. Thus we examined whether there are significant differences in geographical and historical characteristics of the (two or more) areas of partitioned ethnicities across the border.

Table 1-Panel *B* reports the mean value of variables reflecting geography (surface area, area under water, elevation, suitability for agriculture), the disease environment (malaria stability index), natural resources (presence of a diamond mine and an oil deposit), and early

development. Following Nunn and Wantchekon (2009) we proxy pre-colonial development using an indicator variable of whether a city with population larger than 20,000 was in a country-ethnic region in 1400. We proxy development around independence using population density in 1960. The units of observation are adjacent areas of partitioned ethnic groups. The neighboring regions are assigned to either the relatively high or the relatively low institutional quality country.[22] We also report the standard errors of the difference in means corrected for spatial correlation.

Table $1-B$ reveals some interesting patterns that support our RD design. First, there are no systematic differences in geography, the disease environment, and natural resources across the border. Second, there is no statistical difference in population density at independence across adjacent partitions of the same ethnic group. In a Malthusian regime where richer areas are more densely populated, this implies that there were no systematic differences in economic performance within split ethnicities whose partitions following independence would come to be subject to different national level institutions. Likewise, there are no differences in the proxy measure of early development (major city in 1400). Third, the only covariate that is significantly correlated with the treatment is distance from the capital city. Partitions falling in the relatively high rule of law countries appear to be closer to the capital city of the country. This correlation is driven by Sudan and the Democratic Republic of Congo, two of the largest in terms of size countries in Africa that also score very low in institutional development (see Hansson and Olsson (2010)). In light of this finding in the empirical analysis we include the RD polynomial to control in a flexible way for the location of each partition. Overall, the results in Table 1 - Panel B support the claim that national boundaries in Africa were arbitrarily drawn;[23] thus the border analysis offers a quasi-experimental setting for identifying the causal effect of contemporary national institutions on regional development.[24]

[22] For example the region of the Ababda in Sudan is assigned to the group of low institutional quality countries, while the adjacent region of the Ababda in Egypt is assigned to the group of relatively high institutional quality countries. Partitioned in more than two countries ethnic groups would appear as having three or more pairs of adjacent partitions. For example, the Azande, which is split into three countries features 3 pairs of adjacent partitions: the Sudan-Congo pair, the Sudan-Central African Republic one, and the partitions along the border of Congo and the Central African Republic.

[23] The history of the border between Nigeria and Cameroon offers an additional piece of qualitative evidence on the arbitrariness of the colonial border design. Although the overwhelming majority of the borders in Africa did not change following independence, the border separating Nigeria and Cameroon was redrawn in 1961 following a UN-sponsored referendum involving the local communities. Perhaps not surprisingly, our identification of partitioned groups reveals no historical homeland to be significantly impacted by this locally engineered border segment (Figure 6*b*).

[24] We also formally examined (running linear probability models, probits, and logits) whether geography, the disease environment, proxy measures of early development, and ethnic characteristics predict which ethnicities got partitioned or not. With the exceptions of total area and area under water which entered positively and significantly, none of the other variables correlates with the event of partitioning. Moreover the R^2 of these specifications (estimated across the 835 ethnicities) were well below 0.10.

Within Ethnicity Results Table 3 presents the results of the within ethnicity specifications. The cross-sectional estimates in columns (1), (3) and (5) echo the findings of Table 2. Across partitioned groups, areas falling in countries with higher institutional quality have systematically higher levels of development. In columns (2), (4) and (6) we add a vector of ethnicity constants (i.e. estimating (1) with a_i). The ethnicity fixed effects account for all hard-to-observe tribe-specific cultural and institutional factors that affect regional economic development. The coefficients on rule of law drop dramatically and become statistically indistinguishable from zero. The insignificance is not driven by a decrease in the precision of the estimated coefficients, since the standard errors remain largely unchanged (if anything in the Poisson specifications the standard errors fall). The results are similar when we use the control of corruption index to measure national institutional development (in columns (5)-(8)). Differences in light density within ethnicities partitioned by national boundaries cannot be explained by differences in national institutions. Note that although the LS estimates on rule of law and control of corruption without ethnicity fixed effects are not statistically different from those with ethnicity fixed effects, this is not the case for the Poisson estimates where the coefficients on national institutions are statistically different when ethnicity fixed effects are introduced. Figures 7*a* and 7*b* below illustrate the lack of a systematic within ethnicity correlation between light density and institutional quality at the national level.

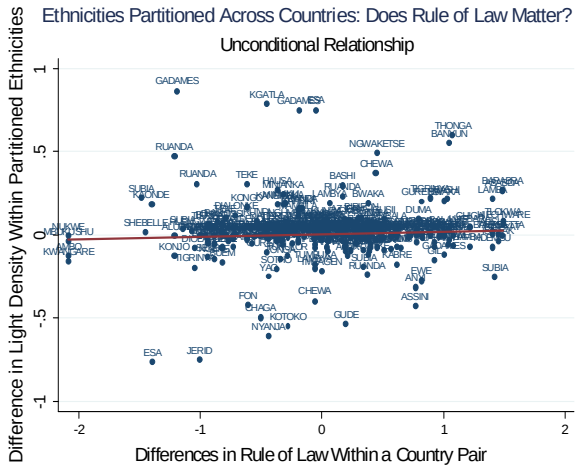

Figure 7*a*

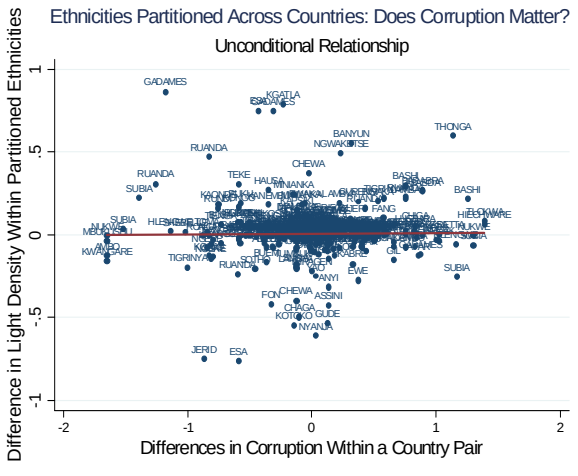

Figure 7*b*

In light of the results reported in Table 2 that the correlation between rule of law and light density weakens but remains significant when one controls for ethnic institutions, the findings in Table 4 suggest that the ethnicity fixed effects capture on the top of the measured tribal institutions other cultural and unobserved ethnic features. To gauge, for example, how much of the decline in the fixed effects estimate of rule of law is due to the inclusion of the precolonial

jurisdictional hierarchy consider the following: In the specification of column (1) in Table 3 when we restrict the estimation on partitioned groups for which information on tribal institutions is available the Poisson estimate on rule of law is 1.158. If we introduce ethnicity fixed-effects (column 2-Table 3), then the coefficient drops by 60% (0.472) and becomes insignificant. If we add the jurisdictional hierarchy index instead, the estimate on rule of law drops by 30% to 0.832. So, from the 60% drop in the magnitude of rule of law that is due to ethnicity fixed effects almost half of the decline can be ascribed to variation in precolonial ethnic institutions.

4.3 Precolonial Ethnic Institutions and Regional Development

The preliminary results in Table 2 reveal a significant association between the complexity of precolonial political structures and regional development. The correlation between pre-colonial ethnic institutions and regional development retained significance when we controlled for geography, the disease environment, and urbanization. Yet a concern with the previous estimates is that the positive correlation between local institutions and regional development is driven by country-level characteristics, reflecting national policies and institutions, or the type of colonization and the identity of the colonizing power, etc. In Table 4 we thus estimate within country specifications associating regional development with pre-colonial ethnic institutions (i.e. estimating (1) with a_c).

Jurisdictional hierarchy beyond the local community level Column (1) reports estimates on the effect of jurisdictional hierarchy on regional development. The LS (Poisson) coefficient is 0.054 (0.476); both estimates are significant at the 99% level. The estimates are only slightly smaller than the analogous unconditional specifications (reported in Table 2 column (5)), suggesting that common to all ethnicities country factors were not driving the positive correlation. In columns (2) and (3) we control for geography (with the set of geographic controls and the RD polynomial) and population density respectively. Column (4) reports the most restrictive specifications, where besides including country fixed-effects, we also condition on population density and geography. Accounting for geographical factors is important as there is a positive correlation between land suitability for agriculture and ethnic institutional development (see Fenske (2009)). The coefficient on the jurisdictional hierarchy beyond the local community level falls, but retains statistical significance at standard confidence levels. The LS estimate (0.035) implies that regional development decreases by approximately 10% as one moves from areas where ethnic groups with strong pre-colonial institutions reside (i.e. have a jurisdictional hierarchy index equal to 3; for example the Luba in Congo), to regions populated by ethnicities without any statehood experience before colonization (the index equals 0; for example the Songe in Congo). Likewise the Poisson specifications imply that moving from

regions populated by stateless societies to ethnic areas with strong precolonial institutions is associated with a one standard deviation increase in the dependent variable ($3 * 0.155 \simeq 0.47$).

Political centralization For comparability with Gennaioli and Rainer (2006, 2007) and Nunn (2008) in columns (5) to (8) we use an alternative indicator of pre-colonial institutions, based on the jurisdictional hierarchy index. Following these authors, we construct a dummy variable of pre-colonial political centralization (statehood) that takes the value of zero when Murdock's jurisdictional hierarchy indicates that the tribe lacks a centralized political organization or is part of a small chiefdom and 1 otherwise. Experimenting with the re-scaled index is also useful, because the aggregation might account for measurement error in the jurisdictional hierarchy (we formally examine the implications of measurement error in the next section). The within-country coefficient on the political centralization indicator variable in column (5) is positive and significant at standard confidence levels. The estimate retains significance, when we control for geography (in (6)), current levels of urbanization (in (7)) or both (in (8)). These results crucially advance the novel findings of Gennaioli and Rainer (2007), by showing that even when one accounts for regional geographical endowments and country fixed-effects, the correlation between pre-colonial political centralization and regional development remains strong. The estimates are supportive to Herbst's (2000) and Olson's (1987) conjecture on the importance of pre-colonial political institutions in the process of African development.[25] Moreover, these results are in line with the cross-country evidence of Bockstette, Chanda, and Putterman (2002) on the role of state antiquity and statehood experience on contemporary development.

Class stratification In columns (9)-(12) we report specifications using Murdock's (1967) class stratification index. While our focus is not on the role of the social structure per se, we use the class stratification index as a proxy for the presence and protection of property rights. Studies in sociology and anthropology document a strong correlation between ethnic class stratification and property rights protection, as well as strong political centralization (e.g. Rudmin (1995)). Indeed in our sample the class stratification index is strongly correlated with the jurisdictional hierarchy measure (0.63). The main explanation is that in weakly institutionalized societies economic inequality can lead to some form of property rights protection, as the elite has the incentive to establish constraints on the executive and against expropriation from

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

the masses (see Bates (1981)). In the same spirit Acemoglu, Bautista, Querubin, and Robinson (2008) argue that economic inequality might be conducive to development by constraining distortionary policies (such as expropriation), while Goldstein and Udry (2008) show a positive link between political power inside the local community (class stratification) and land tenure (property rights) in rural Gambian communities.

The coefficient on the class stratification index is positive and highly significant across all permutations. The LS estimate in column (12) implies that a movement of the index from communal societies (such as the Chamba in Nigeria where class stratification equals 0) to highly complex stratified societies (such as the Yoruba in Nigeria where the index equals 4) is associated with a 10% increase in regional development. The Poisson estimates imply that a 2 point increase in the class stratification index increases light density by one standard deviation. The positive association between class stratification and regional development, though surprising at first glance, is in line with recent works in Latin America. For example, Acemoglu, Bautista, Querubin, and Robinson (2008) and Naritomi, Soares, and Assunção (2009) document a significant positive cross-regional correlation between historical economic inequality and current development in Colombia and Brazil respectively. The positive within country effect of early class stratification on development is also in line with Dell's (2009) novel finding that the early concentration of large rural estates (the "haciendas"), in regions not affected by the colonial forced mining labor system in Peru ("the mita"), is associated with better public goods provision today.

The evidence in Table 4 suggests that tribal characteristics associated with the precolonial institutional arrangements and societal structure exert a significant effect on contemporary regional economic outcomes. Given that these ethnicity-specific characteristics remain robust to the inclusion of a host of correlates at a very fine level and country-specific unobservables highlights the importance of deeply rooted historical determinants of comparative economic development.

4.4 Sensitivity Analysis

4.4.1 Migration

African scholars and anecdotal evidence suggest that national boundaries across Africa are poorly enforced; this is due to poor demarcation, geographic conditions (desert areas in the North; rainforest in Central Africa), lack of border patrolling, and because African leaders do not bother when people move across the national borders. This poses a threat to our identification strategy regarding the effect of the national institutions. To the extent that people migrate to take advantage of higher incomes in regions with higher levels of institutional qual-

ity, mobility across national boundaries may attenuate income differences across the border. Furthermore, to the extent that mobility barriers are much lower within the same ethnicity compared to moving across areas of different ethnic groups, once we focus within partitions of the same ethnic group, then the ease of mobility would further attenuate any differences in regional development per capita caused by changes in the national level institutions. This scenario predicts that if institutions matter then as a result of the migration towards the partitions located in the high institutional quality country, population density should be systematically higher in the latter.

Although in many of the specifications in Table 3 we control for population density, we explore in detail this hypothesis. Table 5 reports specifications with log population density as the dependent variable. As the variable is normally distributed, we only report LS estimates. There is no systematic association between national institutions and population density in areas of partitioned ethnicities. This suggests that the insignificant within ethnicity relationship between country-level institutions and regional development is not driven by migration towards the partition located in the country with better functioning national institutions. Moreover, to the extent that population density reflects (to some degree at least) regional development, the insignificant within ethnicity coefficient on rule of law and corruption in columns (2), (3), (5), and (6) in Table 5 provides additional evidence that national level institutions are not systematically related to regional economic performance.

4.4.2 Robustness Checks

We performed many robustness checks to investigate the stability of the patterns shown in Tables 3 and 4**.** Table 6 summarizes the main sensitivity checks. For brevity, we report results proxying national institutions with the rule of law index and proxying pre-colonial ethnic institutions with the binary political centralization index.

Sub-Saharan Africa only Columns (1)-(4) report estimates excluding North Africa from the analysis to account for the different timing and type of colonization in Sub-Saharan countries as compared to North Africa. The Europeans had established relationships from the ancient times with North Africa, while contacts with most Sub-Saharan regions were limited till mid 19th century. Column (1) reveals a positive and significant correlation between rule of law and regional development. The coefficient is quite similar to the estimate in the full sample (see column (1) - Table 3). Yet once we include ethnicity fixed-effects, the coefficient on rule of law declines by 70% and becomes statistically indistinguishable from zero. Columns (3)-(4) report cross-sectional and within-country estimates exploring the effect of precolonial ethnic institutions on regional development in Sub-Saharan countries. There is a strong positive

correlation between pre-colonial political centralization and regional development, even when we include country fixed-effects and control for geography at a very fine level.

Large differences in institutional quality One may argue that the lack of within ethnicity (across the border) correlation between institutions and regional development is driven by the small differences in institutional quality among African countries. We thus repeated estimation adding squares and higher polynomial terms of the institutional quality indicators searching for nonlinearities, without detecting any significance. In columns (5)-(6) we report specifications estimated for two-way partitioned ethnic groups residing across country pairs with large (defined as higher than the median) differences in rule of law. Examples of ethnic groups partitioned between countries with large differences in institutional quality are present in many parts of Africa, including, among others, the Tabwa at the border of Congo with Zambia, the Suri between Ethiopia and Sudan, the Seke between Equatorial Guinea and Gabon, and the Nafana between Ivory Coast and Ghana. While the cross-sectional correlation between rule of law and light density is positive and significant, the correlation weakens considerably and becomes insignificant when we include ethnicity fixed-effects.

Additional Controls Although we control for geographical and ecological features at a very fine level, one may still argue that some other variable is driving the results. In columns (7)-(10) we include additional controls in an effort to mitigate concerns related to omitted variables bias. To account for the potential negative effect of natural resources on development (via spurring conflict for example), we augment the specification with indicator variables that take on the value one when a diamond mine or an oil/gas field is in the (country-ethnicity) area and zero otherwise. We also control for the pre-colonial level of urbanization, including a dummy variable that takes on the value one when a city with a population larger than $20,000$ in 1400 was in a tribal area. Additionally, to make sure that the estimates on precolonial institutions are not driven by differences in historical tribal levels of development we controlled for the pre-colonial settlement pattern of each ethnic group as recorded by Murdock (1967). The classification ranges from fully nomadic ethnic groups to groups living in complex settlements and the categories are listed in order of increasing economic and social development. Both the pre-colonial settlement pattern and the presence of a historical urban center are important controls because Bates (1983) argues that political centralization was higher in densely populated areas. We also include a dummy variable that takes on the value 1 for areas that belong to ethnic groups partitioned by the national border. Columns (7)-(10) present the results with this augmented set of controls. First, while in the cross-section of ethnicities there is a strong correlation between rule of law and regional development (in (7)), the correlation

weakens and turns insignificant when we focus within ethnicities (in column (8)). Second, the precolonial centralization index enters with a positive and significant coefficient both in the cross-sectional and the within-country estimation. The coefficient is also stable (in both LS and Poisson), quite similar to the more parsimonious specifications in Table 4.[26]

5 Further Evidence

5.1 Heterogeneous Effects of National Institutions

Besides perturbing the empirical specification to check the sensitivity of our results, we also searched for potential differential effects of institutional quality.[27]

Distance to the Capital City African historians and political scientists (e.g. Herbst (2000)) have long argued that the European's presence in Africa with some exceptions was limited to the coastline and the capital cities. Hence, colonial institutional arrangements, reflected through persistence on today's institutional quality, would have limited reach far from the capital cities. Along the same lines, several scholars have argued that due to the lack of the necessary infrastructure (roads, transportation system) and limited state capacity, nationwide institutions have minimal impact far from the capital cities (e.g. Dowden (2008)). To explore this hypothesis we estimated specifications associating country institutions with regional development separately for ethnic areas whose partitions are both close to or far from the capital city (using as a threshold the median distance of the centroid of each ethnic area from the capital city).

Table 7, columns (1)-(4) report the LS (in Panel A) and Poisson (in Panel B) estimates. The cross-sectional coefficients in (1)-(2) reveal an interesting pattern. While rule of law correlates positively with regional development in ethnic areas both close and far from the capital city, the estimate is much larger for ethnic areas close to the capital city. The LS coefficient on rule of law is ten times larger for partitioned ethnicities residing close to the capital city

[26] The coefficient on the diamond dummy is negative though not always statistically significant. The oil dummy enters with a positive and in most specifications significant estimate. While this seems to contradict the negative cross-country correlation, it is in line with the cross-region results in La Porta, Lopez-de-Silanes, and Shleifer (2010). The positive correlation between luminosity and oil field may also partially reflect fires of oil platforms as captured by the satellite. The unconditional estimate on the city in 1400 indicator is positive and significant suggesting that contemporary regional development is much higher in areas where a major city was before colonization. Yet the coefficient turns insignificant when we include the regression discontinuity polynomial or/and other geography control variables. The partitioned indicator enters with a negative estimate that in some model permutations is statistically significant. Finally, the settlement pattern enters positively, but does not attain significance at conventional levels.

[27] We also searched for potential interactions between national and ethnic institutions, without however detecting anything significant. Similarly, we interacted pre-colonial institutions with proxy measures of state capacity (e.g. Acemoglu (2005); Besley and Persson (2008)), such as tax revenues as a share of GDP and income tax receipts as a share of GDP, without however uncovering any robust patterns.

compared to those located far from it (0.21 and 0.02 respectively).[28] This finding is interesting in light of works on optimal country size (e.g. Alesina and Spolaore (2003)) and the emerging literature on state capacity (e.g. Acemoglu (2005); Besley and Persson (2008, 2010)).[29] Yet this difference as well as the coefficient estimates for both types of partitioned ethnic groups drop to zero when we exploit the within-ethnicity variation, thus replicating the pattern uncovered in Table 3.

Border Artificiality Influential scholars argue that colonization had a devastating long-run impact in Africa and other parts of the world because the political boundaries designed by Europeans partitioned ethnicities destroying the pre-existing social infrastructure (e.g. Dowden (2008); Alesina, Easterly, and Matuszeski (2010); Englebert, Tarango, and Carter (2002)). In Table 7 columns (5)-(8) we examine whether there is a differential effect of nationwide institutions on regional development depending on whether the partitioned ethnicities are separated by borders that are more likely to be natural as compared to boundaries that are more likely to be artificial. We differentiate between national boundaries using information on whether there are significant geographic barriers. We classify borders being natural where the terrain (across a 5 kilometer buffer zone from the border line) is either highly rugged or contains water barriers or is dominated by desert. The cross-sectional estimates in (5)-(6) show that the correlation between institutional quality and local development is somewhat stronger across ethnic areas partitioned by natural borders as compared to tribal areas partitioned by national borders that are more likely to be artificial. Nevertheless, when we include ethnicity fixed-effects the within ethnic group correlation between rule of law and regional development becomes insignificant for both types of borders.

5.2 Measurement Error in Precolonial Ethnic Institutions

The significant effect of pre-colonial ethnic institutions and class stratification on regional development is striking not only because it does not seem to be driven by geography, natural resources, endowments, and country unobservables, but because the underlying data almost certainly contain significant measurement error. While the compilation of the data was a life-long project for George Peter Murdock that involved field work and extensive research,

[28] This difference in the coefficients is statistically significant at the 99% level. In the previous version of the paper we obtained similar results by adding an interaction term between distance to the capital city and the rule of law index. The interaction term entered with a significantly negative estimate, suggesting that the correlation of rule of law index and local development decays for regions further away from the capital city.

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

given the limited resources and the nature of the endeavor, measurement error is most likely present. Since there is no a priori reason to expect that Murdock misclassified the institutional traits of ethnicities in a systematic way, our estimates are most likely attenuated. Although classical error-in-variables produces conservative estimates, it is intriguing to try to account for measurement error. We thus combine the two mismeasured proxies of precolonial local institutional in a 2SLS empirical framework, since under error orthogonality this approach yields unbiased estimates (Wooldridge (2002)).[30]

Table 8 reports country fixed-effects 2SLS specifications. To be conservative in all empirical models we include population density, the RD polynomial, and the standard set of geographic/ecological controls. For the specifications in columns (1) and (2) we use class stratification as an "instrument" for the jurisdictional hierarchy index and the political centralization index respectively. Reversely in column (3) we use the jurisdictional hierarchy beyond the local community level as an "instrument" for ethnic class stratification. Panel B illustrates the strong positive correlation (first-stage fit) between these proxy measures of pre-colonial institutions (the F-score of excluded instrument exceeds 70 in all specifications). Panel A reports the 2SLS estimates, while Panel C presents for comparability the corresponding OLS estimates. All three proxy measures of pre-colonial institutions enter the second stage with highly significant estimates. Moreover, in all specifications the 2SLS estimates are larger than the analogous LS estimates, suggesting that our previous estimates were conservative due to attenuation. The 2SLS coefficient on the jurisdictional hierarchy index (0.055) is 70% higher than the LS estimate (0.033), while the 2SLS coefficient on the political centralization index (0.117) is almost two times larger than the LS estimate (0.066). Likewise the 2SLS coefficient on class stratification (0.037) is also significantly larger than the LS estimate (0.021).

Table 8 columns (4)-(6) report 2SLS estimates associating log light density and precolonial ethnic institutions in areas where there is at least some light. Focusing on the intensive margin of light density is useful because it guarantees that we focus on areas with considerable economic activity. Moreover when we ignore unlit areas log light density is normally distributed and thus linear models are appropriate. The LS estimates (in Panel C) suggest that, conditional on the area being lit, regional development is significantly higher in areas where ethnicities with complex political institutions and societal structure reside. Similar to the pattern found for the maximum sample (in columns (1)-(3)), the 2SLS coefficients are significantly larger than the corresponding LS estimates suggesting that measurement error was producing attenuated

[30] To account for measurement error, we also run regressions using the principal component of the jurisdictional hierarchy and the class stratification index. The principal component that reflects the pre-existing complexity of the society enters in all permutations with a highly significant estimate implying larger effects than the estimates in Table 4.

estimates. While the 2SLS specifications might still not uncover the true effect (because the errors in the variables are correlated), they clearly suggest that if anything the effect of pre-colonial institutional traits on contemporary African development established in the previous Tables is quantitatively large.

5.3 The Rule of the Divided

Our finding that within partitioned ethnicities national level institutions do not explain differences in economic performance raises the question on whether these are just unruly areas or there is some other local institutional arrangement that shapes economic performance. Table 9 reports within country specifications examining the impact of ethnic institutional structures in regions of partitioned ethnic groups.[31] For completeness we report results with all three proxies of early ethnic-specific institutional traits. The coefficient on the jurisdictional hierarchy beyond the local community level in columns (1)-(2) is positive and highly significant. We obtain similar results when we use the binary political centralization index in columns (3)-(4). The LS estimate when we include all control variables and country fixed-effects implies that regional economic development is 9.5% higher in border regions populated by partitioned ethnicities characterized by hierarchical political centralization. Likewise, the Poisson specifications reveal that light density is 0.75 points (1.5 standard deviations) higher in border areas of centralized ethnicities as compared to areas of ethnicities with lack of political centralization. The coefficients for both the political centralization index and the jurisdictional hierarchy measure are if anything higher than the analogous estimates in the full sample (see Table 4). We obtain similar -though somewhat weaker- results when we use the class stratification index in columns (5)-(6)). Overall, the evidence suggests that historical ethnicity-specific institutional traits exert an important influence on comparative development across partitioned ethnic groups where formal countrywide institutions have been shown to lack explanatory power.

6 Conclusion

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

[31] We obtain similar results when we focus on all border groups (i.e. focusing not only on partitioned ethnicities but also on ethnicities that reside in areas adjacent to the national border).

areas. Exploiting within-country across-ethnicity variation in ethnic pre-colonial institutions as well as within ethnicity across-country variation on contemporary country level institutions, we document new empirical regularities on the role of institutional structures on African development.

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

Second, we explore the large heterogeneity in tribal institutions and examine their effect on regional development. In line with an influential conjecture among African scholars and historians, we show that ethnic pre-colonial institutions still exert a significant effect on contemporary regional development. The strong within country correlation between ethnic institutions, i.e. political centralization and class stratification, and regional development is also present along areas populated by ethnic groups partitioned by the national borders. Our findings contribute to the literature on the role of contemporary and historical determinants of economic development and suggest that in Africa the relevant unit of analysis is the ethnic group rather than the country.

Besides these findings, our codification of partitioned ethnic groups and the combination of high resolution regional data on development (such as satellite light density at night) with historical measures on culture and institutions provide a platform for subsequent research. One could employ our approach to shed light on the perennial debate regarding the fundamental determinants of comparative economic development across countries, examining for example the effect of human capital, public policies, and democracy, on economic performance. We intend to tackle some of these questions in future research.

7 Data Appendix

7.1 Variables at the ethnicity-country level

Light Density at Night: Light Density is calculated at a tribe-country level averaging light density observations across pixels that fall within the unit of analysis. To smooth weather variation we use the average of the values in 2007 and 2008.
Source: Available at http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html.

Population Density: Log population density per sq. km. in 2000 and in 1960. In the regressions we exclude unpopulated areas in 2000. *Source: UNESCO (1987). Available at: http://na.unep.net/datasets/datalist.php.*

Area: Log Surface area in 1000*s* of sq. km. *Source: Global Mapping International, Colorado Springs, Colorado, USA.*

Water Area: Log (1 + total area within an ethnic group district covered by rivers or lakes in sq. km.). *Source: Constructed using the "Inland water area features" dataset from Global Mapping International, Colorado Springs, Colorado, USA. Series name: Global Ministry Mapping System.*

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

Land Suitability for Agriculture: Average land quality for cultivation within the area of each ethnic-country observation. The index is the product of two components capturing the climatic and soil suitability for cultivation. *Source: Michalopoulos (2008); Original Source: Atlas of the Biosphere. Available at http://www.sage.wisc.edu/iamdata/grid_data_sel.php.*

Malaria Stability Index: The index takes into account the prevalence and type of mosquitoes indigenous to a region, their human biting rate, their daily survival rate, and their incubation period. The index has been constructed for 0.5 degree by 0.5 degree grid-cells globally. *Source: Kiszewski, Mellinger, Spielman, Malaney, Sachs, and Sachs (2004)*

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

Sea Distance: The geodesic distance of the centroid of each ethnic group in a country from the nearest coastline, measured in 1000*s* of km's. *Source: Global Mapping International, Colorado Springs, Colorado, USA. Series name: Global Ministry Mapping System. Series issue: Version 3.0*

Petroleum: Indicator variable that takes on the value of one if an oil field is found in the region of ethnic group i in country c. *Source: The Petroleum Dataset v.1.1 contains*

information on all known on-shore oil and gas deposits throughout the world. http://www.prio.no/CSCW/Datasets/Geographical-and-Resource/Petroleum-Dataset/Petroleum-Dataset-v11/

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

City in 1400: Indicator variable that takes on the value of one if a city with a population larger than $20,000$ in 1400 was in the area of ethnic group i in country c. *Source: Chandler (1987).*

Split: Indicator variable that equals 1 if at least 10% of the historical homeland of an ethnic group is partitioned into different countries. *Source: Calculated intersecting Murdock's (1959) ethnic map of Africa with the Digital Chart of the World (DCW) shapefile. The latter contains the polygons delineating the international boundaries in 1992.*

7.2 Country-level variables

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

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

7.3 Pre-Colonial Ethnicity-level variables

Jurisdictional Hierarchy beyond Local Community: Ordered variable ranging from 0 to 4 indicating the number of jurisdictional levels (political complexity) in each society above the local level**.** A 0 indicates stateless societies, 1 and 2 indicate petty and large paramount chiefdoms (or their equivalent), 3 and 4 indicate large states. Since we have only 3 ethnicities where the index equals 4 we assign to these ethnicities a score of 3. *Source: Murdock (1967); variable code in the Entholinguistic Atlas v32; A revised version of Murdock's Atlas has been*

made available by J. Patrick Gray at:
http://eclectic.ss.uci.edu/~drwhite/worldcul/EthnographicAtlasWCRevisedByWorldCultures.sav.

Centralization Indicator: This binary index takes the value 0 if the Jurisdictional Hierarchy beyond Local Community variable equals 0 or 1. The index takes on the value 1 if the Jurisdictional Hierarchy Beyond Local Community variable equals 2,3, and 4. This aggregation follows Gennaioli and Rainer (2006, 2007). *Source: Murdock (1967).*

Class Stratification: Ordered variable ranging from 0 to 4 quantifying "*the degree of class differentiation, excluding purely political and religious statuses*". A zero score indicates "*absence of significant class distinctions among freemen, ignoring variations in individual repute achieved through skill, valor, piety, or wisdom.*" A score of 1 indicates "*the presence of wealth distinctions, based on possession or distribution of property, which however have not crystallized into distinct and hereditary social classes.*" A score of 2 indicates "*elite stratification in which an elite class derives its superior status from control over scarce resources, particularly land, and is thereby differentiated from a propertyless proletariat or serf class*". A score of 3 indicates a "*dual stratification into a hereditary aristocracy and a lower class of ordinary commoners or freemen, where traditionally ascribed noble status is at least as decisive as control over scarce resources.* A score of 4 indicates "*complex stratification into social classes correlated in large measure with extensive differentiation of occupational statuses.*" *Source: Murdock (1967); variable code in the Ethnolinguistic Atlas v67.*

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

References

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

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

Acemoglu, D., and M. Dell (2008): "Productivity Differences Between and Within COuntries," *American Economic Journal: Macroeconomics*, 2(1), 169–188.

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," *Handbook of Economic Growth*, 91, 386–472.

Acemoglu, D., D. Ticchi, and A. Vindigni (2009): "Emergence and Persistence of Inefficient States," Journal of the European Economic Association, forthcoming.

Alcalá, F., and A. Ciccone (2004): "Trade and Productivity," *Quarterly Journal of Economics*, 119(2), 612–645.

Alesina, A., W. Easterly, and J. Matuszeski (2010): "Artificial States," Journal of the European Economic Association, forthcoming.

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

Ashraf, Q., and O. Galor (2008): "Human Genetic Diversity and Comparative Economic Development," Working Paper, CEPR No. 6824.

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

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

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

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

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

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

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

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

Chandler, T. (1987): *Four Thousand Years of Urban Growth: An Historical Census*. Edwin Mellon Press, Lewiston, NY.

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

Ciccone, A., and M. Jarocinski (2010): "Determinants of Economic Growth: Will Data Tell?," American Economic Journal Macroeconomics, forthcoming.

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

Deaton, A., and A. Heston (2010): "Understanding PPPs and PPP-based National Accounts," American Economic Journal: Macroeconomics, forthcoming.

Dell, M. (2009): "The Persistent Effects of Peru's Mining Mita," Econometrica, forthcoming.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Gopinath, G., P.-O. Gourinchas, C.-T. Hsieh, and N. Li (2010): "International Prices, Costs and Mark-up Differences," American Economic Review, forthcoming.

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

Hansson, G., and O. Olsson (2010): "Country Size and the Rule of Law: Resuscitating Montesquieu," European Economic Review, forthcoming.

Henderson, V. J., A. Storeygard, and D. N. Weil (2009): "Measuring Economic Growth from Outer Space," NBER Working Paper No. 15199.

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

Holmes, T. J. (1998): "The Effect of State Policies on the location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, 106(4), 667–705.

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

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

Johnson, S., W. Larson, C. Papageorgiou, and A. Subramanian (2009): "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates," Working Paper NBER No. 15455.

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

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

La-Porta, R., F. L. de Silanes, A. Shleifer, and R. Vishny (1998): "Law and Finance," *Journal of Political Economy*, 106(6), 1113–1155.

——— (1999): "The Quality of Government," *Journal of Law Economics and Organization*, 15(1), 222–279.

——— (2010): "Human Capital and Regional Development," mineo, Harvard University.

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

McCallum, J. (1995): "National Borders Matter: Canada-U.S. Regional Trade Patterns," *American Economic Review*, 85, 615–623.

Michalopoulos, S. (2008): "The Origins of Ethnic Diversity," Working Paper, mimeo Tufts University.

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

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

MILLER, D., C. A.CAMERON, AND J. GELBACH (2006): "Robust Inference with Multi-Way Clustering," NBER Technical Working Paper No. T0327.

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

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

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

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

NARITOMI, J., R. R. SOARES, AND J. J. ASSUNÇÃO (2009): "Institutional Development and Colonial Heritage within Brazil," Working Paper, IZA DP No. 4276.

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

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

NUNN, N., AND D. PUGA (2010): "Ruggedness: The Blessing of Bad Geography in Africa," Review of Economics and Statistics, forthcoming.

NUNN, N., AND L. WANTCHEKON (2009): "The Slave Trade and the Origins of Mistrust in Africa," Working Paper, NBER No. 14783.

OLSON, M. (1987): "Diseconomies of Scale and Development," *Cato Journal*, 7(1), 77–98.

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.

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

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

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

——— (1970): "Tribal innovators: Tswana chiefs and social change, 1795-1940," *London School of Economics Monographs on Social Anthropology*, (43).

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

Silva, J. S., S. Tenreyro, and F. Windmeijer (2010): "Is it Different for Zeros? Discriminating Between Models for Non-Negative Data with Many Zeros," CeMMAP working papers CWP20/10.

Spolaore, E., and R. Wacziarg (2005): "Borders and Growth," *Journal Of Economic Growth*, 10(4).

——— (2009): "The Diffusion of Development," *Quarterly Journal of Economics*, 124(2).

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

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

Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Young, A. (2009): "The African Growth Miracle," mimeo, London School of Economics.

Table 1 - Panel A: Summary Statistics

variable	Obs.	mean	st. dev.	p25	median	p75	min	max
Light Density	1218	0.198	0.521	0.000	0.018	0.126	0.000	4.653
Ln (1 + Light Density)	1218	0.131	0.271	0.000	0.017	0.119	0.000	1.732
Ln(Light Density)	862	-2.956	2.056	-4.255	-2.930	-1.542	-10.431	1.537
Rule of Law	1218	-0.851	0.583	-1.266	-0.888	-0.464	-1.912	0.708
Control of Corruption	1218	-0.776	0.494	-1.048	-0.873	-0.468	-1.590	0.722
Ln(Population Density)	1218	2.732	1.789	1.831	3.004	3.957	-6.673	6.219
Distance from the Capital City	1218	0.629	0.623	0.265	0.426	0.715	0.010	3.174
Distance from the Sea	1218	0.654	0.442	0.264	0.610	0.973	0.005	1.805
Ln(Area)	1218	1.946	1.721	0.940	2.090	3.122	-4.032	6.202
Ln(Water Area)	1218	0.262	0.424	0.003	0.091	0.323	0.000	3.119
Mean Elevation	1218	0.627	0.438	0.298	0.493	0.958	0.000	2.181
Land Suitability For Agriculture	1218	0.409	0.239	0.254	0.424	0.572	0.001	0.979
Malaria Stability Index	1218	0.732	0.334	0.525	0.894	1.000	0.000	1.000
Diamond Mine Indicator	1218	0.091	0.288	0.000	0.000	0.000	0.000	1.000
Major City in 1400 AD Inidcator	1218	0.020	0.139	0.000	0.000	0.000	0.000	1.000
Oil Deposit Indicator	1218	0.072	0.259	0.000	0.000	0.000	0.000	1.000
Split Indicator	1218	0.424	0.494	0.000	0.000	1.000	0.000	1.000
Jurisdictional Hierarchy	669	1.196	0.950	0.000	1.000	2.000	0.000	3.000
Political Centralization	669	0.356	0.479	0.000	0.000	1.000	0.000	1.000
Class Stratification	591	1.359	1.417	0.000	1.000	3.000	0.000	4.000

The Table reports descriptive statistics for all variables employed in the empirical analysis. The Data Appendix gives detailed variable definitions and data sources.

Table 1 - Panel B: Regression Discontinuity Design

	Partitioned Ethnic Group Areas		
	High Rule of Law	Low Rule of Law	Standard Error
Malaria Stability	0.747	0.747	[0.009]
Elevation	640.606	624.632	[12.316]
Land Area	19.784	19.811	[2.198]
Water Area	0.487	0.420	[0.058]
Suitability for Agriculture	0.408	0.400	[0.008]
Distance from the Capital City	0.479	0.744	[0.082]***
Distance from the Sea	0.664	0.657	[0.008]
Diamond Mine Indicator	0.107	0.104	[0.023]
Oil/Petroleum Deposit Indicator	0.062	0.032	[0.019]
Population Density in 1960	15.159	14.163	[1.723]
Major City in 1400 AD Indicator	0.016	0.019	[0.007]
Observations	309	309	

The unit of observation is adjacent partitions of ethnic groups divided by national boundaries. Standard errors for the difference in means in brackets are double clustered at the common border and ethnic-group dimension. Standard errors for the difference in means in parentheses are clustered at the country each partition belongs to. The Data Appendix gives detailed variable definitions and data sources. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 2 - Preliminary Evidence:
National Contemporary Institutions, Precolonial Local Ethnic-Specific Institutions, and Regional Development

	National Contemporary Institutions				Ethnicity-specific Pre-Colonial Institutions				National and Ethnic Institutions			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: OLS												
Rule of Law	0.1105**	0.1076**	0.1316***	0.1046**					0.1240**	0.1224**	0.1431**	0.1188**
	(0.0470)	(0.0509)	(0.0499)	(0.0418)					(0.0518)	(0.0618)	(0.0582)	(0.0523)
Jurisdictional Hierarchy					0.0675***	0.0601***	0.0558**	0.0457**	0.0579***	0.0529***	0.0451**	0.0406**
					(0.0257)	(0.0219)	(0.0231)	(0.0179)	(0.0213)	(0.0190)	(0.0188)	(0.0163)
adjusted R-squared	0.057	0.185	0.238	0.295	0.053	0.200	0.245	0.311	0.112	0.246	0.301	0.348
Panel B: Poisson ML												
Rule of Law	0.8606***	0.8071***	0.9747***	0.7330**					0.9983***	0.7921**	0.8413***	0.7621***
	(0.2125)	(0.3462)	(0.3297)	(0.2918)					(0.2250)	(0.3745)	(0.2397)	(0.2527)
Jurisdictional Hierarchy					0.5386***	0.4664***	0.3950***	0.3039***	0.4414***	0.3779***	0.3017***	0.2470**
					(0.1424)	(0.1338)	(0.1367)	(0.1109)	(0.1262)	(0.1203)	(0.1168)	(0.1041)
Log Likelihood	-625.91	-541.34	-499.52	-458.61	-365.62	-310.92	-287.06	-271.47	-343.00	-297.38	-274.36	-261.88
RD Polynomial	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Geographic Controls	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Observations	1218	1218	1218	1218	672	672	672	672	669	669	669	669

Panel A reports OLS estimates and Panel B reports Poisson ML estimates associating regional development with contemporary national institutions (in columns (1)-(4), (9)-(12)) and pre-colonial ethnic institutions (in columns (5)-(12)). The dependent variable is light density at night from satellite at the ethnicity-country level. In the OLS specifications satellite light density at night is expressed in logs, while in the Poisson estimates satellite light density at night is expressed in levels. The proxy for contemporary national institutions is World Bank's Governance Matters rule of law index. The proxy for pre-colonial ethnic institutions is Murdock's (1967) index of jurisdictional hierarchy beyond the local community level before colonization. The regression discontinuity (RD) polynomial in columns (2)-(4), (6)-(8), and (10)-(12) is a cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast. The specifications in columns (3), (4), (7), (8), (11), and (12) also include the log population density. The set of geographic controls included in columns (4), (8) and (12), include log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: OLS												
Rule of Law	0.0939*	0.0241	0.1132**	0.0251	0.0932**	0.0246						
	(0.0497)	(0.0517)	(0.0502)	(0.0469)	(0.0457)	(0.0497)						
Control of Corruption							0.0985	0.0384	0.1233**	0.0445	0.0990*	0.0431
							(0.0635)	(0.0703)	(0.0629)	(0.0644)	(0.0585)	(0.0741)
adjusted R-squared	0.167	0.684	0.21	0.721	0.249	0.725	0.161	0.685	0.204	0.722	0.245	0.726
within R-squared	__	0.0566	__	0.1661	__	0.179	__	0.0586	__	0.1695	__	0.1819
Panel B: Poisson ML												
Rule of Law	0.8780**	-0.0187	1.0575**	0.1850	0.8362*	0.066						
	(0.4309)	(0.3164)	(0.5232)	(0.1903)	(0.4449)	(0.1941)						
Control of Corruption							0.9277**	0.0864	1.2114**	0.2282	0.9339**	0.0888
							(0.4702)	(0.4239)	(0.5063)	(0.1947)	(0.4389)	(0.2137)
Log Likelihood	-625.91	-542.52	-499.52	-458.61	-365.62	-328.10	-192.58	-124.80	-177.38	-113.74	-164.08	-113.15
Ethnicity Fixed-Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Geographic Controls	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Observations	503	503	503	503	503	503	503	503	503	503	503	503

Panel A reports OLS estimates and Panel B reports Poisson ML estimates associating regional development with contemporary national institutions as proxied by a rule of law index (in columns (1)-(6)) and a control of corruption index (in columns (7)-(12)) in areas of partitioned ethnicities. The dependent variable is light density at night from satellite at the ethnicity-country level. In the OLS specifications satellite light density at night is expressed in logs, while in the Poisson estimates satellite light density at night is expressed in levels. Odd-numbered specifications report cross-sectional estimates. Even-numbered specifications report within-country estimates. The specifications in even-numbered columns include a set of ethnicity fixed-effects (constants not reported). In all specifications we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast. The specifications in columns (3)-(6) and (9)-(12) also include the log population density. The set of geographic controls included in columns (4), (6), (11), and (12) include log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 4: Pre-colonial Ethnic Institutions and Regional Development within African Countries

	Jurisdictional Hierarchy				Political Centralization				Class Stratification			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: OLS												
Ethnic Institutions	0.0538***	0.0434***	0.0419***	0.0352***	0.0878***	0.0813***	0.0707***	0.0658***	0.0251**	0.0210**	0.0308***	0.0250**
	(0.0182)	(0.0166)	(0.0143)	(0.0135)	(0.0310)	(0.0287)	(0.0233)	(0.0224)	(0.0099)	(0.0088)	(0.0109)	(0.0107)
adjusted R-squared	0.418	0.557	0.519	0.607	0.410	0.555	0.515	0.606	0.550	0.645	0.436	0.578
within R-squared	0.044	0.271	0.209	0.354	0.030	0.269	0.202	0.353	0.228	0.391	0.033	0.276
Panel B: Poisson ML												
Ethnic Institutions	0.4767***	0.2670***	0.2140**	0.1545*	0.7546***	0.5228**	0.4530**	0.3705**	0.1706***	0.1557***	0.2613***	0.2103***
	(0.0999)	(0.1029)	(0.0972)	(0.0804)	(0.2157)	(0.2028)	(0.1776)	(0.1444)	(0.0499)	(0.0270)	(0.0655)	(0.0485)
Log Likelihood	-271.37	-233.46	-229.12	-213.82	-275.44	-233.39	-228.80	-213.29	-246.47	-206.73	-201.44	-185.92
Country Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Observations	672	672	672	672	672	672	672	672	593	593	593	593

Panel A reports country fixed-effects OLS estimates and Panel B reports country fixed-effects Poisson ML estimates associating regional development with pre-colonial ethnic institutions. The dependent variable is light density at night from satellite at the ethnicity-country level. In the OLS specifications satellite light density at night is expressed in logs, while in the Poisson estimates satellite light density at night is expressed in levels. In columns (1)-(4) we measure pre-colonial ethnic institutions using Murdock's (1967) index of jurisdictional hierarchy beyond the local level. In columns (5)-(8) we use the binary (0-1) Gennaioli and Rainer (2006, 2007) political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local level. In columns (9)-(12) we use Murdock's (1967) class stratification index. All specifications include a set of country fixed-effects (constants not reported). In even-numbered specifications we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast and a set of geographic controls. The set of geographic controls includes log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. The specifications in columns (3), (4), (7), (8), (11), and (12) also include the log population density. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 5: Contemporary National Institutions and Regional Population Density across and within Partitioned Ethnicities

	(1)	(2)	(3)	(4)	(5)	(6)
Rule of Law	-0.3244	0.0744	0.0069			
	(0.4245)	(0.2767)	(0.3107)			
Control of Corruption				-0.4227	0.0195	-0.0965
				(0.4928)	(0.3357)	(0.3722)
adjusted overall R-squared	0.012	0.801	0.829	0.014	0.801	0.830
within R-squared	__	0.006	0.029	__	0.001	0.026
Ethnicity Fixed-Effects	No	Yes	Yes	No	Yes	Yes
RD Polynomial	No	No	Yes	No	No	Yes
Geographic Controls	No	No	Yes	No	No	Yes
observations	503	503	503	503	503	503

The Table reports cross-sectional and within-ethnicity OLS estimates associating regional population density with contemporary national institutions, as proxied by a rule of law index (in columns (1)-(3)) and a control of corruption index (in columns (4)-(6)) in areas of partitioned ethnicities. The dependent variable is log population density at the ethnicity-country level. Odd-numbered specifications report cross-sectional estimates. Even-numbered specifications report within-country estimates. The specifications in even-numbered columns include a set of ethnicity fixed-effects (constants not reported). In the specifications in columns (3) and (6) we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast and a set of geographic control variables. The set of geographic controls includes log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family of each tribe dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 6 - Sensitivity Analysis:
National Contemporary Institutions, Precolonial Local Ethnic-Specific Institutions, and Regional Development

	Excluding North Africa				Large Differences in Rule of Law		Additional Controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: OLS										
Rule of Law	0.1043**	0.0317			0.1694**	0.0617	0.1398***	0.031		
	(0.0485)	(0.0518)			(0.0708)	(0.0728)	(0.0526)	(0.0503)		
Political Centralization			0.0797**	0.0605***					0.0902***	0.0640***
			(0.0387)	(0.0218)					(0.0348)	(0.0216)
adjusted R-squared	0.265	0.710	0.284	0.565	0.431	0.877	0.37	0.73	0.336	0.606
Panel B: Poisson ML										
Rule of Law	1.0169**	0.1557			1.1055***	0.3526	0.8665***	0.3028		
	(0.4807)	(0.1950)			(0.3737)	(0.2738)	(0.2637)	(0.2905)		
Political Centralization			0.4214*	0.3045***					0.5502**	0.3915***
			(0.2470)	(0.0873)					(0.2539)	(0.1498)
Log Likelihood	-139.04	-99.83	-174.28	-151.61	-43.48	-34.98	-453.47	-112.91	-271.40	-212.81
Ethnicity Fixed-Effects	No	Yes	No	No	No	Yes	No	Yes	No	No
Country Fixed-Effects	No	No	No	Yes	No	No	No	No	No	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	480	480	636	636	160	160	686	503	669	669

Table 6 Notes

Panel A reports OLS estimates and Panel B reports Poisson ML estimates associating regional development with contemporary national institutions (in columns (1), (2), (5)-(8)) and pre-colonial ethnic institutions (in columns (3), (4), (11), 12)). The dependent variable is light density at night from satellite at the ethnicity-country level. In the OLS specifications satellite light density at night is expressed in logs, while in the Poisson estimates satellite light density at night is expressed in levels. The proxy for contemporary national institutions is World Bank's Governance Matters rule of law index. The proxy for pre-colonial ethnic institutions is the Gennaioli and Rainer (2006, 2007) binary (0-1) political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local level. The specifications in columns (2), (6), and (8) include a set of ethnicity fixed-effects (constants not reported). The specifications in columns (4) and (10) include a set of country fixed-effects (constants not reported). In all specifications we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast, a set of geographic controls, and log population density. The set of geographic controls includes log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. Columns (1)-(4) report estimates only in Sub-Saharan countries (excluding North Africa). Columns (5) and (6) focus on ethnicities partitioned in 2 countries and report cross-sectional and within-ethnicity results across pairs of countries with large differences in the rule of law index (larger than the median difference). The specifications in column (7)-(10) also include indicator variables on: (i) whether the ethnicity is partitioned by a national border or not, (ii) whether a major city in 1400 is in the ethnicity's historical homeland, (iii) whether a diamond mine is present or not, and (iv) whether an oil/petroleum field is present or not. These specifications also include an index of precolonial settlement patterns (ranging from nomadic ethnicities to groups residing in complex settlements). The Data Appendix gives detail variable definitions and data sources. In the specifications with country fixed-effects we also report the within country R-squared. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 7 - Heterogeneous Effects:
National Institutions and Regional Development Across and Within Partitioned Ethnicities

	Distance to the Capital				Border Type			
	Close	Far	Close	Far	Natural	Artificial	Natural	Artificial
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: OLS				
Rule of Law	0.2055**	0.0215**	-0.0127	0.0075	0.1998**	0.0636***	0.0993	-0.0245
	(0.0845)	(0.0105)	(0.0903)	(0.0237)	(0.0824)	(0.0225)	(0.1243)	(0.0525)
adjusted R-squared	0.505	0.402	0.906	0.825	0.428	0.342	0.877	0.739
within R-squared	__	__	0.307	0.455	__	__	0.281	0.380
				Panel B: Poisson ML				
Rule of Law	1.2072***	0.5277***	0.3596	0.0911	1.2518***	1.1261***	0.2341	-0.4698
	(0.4510)	(0.1400)	(0.2570)	(0.3960)	(0.5850)	(0.2750)	(0.4180)	(0.4310)
Log Likelihood	-32.62	-14.35	-26.19	-11.67	-64.88	-19.08	-48.11	-16.68
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	No	No	Yes	Yes	No	No	Yes	Yes
Observations	92	96	92	96	180	126	180	126

Panel A reports OLS estimates and Panel B reports Poisson ML estimates associating regional development with contemporary national institutions as proxied by a rule of law in areas of partitioned ethnicities. The dependent variable is light density at night from satellite at the ethnicity-country level. In the OLS specifications satellite light density at night is expressed in logs, while in the Poisson estimates satellite light density at night is expressed in levels. Odd-numbered specifications report cross-sectional estimates. Even-numbered specifications report within ethnicity estimates (ethnicity constants not reported). In all specifications we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast, a set of geographic controls, and log population density. The set of geographic controls includes log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. All specifications focus on ethnicities partitioned in 2 countries (i.e. partitioned ethnic groups in 3 or more countries are excluded). In columns (1) and (3) we restrict estimation in areas of partitioned ethnicities that are both close to the capital city (less than the median distance). In columns (2) and (4) we restrict estimation in areas of partitioned ethnicities that are both far from the capital city (more than the median distance). In columns (5) and (7) we restrict estimation in areas of partitioned ethnicities that are more likely to be natural, i.e. separated by significant geographic barriers. In columns (6) and (8) we restrict estimation in areas of partitioned ethnicities that are more likely to be artificial (are not separated by significant geographic barriers). The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 8 : Pre-colonial Ethnic Institutions and Regional Development within African Countries
Measurement Error and Intensive Margin Analysis

	All Observations			Intensive Margin		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: 2SLS			
Jurisdictional Hierarchy	0.0554***			0.4306***		
	(0.0198)			(0.1586)		
Political Centralization		0.1167***			0.9013***	
		(0.0409)			(0.3204)	
Class Stratification			0.0374***			0.1586*
			(0.0132)			(0.0816)
			Panel B: 1st Stage			
Instrument	Class Strat	Class Strat	Juris Hier	Class Strat	Class Strat	Juris Hier
coefficient	0.3824	0.1817	0.8766	0.3826	0.1287	0.8358
standard error	(0.0408)	(0.0215)	(0.0758)	(0.0400)	(0.0209)	(0.0917)
First-stage *F* -score	87.73 [0.00]	71.86 [0.00]	134.88 [0.00]	91.34 [0.00]	76.13 [0.00]	83.06 [0.00]
			Panel C: Corresponding OLS			
coefficient	0.0328***	0.0663***	0.0212***	0.1326*	0.2766**	0.1647***
standard error	(0.0125)	(0.0209)	(0.0078)	(0.0730)	(0.1321)	(0.0589)
Country Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	560	560	560	424	424	424

The Table reports country fixed-effects 2SLS estimates associating regional development with pre-colonial ethnic institutions. The dependent variable is light density at night from satellite at the ethnicity-country level. In columns (4)-(6) we exclude unlit country-ethnic areas and focus on the "intensive margin" of luminosity. Panel A gives the estimates of the second stage. Panel B gives the estimates of the first-stage. Panel C gives the corresponding OLS estimates. In columns (1) and (4) we instrument jurisdictional hierarchy beyond the local community level with class stratification. In columns (2) and (5) we instrument the binary (0-1) Gennaioli and Rainer (2006, 2007) political centralization index with class stratification. In columns (3) and (6) we instrument class stratification with the jurisdictional hierarchy beyond the local community level index. All specifications include a set of country fixed-effects (constants not reported). In all specifications we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast, a set of geographic controls, and log population density. The set of geographic controls includes log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. The Data Appendix gives detailed variable definitions and data sources. The Table also reports the first-stage F-score of the excluded instrument. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family of each tribe dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Table 9: The Rule of the Divided
Pre-colonial Ethnic Institutions and Regional Development in Partitioned Ethnicities

	Jurisdictional Hierarchy		Political Centralization		Class Stratification	
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: OLS			
Ethnic Institutions	0.0583***	0.0546***	0.0905**	0.0954***	0.0232**	0.0162*
	(0.0186)	(0.0156)	(0.0395)	(0.0333)	(0.0116)	(0.0093)
adjusted R-squared	0.560	0.665	0.551	0.663	0.604	0.711
within R-squared	0.064	0.403	0.043	0.399	0.033	0.426
			Panel B: Poisson ML			
Ethnic Institutions	0.4067***	0.2621**	0.9011***	0.7158***	0.1845***	0.1263**
	(0.1163)	(0.1278)	(0.2296)	(0.2455)	(0.0695)	(0.0560)
Log Likelihood	-89.76	-83.99	-89.15	-83.45	-80.33	-74.22
Country Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
RD Polynomial	No	Yes	No	Yes	No	Yes
Population Density	No	Yes	No	Yes	No	Yes
Geographic Controls	No	Yes	No	Yes	No	Yes
Additional Controls	No	Yes	No	Yes	No	Yes
Observations	309	309	309	309	261	261

Panel A reports country fixed-effects OLS estimates and Panel B reports country fixed-effects Poisson ML estimates associating regional development with pre-colonial ethnic institutions in areas of partitioned ethnicities. The dependent variable is light density at night from satellite at the ethnicity-country level. In the OLS specifications satellite light density at night is expressed in logs, while in the Poisson estimates satellite light density at night is expressed in levels. In columns (1) and (2) we measure pre-colonial ethnic institutions using Murdock's (1967) index of jurisdictional hierarchy beyond the local level. In columns (3) and (4) we use the binary (0-1) Gennaioli and Rainer (2006, 2007) political centralization index that is based on Murdock's (1967) jurisdictional hierarchy beyond the local level. In columns (5) and (5) we use Murdock's (1967) class stratification index. All specifications include a set of country fixed-effects (constants not reported). In even-numbered specifications we include a regression discontinuity (RD) cubic polynomial in distance of the centroid of each ethnicity-country area from the capital city of each country and distance from the closest sea coast, a set of geographic controls, and additional control variables. The set of geographic controls includes log surface area, log area under water (lakes, rivers, and other streams), land suitability for agriculture, elevation, and a malaria stability index. The set of additional control variables includes indicators on: (i) whether the ethnicity is partitioned by a national border or not, (ii) on whether a major city in 1400 is in the ethnicity's historical homeland, (iii) whether a diamond mine is present or not, and (iv) whether an oil/petroleum field is present or not. These specifications also include an index of precolonial settlement patterns (ranging from nomadic ethnicities to groups residing in complex settlements). The specifications in even-numbered columns also include the log population density. The Data Appendix gives detailed variable definitions and data sources. Standard errors reported in parentheses are adjusted for double clustering at the country-dimension and the ethno-linguistic family of each tribe dimension. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level respectively.

Appendix Table 1 - Partitioned Ethnicities and Countries they Belong to

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

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

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

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

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

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

Appendix Table 1 reports the name of partitioned ethnic groups (as coded by Murdock (1959)) and the percentage of the historical homeland of the split ethnic groups that fall into more than one country. Section 2 in the paper gives details on our approach in identifying partitioned ethnicities.