

DISCUSSION PAPER SERIES

No. 10258

**FURTHER EVIDENCE ON THE LINK
BETWEEN PRE-COLONIAL POLITICAL
CENTRALIZATION AND COMPARATIVE
ECONOMIC DEVELOPMENT IN AFRICA**

Stelios Michalopoulos and Elias Papaioannou

DEVELOPMENT ECONOMICS



Centre for Economic Policy Research

FURTHER EVIDENCE ON THE LINK BETWEEN PRE-COLONIAL POLITICAL CENTRALIZATION AND COMPARATIVE ECONOMIC DEVELOPMENT IN AFRICA

Stelios Michalopoulos and Elias Papaioannou

Discussion Paper No. 10258

November 2014

Submitted 23 October 2014

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **DEVELOPMENT ECONOMICS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Stelios Michalopoulos and Elias Papaioannou

FURTHER EVIDENCE ON THE LINK BETWEEN PRE-COLONIAL POLITICAL CENTRALIZATION AND COMPARATIVE ECONOMIC DEVELOPMENT IN AFRICA[†]

Abstract

We examine the link between pre-colonial statehood and contemporary regional African development, as reflected in satellite images on light density at night. We employ a variety of historical maps to capture the former. Our within-country analysis reveals a strong positive correlation between pre-colonial political centralization and contemporary development (and urbanization). If anything, the association strengthens when we account for measurement error on the historical maps of pre-colonial political organization.

JEL Classification: O10, O40 and O43

Keywords: africa, development, ethnicity, institutions and state capacity

Stelios Michalopoulos smichalo@brown.edu
Brown University

Elias Papaioannou eliaspapaioannou@london.edu
London Business School and CEPR

[†] Stelios Michalopoulos, Brown University, Department of Economics, 64 Waterman St., Robinson Hall, Providence RI, 02912, USA. Email: smichalo@brown.edu. Web: <https://sites.google.com/site/stelioecon/>. Elias Papaioannou, London Business School, Regent's Park, 1-26 Sussex Place, London NW1 4SA, United Kingdom. Email: eliaspapaioannou@london.edu. Web: <https://sites.google.com/site/papaioannouelias/>. All errors are our sole responsibility.

1 Introduction

1.1 Previous Research

In Michalopoulos and Papaioannou (2013) we examined empirically the role of deeply-rooted pre-colonial ethnic institutions on comparative regional development within African countries. We combined information on the spatial distribution of ethnicities before colonization (as portrayed in George Peter Murdock’s [1959 map]) and anthropological sources describing political organization at the time of colonization across ethnic groups (from Murdock (1967), see Figure 1a),¹ with regional variation in contemporary development, as captured by satellite images of light density at night. This approach follows Henderson, Storeygard, and Weil (2012) and subsequent works (e.g., Chen and Nordhaus (2011), Pinkovskiy (2013) and Hodler and Raschky (2014a)) who proposed proxying economic development (GDP) with luminosity (per area), especially for war-prone countries and for regional analyses where output data are scarce.²

The analysis establishes that there is a strong long-run association between pre-colonial political centralization and regional development across ethnic homelands. Our study, thus, offers supportive evidence to the arguments in the African historiography, dating at least to the important work of Fortes and Evans-Pritchard (1940) on the salience of deeply-rooted political organizational structures (see, also, Acemoglu and Robinson (2012) and Diamond (1997) for eloquent narratives). Our findings also complement cross-country works showing a positive correlation between early statehood and subsequent development in a global sample (Bockstette, Chanda, and Putterman (2002) and Putterman and Weil (2010)) as well as across African states (Gennaioli and Rainer (2006, 2007)). While this correlation does not necessarily imply a casual relationship (from pre-colonial) political centralization to regional development, it appears quite robust.³ While hard-to-account-for unobserved features -related to geography, culture, or early development- may explain this long-run relationship, it is intriguing that ethnic-specific legacies dating back to mid-late 19th century still matter nowadays.

The link between pre-colonial centralization and regional development has been verified by a number of subsequent studies. For example, using two-stage-least-squares models, Fenske (2014) and Alsan (2014) show that the geography and ecology-predicted component of pre-colonial political centralization is also a robust correlate of regional development. Moreover, a

¹This proxy of political centralization describes the number of political jurisdictions above the village level. A 0 indicates stateless societies. 1 indicates small chiefdoms. 2 designates paramount (large) chiefdoms; and 3 and 4 indicate groups that were part of large states.

²See also Hodler and Raschky (2014b), Hodler and Raschky (2014a), and Storeygard (2014), among others

³In particular, the within-country association is not driven by differences in local geographic features or by observable ethnic-specific cultural and economic variables. Moreover, Michalopoulos and Papaioannou (2013) also show that luminosity correlates strongly with early political centralization, even when they condition on dozens of alternative pre-colonial ethnic characteristics, such as occupational specialization, economic organization, the presence of polygyny, slavery, and proxies of early development (using data from Murdock (1967)).

similar link seems to be also present outside Africa. Focusing on North America, Dippel (2014) finds a strong positive correlation between the degree of pre-colonial political centralization of indigenous tribes and contemporary development across indigenous Native American Reservations. Likewise, Chiovelli (2014) uncovers a robust link between pre-colonial statehood and sub-national development (as captured by luminosity) within Latin American countries.

The increasing number of studies that use Murdock’s classification of pre-colonial centralization as a proxy for the complexity of pre-colonial institutional complexity is subject to the critique that this classification may contain sizeable and (perhaps non-classical) measurement error. This note alleviates such concerns by providing an alternative measure of pre-colonial centralization in Africa and showing that the empirical regularities established by Michalopoulos and Papaioannou (2013) remain intact.

1.2 Measuring Pre-colonial Political Organization

Murdock’s data and map clearly contain noise. Measurement error in Murdock’s (1967, 1959) data is present for (at least) the following reasons:

1. Murdock (1959, 1967) relied mostly on secondary historical resources that (unfortunately) for Africa are quite scarce.
2. There is always some subjectivity involved when classifying historical political regimes. (Regime classification is tricky even with contemporary data).⁴
3. Murdock (1959) assigned each part of Africa to a single dominant group, while – to some extent at least- ethnicities overlap; and naturally population mixing is higher closer to ethnic boundaries and in the capitals.
4. The maps of ethnic boundaries are (rough) approximations.⁵
5. Another problem with Murdock’s data is their limited availability. Murdock’s Ethnolinguistic Atlas (1967) reports pre-colonial ethnic-specific economic, institutional, and cultural traits only for a subset of the ethnic groups reported in .
6. Moreover, scholars in African historiography maintain that the impact of statehood in pre-colonial Africa was evident in the centres of pre-colonial states, kingdoms, and empires

⁴See, for example, Acemoglu, Naidu, Restrepo, and Robinson (2014), Papaioannou and Siourounis (2008), and Munck and Verkuilen (2002), among others.

⁵Michalopoulos and Papaioannou (2013) (page 143) write “*while national borders are accurately delineated, drawing error in Murdock’s map on the exact location of ethnic boundaries is likely to be nontrivial.*” Moreover, “*since Murdock’s map, originally printed in the end of his book on African ethnicities, is available at a small scale, its digitization magnifies any noise inherent to the initial border drawing.*”

decaying for regions in the periphery (see Herbst (2000) for an eloquent overview, and Bandyopadhyay and Green (2012) for recent evidence from Uganda). Thus, Murdock’s data are likely more accurate for the historical cores of the respective ethnic homelands.

Linear regression analysis produces attenuated (biased towards) zero estimates, when the explanatory variable (pre-colonial political centralization) contains error that takes the "classical" form, i.e., it is orthogonal to the outcome variable and the key controls, see Wooldridge (2002). However, it is hard knowing *ex ante* whether error-in-variables takes the classical form, as for example, Murdock may have systematically misclassified ethnic groups with strong polities as compared to acephalous ones.

Paper and Structure To account for these shortcomings related to error-in-variables, in this paper we use recently assembled data on pre-colonial statehood from Besley and Reynal-Querol (2014) (that in turn are based on a plethora of historical accords, maps, and archives) and re-examine the link between pre-colonial statehood and contemporary development across ethnic homelands. In Section 2 we present the data and detail the empirical specification. Section 3 gives the baseline OLS results. In Section 4 we present both two-stage-least-squares and OLS estimates that combine Murdock’s data with the newly assembled data in an effort to account for error-in-variables.

2 Data and Empirical Specification

2.1 Data

To capture pre-colonial political centralization, we employ recently compiled data by Timothy Besley and Marta Reynal-Querol (The Legacy of Historical Conflict. Evidence from Africa. *American Political Science Review*, 2014) that in turn are based on historical maps put together by Cioffi-Revilla (1996), O’Brien (1999) and Brecke (1999). Besley and Reynal-Querol (2014) provide georeferenced information on the reach of large kingdoms and empires just before the colonial era in Africa (Figure 1*b*).

Using this map we construct three alternative measures of pre-colonial statehood. First, we define a binary variable that takes on the value one when the centroid of a country-ethnic region (in Murdock’s Atlas of 1959) falls within the boundaries of a pre-colonial empire and zero otherwise. We label this measure “*pre-colonial political centralization 1*”. Overall 302 of a total of 1,218 country-ethnic homelands in Murdock’s map (24.8%) have been part of a large pre-colonial state, while the remaining 916 ethnic regions (75.2%) were not being part of a pre-colonial state. Second, we construct an alternative indicator of political centralization

that takes on the value one if *any* part of the ethnic homeland has been part of a pre-colonial state. This results in a sample of 444 (36.5%) ethnic homelands classified as being (at least partially) part of a pre-colonial kingdom and 774 (63.5%) country-ethnic observations, whose homelands have not been captured by one of the large pre-colonial African states. We label this measure as “*pre-colonial political centralization 2*”. Third, we estimated the distance of each country-ethnic region to the boundaries of the closest pre-colonial kingdom-empire.

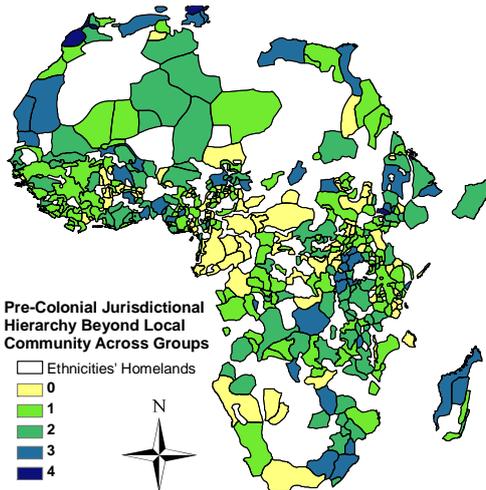


Figure 1a: Murdock's data

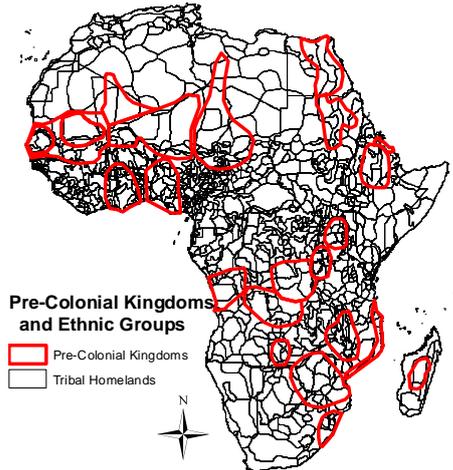


Figure 1b: O' Brecke data

2.2 Empirical Specification

We estimate OLS specifications associating regional development ($y_{i,c}$) in the ancestral homeland of ethnicity i in country c with these proxies of pre-colonial political centralization (IQL_i).

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

As dependent variable we use various transformations of (log) luminosity. In most specifications we control for the log of population density ($PD_{i,c}$) at the country-ethnicity level, so as to capture the role of pre-colonial institutions on development conditional on urbanization. Since African development may still be characterized by Malthusian dynamics, we also estimate models with log population density as the dependent variable (see also Michalopoulos and Papaioannou (2013)).⁶ In all specifications we include a vector of geographical and location controls ($X'_{i,c}$); the "geography" set includes elevation, land suitability for agriculture, the log of land area, the log of the area under water, a malaria stability index, a diamond mine indicator and an oil field indicator, while the "location" controls include distance of the centroid of

⁶We drop country-ethnic areas with zero population in 2000 (though this plays absolutely no role).

each ethnicity–country area from the respective capital city, distance from the closest sea coast, and distance from the national border. In all specifications, we include country fixed effects (a_c), so to account for time-invariant country-level factors, related to the type of colonization, the identity of the colonial power, national politics and institutions.⁷ This specification and conditioning set follows Michalopoulos and Papaioannou (2013) and subsequent works (Hodler and Raschky (2014a), Alsan (2014), Fenske (2013), among others). The Data Appendix gives detailed variable definitions and data sources.

3 Baseline Results

Table 1 reports within-country OLS estimates using as dependent variable various transformations of luminosity and population density.⁸ In panel *A* and panel *B* we use the pre-colonial political centralization measures 1, and 2, respectively. In panel *C* we use the natural logarithm of (0.1 plus distance to the boundaries of pre-colonial states of each ethnic region within a country). In columns (1)-(2) the dependent variable is the log of mean luminosity in 2007 and 2008 adding 0.01; this is the benchmark outcome (development proxy) in Michalopoulos and Papaioannou (2013) and related works (e.g., Hodler and Raschky (2014a), Fenske (2014), and Alsan (2014)). In columns (3)-(4) we focus on the “intensive” margin of luminosity (using as a dependent variable the logarithm of mean luminosity in 2007 – 2008), effectively excluding from the estimation low-development ethnic regions.⁹ In columns (5)-(6) we use as dependent variable the log of luminosity per capita (i.e., log luminosity in 2007 – 2008 minus log population density in 2000) without controlling for log population density and with log population density on the RHS. In column (7) we estimate models with log population density in 2000 as the dependent variable.

Let us start our discussion with the baseline specifications in (1)-(2), where we add a small number (0.01) in the luminosity before taking the logarithm so as to use information from all populated ethnic areas. The coefficient in column (1) on the indicator of pre-colonial kingdoms and empires is positive and significant at standard confidence levels. This applies to both proxies of pre-colonial political centralization (panels *A* and *B*). The same applies when we regress log luminosity on the natural logarithm of distance to pre-colonial states in panel *C*; log (0.01 +

⁷For ethnicities that fall into more than one country each partition is assigned to the corresponding country *c*. For example, regional light density in the part of the Maasai in Tanzania is assigned to Tanzania, while the adjacent region of the Maasai in Kenya is assigned to Kenya. After intersecting Murdock’s ethnolinguistic map with the 2000 Digital Chart of the World we drop ethnic partitions of less than 100 km^2 , as such small partitions are most likely due to the lack of precision in the underlying mapping.

⁸In all specifications we employ the approach of Cameron, Gelbach, and Miller (2011) and cluster standard errors along the country and at the ethnic-family dimensions. (Murdock assigns the 834 groups into 96 ethnolinguistic families).

⁹This follows the specifications reported in Table 3, Panel *B* in Michalopoulos and Papaioannou (2013).

distance to the centroid of a nearest pre-colonial kingdom) enters with a significantly negative elasticity implying higher levels of development in areas within and close to the boundaries of large pre-colonial states. Consistent with the estimates in Michalopoulos and Papaioannou (2013) (Tables 2 – 3), the coefficients on the proxies of pre-colonial political centralization drop once we control for log population density (in column (2)).¹⁰ Yet the coefficient on the binary political centralization index retains significance and is quite similar to the estimate in Michalopoulos and Papaioannou (2013) (Table III, column (9)). The magnitude suggests that luminosity is 35% ($\exp(0.298) - 1 = 0.347$) higher in ethnic homelands where politically centralized societies reside (e.g., Yoruba in Nigeria), as compared to stateless societies or small chiefdoms (e.g., the Sokoto or the Tiv in Nigeria).

The same pattern emerges when we restrict estimation to lit areas (in (3)-(4)). The pre-colonial political centralization indicators enter with highly significant coefficients. The estimate in column (4) implies that conditional on numerous observable features and contemporary population density, luminosity is roughly 50% higher in ethnic homelands that before the colonial era were parts of large kingdoms ($\exp(0.42) - 1 = 0.52$). Likewise, log distance to pre-colonial empires enters with a highly significant (at the 99% confidence level) negative coefficient implying that contemporary development is lower in ethnic areas far from pre-colonial states. Note again that the coefficient drops, once we control for log population density, a clearly relevant variable (the coefficient on log population density is 0.73 with a double-clustered standard error of 0.08).

In columns (5)-(6) we use as dependent variable the log of luminosity per capita (i.e., log luminosity in 2007 – 2008 minus log population density in 2000). In column (6) we control for log population density, while in (5) we do not. So, column (5) reports the “restricted” version of the model in column (6), the “restriction” being that the coefficient on log population density is 1. While this restriction is clearly rejected, we report these models as some scholars also do so for robustness (Hodler and Raschky (2014a)). The coefficient on the pre-colonial political centralization binary index is positive (0.33 and 0.36) and significant at the 98% – 99% confidence level. Likewise, log distance to the nearest pre-colonial kingdom enters with a significantly negative estimate. So, while one should not be estimating a “restricted” model, since the restriction of a unit-elasticity of luminosity-population density does not hold in the data, even in this case there is a strong relation between log luminosity per capita and being part of a pre-colonial kingdom.¹¹ The models in column (6) also include on the RHS the log

¹⁰Log population density enters with a significantly positive coefficient, around 0.5.

¹¹It is important to note that using luminosity per capita does not avoid the “bad control” problem (Angrist and Pischke (2008)) since employing the former as the dependent variable it is as if controlling by population density and constraining the coefficient to be equal to 1.

of population density in 2000 -that clearly belongs to the model as it enters with a highly significant coefficient, -0.27 (see also Hodler and Raschky (2014a), Table *II*, columns (7)-(8)). To the extent that some parts of Africa may be characterized by Malthusian dynamics and since luminosity is a far-from-perfect proxy of local output, in column (7) we use log population density as the dependent variable. Pre-colonial statehood is systematically linked to contemporary population density.

Sensitivity Analysis 1. Population In Appendix Table 1 we examine the robustness of these correlations dropping “small” -in terms of population- ethnic areas. In column (1) we exclude from the estimation country-ethnic areas with small population (bottom 1% of the population distribution in 2000). In (2) and (3) we drop observations (country-ethnic areas) in the bottom 5th and in the bottom 10th percentile of population. In (4) and (5) we report even more restrictive models, dropping a quarter or even half of the sample. In all permutations the coefficient on the pre-colonial political centralization is positive and statistically significant at standard confidence levels. Moreover, the estimate is quite stable across the various subsamples despite the fact that trimming the sample obviously restricts the variation in our dependent variable of interest. In the baseline specifications in column (2) where we use all observations and condition on log population density the coefficient on the pre-colonial political centralization index is $0.29 - 0.35$, very close to the estimate in the full sample (0.30).

Sensitivity Analysis 2. Outside Capital Cities Motivated by the findings of Michalopoulos and Papaioannou (2014a) that the impact of national institutions (e.g., rule of law) on regional development decays further from the capitals and the results in Michalopoulos and Papaioannou (2014b) showing that pre-colonial ethnic-specific political institutions and chiefs’ role is especially strong outside these urban hubs, in Appendix Table 2 we examine the association between log luminosity and pre-colonial political centralization excluding areas close to the capitals. In the latter, population mixing and the dominance of national institutions are likely to diminish the role of ethnic-specific traits. With all various transformations of luminosity, the coefficient on pre-colonial political centralization (or distance to pre-colonial kingdoms) is quantitatively larger when we exclude regions close to the capital (the coefficient estimate increases by roughly 25% depending on the specification).

4 Accounting for Measurement Error

All proxies and maps of pre-colonial centralization are bound to contain noise; thus the OLS estimates (in Table 1 and in Appendix Tables 1 – 2) are also likely attenuated, perhaps con-

siderably, yielding conservative inference.

4.1 2sls Estimates

In an effort to quantify the role of error-in-variables, we estimated within-country two-stage-least-squares (2SLS) specifications “instrumenting” the log distance to the centroid of pre-colonial states with the binary political centralization index of Gennaioli and Rainer (2006, 2007) and Michalopoulos and Papaioannou (2013), which is based on George Peter Murdock’s jurisdictional hierarchy beyond the local community index (Figure 1a). This indicator takes on the value of one for societies organized as paramount chiefdoms or states and zero for ethnicities organized either at the village level or as small (petty) chiefdoms. Under error orthogonality of Murdock’s (1967) classification and O’Brien (1999) maps, this 2sls approach allows accounting for error-in-variables (see Angrist and Pischke (2008), Wooldridge (2002) and Krueger and Lindahl (2001) for an analogous application). Even when the noise in the two variables are correlated, this approach is still useful in gauging the importance of error-in-variables.¹²

Table 2 reports 2sls-country-fixed-effects estimates that extract the common across the two maps component of statehood. For completeness, we report results with all different transformations of luminosity both with and without controlling for log population density. Three key patterns emerge. First, the two proxies (maps) of statehood clearly have a common component, as the first-stage fit is quite strong. The first-stage F -statistic safely exceeds ten, minimizing concerns of “weak instrument” bias. Second, in all permutations the 2sls coefficient is positive and consistently significant. This implies that the component of distance to the centroid of large pre-colonial states in O’Brien’s map, captured by Murdock’s jurisdictional hierarchy beyond the local community index, is a significant correlate of contemporary development (as reflected in luminosity and population density). Third, in all permutations the 2sls coefficient is larger as compared to the respective OLS estimate (reported in panel *C* of Table 1), implying the simple OLS estimates were lower bounds of the true strength of the association. While the 2sls approach is not a panacea (as the errors in the two maps may be correlated), it suggests that the OLS estimates are conservative due to attenuation.

Sensitivity Analysis To gauge the role of error-in-variables in Michalopoulos and Papaioannou (2013), in Appendix Table 3 we report 2sls estimates associating ethnic development with the component of Murdock’s jurisdictional hierarchy index explained by the binary pre-colonial statehood measure from O’Brien (1999) map. In all permutations the 2sls coefficient

¹²We prefer “instrumenting” distance to pre-colonial kingdoms rather than the indicators of statehood, so as to avoid 2sls estimation with a binary endogenous variable and fixed effects. 2sls estimates where we instrument the binary statehood measures with Murdock data also produce significant positive coefficients.

is positive and highly significant (please note that the first-stage fit reported in Panel *B* is also quite strong). The 2sls estimates are also considerably larger than the OLS estimates reported in Michalopoulos and Papaioannou (2013), implying considerable attenuation.

We also estimated 2sls models excluding from the estimation low population regions (and regions close to the capitals). As shown in Appendix Table 4 the 2sls coefficient on Murdock’s jurisdictional hierarchy beyond the local community index is positive and highly significant even when we drop half the sample.

4.2 OLS Estimates

We also estimated (country-fixed-effects) OLS models augmenting the specification with (*i*) a dummy variable that takes on the value of one if either Murdock’s-map-based binary political centralization index is one or the pre-colonial statehood index based on O’Brien’s map equals one and (*ii*) a dummy variable that takes on the value of one when an ethnic group is classified as being political centralized by both Murdock’s data and O’Brien’s map. So the first indicator of political centralization is a noisy proxy, as the two data sources yield different classifications, while the second indicator of political centralization is much more accurate, as both maps assign the ethnic homelands as being part of a large pre-colonial state. This trichotomous classification yields that 126 (out of 667) ethnic groups (18.9%) were most likely part of pre-colonial states, 112 groups (16.7%) were perhaps part of organized states, while the remaining 429 ethnic homelands (64.3%)

Table 3 give the results. Two main patterns are evident. First, the coefficient on the binary political centralization index that switches to one when either source classifies an ethnic homeland as being part of a pre-colonial state is positive. Yet, while this shows that -on average- regional development is higher in the homelands of these groups as compared to acephalous-fragmented societies, in all permutations the estimate is statistically indistinguishable from zero. In line with attenuation stemming from error-in-variables, this insignificance is driven by lower coefficients (see Table 1), rather than by increased standard errors. Second, the coefficient on the binary political centralization index that identifies homelands that were part of pre-colonial states according to *both* maps is positive, large, and statistically significant (at the 99% confidence level) in all but one specification. These results thus further show that both the estimates in Michalopoulos and Papaioannou (2013) and in Table 1 are conservative, due to error on the proxies of pre-colonial statehood.

5 Conclusion

Using historical maps on African kingdoms and empires before colonization, our results in this short note provide further evidence on a robust link between pre-colonial statehood (political centralization) and contemporary African development, as reflected in satellite images of light density at night. The strong within-country link between pre-colonial statehood and ethnic-level development is quite robust and, if anything, strengthens when we (partially at least) account for noise in the underlying maps depicting pre-colonial African centralized states. Further research is needed in establishing causality and identifying the channels/mechanisms linking political centralization with development, both in Africa and elsewhere. Moreover, future work should elaborate on the factors shaping political centralization.

6 Data Appendix (not for publication)

6.1 Outcome Variables

Light Density at Night: Light Density is the average luminosity across pixels that fall within the unit of analysis. We use the average of the values in 2007 and 2008. In the regressions we use $\text{Log}(0.01 + \text{Average Luminosity})$ or $\text{Log}(\text{Average Luminosity})$ or $\text{Log}(\text{Average Luminosity per Capita})$. Available at http://www.ngdc.noaa.gov/dmsp/global_composites_v2.html.

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

6.2 Pre-colonial Political Institutions and Organization

Pre-colonial political centralization 1. Dummy variable that takes on the value one when the centroid of a country-ethnic region (in Murdock's Atlas of 1959) falls within the boundaries of a pre-colonial empire/large kingdom (as show in O' Brien's (1999) map) and zero otherwise Source: O'Brien (1999), Murdock (1959) and Besley and Reynal-Querol (2014).

Pre-colonial political centralization 2. Dummy variable that equals one if any part of the ancestral homeland of ethnic group i in country c has been part of a pre-colonial kingdom or empire. Source: O'Brien (1999), Murdock (1959) and Besley and Reynal-Querol (2014).

Distance to pre-colonial states. Natural logarithm of 0.01 plus distance of each country-ethnic region to the boundaries of the closest pre-colonial kingdom-empire, as shown in O' Brian (1999) map. Source: O'Brien (1999), Murdock (1959) and Besley and Reynal-Querol (2014).

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

<http://eclectic.ss.uci.edu/~drwhite/worldcul/EthnographicAtlasWCRevisedByWorldCultures.sav>.

Binary Political 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).

6.3 Control Variables

Area: Log (land area) at the country-ethnic homeland level.

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

Elevation: Average elevation in km in each ethnicity-country . *Source: National Oceanic and Atmospheric Administration (NOAA) and U.S. National Geophysical Data Center, TerrainBase, release 1.0 (CD-ROM), Boulder, Colorado.*

<http://www.sage.wisc.edu/atlas/data.php?incdataset=Topography>

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

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

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

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

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

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

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

References

- ACEMOGLU, D., S. NAIDU, P. RESTREPO, AND J. A. ROBINSON (2014): “Democracy Does Cause Growth,” NBER Working Paper 20004.
- ACEMOGLU, D., AND J. A. ROBINSON (2012): *Why Nations Fail? The Origins of Power, Prosperity, and Poverty*. Crown Publishers, New York, NY.
- ALSAN, M. (2014): “The Effect of the TseTse Fly on African Development,” *American Economic Review*, forthcoming.
- ANGRIST, J., AND J.-S. PISCHKE (2008): *Mostly Harmless Econometrics*. Princeton University Press, Princeton, NJ.
- BANDYOPADHYAY, S., AND E. GREEN (2012): “Precolonial Centralisation and Contemporary Development in Uganda,” EOPP Discussion Paper Series No. 26, STICERD, LSE.
- BESLEY, T., AND M. REYNAL-QUEROL (2014): “The Legacy of Historical Conflicts. Evidence from Africa,” *American Political Science Review*, forthcoming.
- 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.
- BRECKE, P. (1999): “Violent Conflicts 1400 A.D. to the Present in Different Regions of the World,” Paper prepared for the 1999 Meeting of the Peace Science Society in Ann Arbor, Michigan.
- CAMERON, C. A., J. GELBACH, AND D. MILLER (2011): “Robust Inference with Multi-Way Clustering,” *Journal of Business and Economic Statistics*, 28(2), 238–249.
- CHEN, X., AND W. D. NORDHAUS (2011): “Using Luminosity Data as a Proxy for Economic Statistics,” *Proceedings of the National Academy of Sciences*, 108(21), 8589–8594.
- CHIOVELLI, G. (2014): “Pre-Colonial Centralization, Colonial Activities and Development in Latin America,” mimeo, University of Bologna, Economics Department.
- CIOFFI-REVILLA, C. (1996): “Origins and Evolution of War and Politics,” *International Studies Quarterly*, 40(1), 1–22.
- DIAMOND, J. (1997): *Guns, Germs, and Steel. The Fates of Human Societies*. W.W. Norton and Co, New York, NY.

- DIPPEL, C. (2014): “Forced Coexistence and Economic Development - Evidence from Native American Reservations,” *Econometrica*, forthcoming.
- FENSKE, J. (2013): “Does Land Abundance Explain African Institutions?,” *Economic Journal*, 123(573), 1363–1390.
- (2014): “Ecology, Trade, and States in Pre-colonial Africa,” *Journal of the European Economic Association*, forthcoming.
- FORTES, M., AND E. EVANS-PRITCHARD (1940): *African Political Systems*. Oxford University Press.
- 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.
- HENDERSON, V. J., A. STOREYGARD, AND D. N. WEIL (2012): “Measuring Economic Growth from Outer Space,” *American Economic Review*, 102(2), 994–1028.
- HERBST, J. (2000): *States and Power in Africa*. Princeton University Press, Princeton, NJ.
- HODLER, R., AND P. A. RASCHKY (2014a): “Regional Favoratism,” *Quarterly Journal of Economics*, 129(2).
- (2014b): “Regional Favoratism,” *Quarterly Journal of Economics*, 129(2), 995–1033.
- 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.
- KRUEGER, A. B., AND M. LINDAHL (2001): “Education for Growth: Why and For Whom?,” *Journal of Economic Literature*, 39(4), 1101–1136.
- MICHALOPOULOS, S. (2012): “The Origins of Ethnolinguistic Diversity,” *American Economic Review*, 102(4), 1508–1539.
- MICHALOPOULOS, S., AND E. PAPAIOANNOU (2013): “Pre-colonial Ethnic Institutions and Contemporary African Development,” *Econometrica*, 81(1), 113–152.

- (2014a): “National Institutions and Subnational Development in Africa,” *Quarterly Journal of Economics*, 129(1), 151–213.
- (2014b): “On the Ethnic Origins of African Development: Traditional Chiefs and Pre-colonial Political Centralization,” *Academy of Management (Perspectives)*, forthcoming.
- MUNCK, G., AND J. VERKUILEN (2002): “Conceptualizing and Measuring Democracy. Evaluating Alternative Indices,” *Comparative Political Studies*, 35(1), 5–34.
- MURDOCK, G. P. (1959): *Africa: Its Peoples and their Culture History*. McGraw-Hill, New York, NY.
- (1967): *Ethnographic Atlas: A Summary*. The University of Pittsburgh Press, Pittsburgh, PA.
- O’BRIEN, P. (1999): *Atlas of World History*. Oxford University Press, Oxford, United Kingdom.
- PAPAIOANNOU, E., AND G. SIOUROUNIS (2008): “Democratization and Growth,” *Economic Journal*, 118(10), 1520–1551.
- PINKOVSKIY, M. (2013): “Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night,” mimeo MIT.
- PUTTERMAN, L., AND D. N. WEIL (2010): “Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality,” *Quarterly Journal of Economics*, 125(4), 1627–1682.
- STOREYGARD, A. (2014): “Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa,” mimeo Tufts University, Department of Economics.
- WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

Table 1: Pre-colonial Political Centralization and Contemporary Ethnic Development

Dependent Variable	Ln(Luminosity+0.01)		Ln(Luminosity)		Ln(Luminosity p.c.)		Ln(Pop.Dens)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Pre-colonial Political Centralization Index 1							
Pre-Colonial Pol. Centralization 1	0.5150**	0.2984*	0.6839***	0.4226**	0.3279**	0.4224**	0.4398**
Double-clustered s.e.	(0.2111)	(0.1777)	(0.2031)	(0.1642)	(0.1634)	(0.1642)	(0.1970)
Log Population Density in 2000		0.4923***		0.7345***		-0.2657***	
Double-clustered s.e.		(0.0663)		(0.0803)		(0.0802)	
Adjusted R ²	0.488	0.630	0.529	0.659	0.607	0.626	0.482
Panel B: Pre-colonial Political Centralization Index 2							
Pre-Colonial Pol. Centralization 2	0.4652**	0.2950*	0.6973***	0.4500***	0.3588**	0.4497***	0.3453*
Double-clustered s.e.	(0.2177)	(0.1661)	(0.1628)	(0.1308)	(0.1399)	(0.1308)	(0.1999)
Log Population Density in 2000		0.4929***		0.7314***		-0.2687***	
Double-clustered s.e.		(0.0649)		(0.0799)		(0.0797)	
Adjusted R ²	0.487	0.631	0.532	0.661	0.608	0.628	0.480
Panel C: Log Distance to Pre-colonial Kingdoms-Empires							
Log Distance to Pre-colonial State	-0.0791**	-0.0433*	-0.1043***	-0.0588**	-0.0419*	-0.0588**	-0.0730**
Double-clustered s.e.	(0.0317)	(0.0251)	(0.0271)	(0.0235)	(0.0245)	(0.0235)	(0.0307)
Log Population Density in 2000		0.4892***		0.7293***		-0.2709***	
Double-clustered s.e.		(0.0652)		(0.0804)		(0.0803)	
Adjusted R ²	0.491	0.631	0.532	0.659	0.607	0.626	0.485
Geography and Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1218	1218	886	886	886	886	1218

The table reports country fixed-effects OLS estimates, associating regional development with pre-colonial political centralization, as reflected in maps of large pre-colonial empires-kingdoms taken from Besley and Reynal-Querol (2014), who in turn use data from O' Brian (1999). In columns (1)-(2) the dependent variable is the log(0.01 + satellite light density at night) at the ethnicity-country level. In columns (3)-(4) the dependent variable is the log(satellite light density at night) at the ethnicity-country level. In columns (5)-(6) the dependent variable is the log(satellite light density at night)-log(population density) [Luminosity per capita]. In column (7) the dependent variable is the log(population density in 2000) at the ethnicity-country level. The set of controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The specifications in even-numbered columns also control for log population density in 2000. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

All panels of the table report two-stage-least-squares fixed-effects OLS estimates, associating regional development with the log distance to the boundaries of large pre-colonial kingdoms and empires, as depicted in O' Brien (1999) map (see Figure 1b). The first-stage specification associates log distance to the boundaries of large pre-colonial kingdoms and empires with the Gennaioli and Rainer (2006, 2007) binary political centralization index that is based on George Peter Murdock's 0-4 jurisdictional hierarchy beyond the local community index. A 0 indicates stateless societies, 1 and 2 indicate petty and large paramount chiefdoms (or their equivalent), 3 and 4 indicate large states. In all specifications we report the Cragg-Donald F-statistic that captures the strength of the first-stage. Panel A gives estimates in the full sample. In Panel B we exclude from the estimation small in terms of population country-ethnic homelands, using the 5th percentile of population as the cut-off. In Panel C we exclude from the estimation country-ethnic areas close to capital cities, using the 25th percentile of distance to the capital as the cut-off.

In columns (1)-(2) the dependent variable is the $\log(0.01 + \text{satellite light density at night})$ at the ethnicity-country level. In columns (3)-(4) the dependent variable is the $\log(\text{satellite light density at night})$ at the ethnicity-country level. In columns (5)-(6) the dependent variable is the $\log(\text{satellite light density at night}) - \log(\text{population density})$ [Luminosity per capita]. In column (7) the dependent variable is the $\log(\text{population density in 2000})$ at the ethnicity-country level. The set of controls includes the log of land area, the log of $(1 + \text{land area under water})$ (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The specifications in even-numbered columns also control for log population density in 2000. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Table 3: Pre-colonial Political Centralization and Ethnic Development
Accounting for Measurement Error on Pre-colonial Political Centralization
OLS Estimates Using Both Murdock (1959, 1967) and O' Brecke Data**

Dependent Variable	Ln(Luminosity+0.01)		Ln(Luminosity)		Ln(Luminosity p.c.)		Ln(Pop.Dens)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Political Centralization (One Source)	0.2387	0.1822	0.3602	0.132	0.0500	0.1309	0.1061
Double-clustered s.e.	(0.1737)	(0.1161)	(0.2862)	(0.1824)	(0.1607)	(0.1825)	(0.2448)
Political Centralization (Both Sources)	0.7445***	0.4176***	0.8763***	0.3999**	0.2312	0.4002***	0.6144***
Double-clustered s.e.	(0.1897)	(0.1230)	(0.2408)	(0.1650)	(0.1598)	(0.1650)	(0.2085)
Log Population Density in 2000		0.5322***		0.7382***		-0.2618*	
Double-clustered s.e.		(0.0708)		(0.1380)		(0.1380)	
Adjusted R ²	0.5410	0.6800	0.5640	0.6820	0.6020	0.6200	0.5
Observations	667	667	514	514	514	514	667
Rich Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports country fixed-effects OLS estimates, associating regional development with two proxies of pre-colonial political centralization. Political Centralization (One Source) is an indicator that takes on the value of one if either Murdock's-map-based binary political centralization index is one or the pre-colonial statehood index based on O'Brian's (1999) map equals one (and zero otherwise). So this index captures with some ambiguity whether the country-ethnic homeland was part of a large pre-colonial African kingdom-state. Political Centralization (Both Sources) is an indicator variable that takes on the value of one when an ethnic group is classified as being politically centralized by Murdock's jurisdictional hierarchy index (Figures 1) and its homeland was part of a large kingdom-empire, as depicted in O'Brian's (1999) map (Figure 1b).

In columns (1)-(2) the dependent variable is the log(0.01 + satellite light density at night) at the ethnicity-country level. In columns (3)-(4) the dependent variable is the log(satellite light density at night) at the ethnicity-country level. In columns (5)-(6) the dependent variable is the log(satellite light density at night)-log(population density) [Luminosity per capita]. In column (7) the dependent variable is the log(population density in 2000) at the ethnicity-country level. The set of controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The specifications in even-numbered columns also control for log population density in 2000. The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 1:
Pre-colonial Political Centralization and Ethnic-level Contemporary Development
Heterogeneity with Respect to Population in 2000**

Excluding	<u>1%</u> (1)	<u>5%</u> (2)	<u>10%</u> (3)	<u>25%</u> (4)	<u>50%</u> (5)
Panel A: Dependent Variable - Log(Luminosity+0.01)					
Pre-Colonial Pol. Centralization 2	0.2928*	0.2948**	0.2844**	0.3160**	0.3475*
Double-clustered s.e.	(0.1603)	(0.1414)	(0.1444)	(0.1353)	(0.1948)
Log Population Density	Yes	Yes	Yes	Yes	Yes
Adjusted R ²	0.657	0.678	0.689	0.730	0.746
Observations	1205	1157	1096	913	609
Panel B: Dependent Variable - Log(Luminosity)					
Pre-Colonial Pol. Centralization 2	0.4464***	0.4006***	0.4308***	0.4317***	0.3682**
Double-clustered s.e.	(0.1316)	(0.1207)	(0.1241)	(0.1150)	(0.1701)
Log Population Density	Yes	Yes	Yes	Yes	Yes
Adjusted R ²	0.669	0.682	0.679	0.708	0.721
Observations	884	874	855	767	565
Panel C: Dependent Variable - Log(Luminosity per capita)					
Pre-Colonial Pol. Centralization 2	0.3662***	0.3455***	0.3915***	0.4469***	0.4112**
Double-clustered s.e.	(0.1359)	(0.1214)	(0.1258)	(0.1141)	(0.1634)
Adjusted R ²	0.619	0.614	0.621	0.616	0.613
Observations	884	874	855	767	565
Panel D: Dependent Variable - Log(Population Density 2000)					
Pre-Colonial Pol. Centralization 2	0.3023	0.2806	0.2966*	0.2956**	0.2449***
Double-clustered s.e.	(0.1941)	(0.1851)	(0.1636)	(0.1252)	(0.0787)
Adjusted R ²	0.502	0.523	0.548	0.587	0.641
Observations	1205	1157	1096	913	609
Geography and Location Controls	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	No	Yes	Yes	Yes	Yes

The table reports country fixed-effects OLS estimates, associating regional development with pre-colonial political centralization, as reflected in maps of large pre-colonial empires-kingdoms taken from Besley and Reynal-Querol (2014), who in turn use data from O' Brian (1999). In Panel A the dependent variable is the $\log(0.01 + \text{satellite light density at night})$ at the ethnicity-country level. In Panel B the dependent variable is the $\log(\text{satellite light density at night})$ at the ethnicity-country level. In Panel C the dependent variable is the $\log(\text{satellite light density at night}) - \log(\text{population density})$ [Luminosity per capita]. In Panel D the dependent variable is the $\log(\text{population density in 2000})$ at the ethnicity-country level.

In column (1) we exclude the bottom percentile of population; in columns (2), (3), (4), and (5) we exclude the bottom 5%, 10%, 25% and 50% of (country-ethnicity) observations based on population. The pre-colonial political centralization variable is a dummy variable that takes on the value of one when the centroid of an ethnic homeland falls within the boundaries of a large pre-colonial kingdom-empire. The set of controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. In all columns in Panels A, B, and C we also control for log population density in 2000. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The Data Appendix in Michalopoulos and Papaioannou (2013) gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Appendix Table 2: Pre-colonial Political Centralization and Ethnic Development Heterogeneity. Excluding Ethnic Regions close to the Capital Cities

	Ln(Luminosity+0.01)		Ln(Luminosity)		Ln(Luminosity p.c.)		Ln(Pop.Dens)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Political Centralization Index 1							
Pol. Centralization 1	0.5494**	0.3641*	0.7561***	0.4998**	0.3625*	0.4998**	0.4554**
Double-clustered s.e.	(0.2321)	(0.1906)	(0.2513)	(0.1966)	(0.1947)	(0.1966)	(0.2179)
Log Pop. Dens. 2000		0.4070***		0.6513***		-0.3491***	
Double-clustered s.e.		(0.0690)		(0.0938)		(0.0937)	
Adjusted R ²	0.472	0.598	0.535	0.639	0.634	0.662	0.650
Panel B: Political Centralization Index 2							
Pol. Centralization 2	0.5931***	0.4339**	0.8327***	0.5738***	0.4294***	0.5737***	0.3942*
Double-clustered s.e.	(0.2285)	(0.1744)	(0.2164)	(0.1641)	(0.1645)	(0.1641)	(0.2162)
Log Pop. Dens. 2000		0.4039***		0.6422***		-0.3581***	
Double-clustered s.e.		(0.0671)		(0.0941)		(0.0941)	
Adjusted R ²	0.479	0.603	0.542	0.642	0.637	0.666	0.64
Panel C: Log Distance to Pre-colonial Kingdoms-Empires							
Log Distance to Pre-colonial States	-0.0826**	-0.0530*	-0.1085***	-0.0664**	-0.0434	-0.0664**	-0.0733**
Double-clustered s.e.	(0.0356)	(0.0275)	(0.0360)	(0.0275)	(0.0275)	(0.0275)	(0.0339)
Log Pop. Dens. 2000		0.4033***		0.6462***		-0.3541***	
Double-clustered s.e.		(0.0675)		(0.0945)		(0.0945)	
Adjusted R ²	0.477	0.599	0.538	0.639	0.634	0.662	0.474
Geography-Location Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	913	913	639	639	639	639	913

The table reports country fixed-effects OLS estimates, associating regional development with pre-colonial political centralization, as reflected in maps of large pre-colonial empires-kingdoms taken from Besley and Reynal-Querol (2014), who in turn use data from O' Brian (1999). In columns (1)-(2) the dependent variable is the $\log(0.01 + \text{satellite light density at night})$ at the ethnicity-country level. In columns (3)-(4) the dependent variable is the $\log(\text{satellite light density at night})$ at the ethnicity-country level. In columns (5)-(6) the dependent variable is the $\log(\text{satellite light density at night}) - \log(\text{population density})$ [Luminosity per capita]. In column (7) the dependent variable is the log of population density in 2000. In all columns we exclude from the estimation country-ethnic homelands close to the capitals using a cut-off the 25th percentile of distance to the capital city. In Panel A the pre-colonial political centralization variable is a dummy variable that takes on the value of one when the centroid of an ethnic homeland falls within the boundaries of a large pre-colonial kingdom-empire. In Panel B the pre-colonial political centralization variable is a dummy variable that takes on the value of one when any part of the ethnic homeland falls within the boundaries of a large pre-colonial kingdom-empire.

The set of controls includes the log of land area, the log of $(1 + \text{land area under water})$ (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The Data Appendix in Michalopoulos and Papaioannou (2013) gives detailed variable definitions and data sources. In columns (2), (4), and (6) we also condition on the log of population density in 2000. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 3: Pre-colonial Political Centralization and Ethnic Development
Accounting for Measurement Error on Pre-colonial Political Centralization
2SLS Estimates**

Dependent Variable	Ln(Luminosity+0.01)		Ln(Luminosity)		Ln(Luminosity p.c.)		Ln(Pop.Dens)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: 2sls Estimates							
Jurisdictional Hierarchy	1.1786***	0.8294**	1.3753***	0.9573***	0.6717**	0.9563***	0.7538*
Double-clustered s.e.	(0.3969)	(0.3637)	(0.3246)	(0.3440)	(0.3246)	(0.3439)	(0.3837)
Log Population Density in 2000		0.4633***		0.5949***		-0.4050**	
Double-clustered s.e.		(0.0752)		(0.1624)		(0.1624)	
Panel B: First-Stage Estimates							
Pre-colonial Political Centralization 2	0.4396***	0.4086***	0.4625***	0.4090***	0.4624***	0.4090***	0.4396***
Double-clustered s.e.	(0.1284)	(0.0926)	(0.1493)	(0.1319)	(0.1493)	(0.1319)	(0.1284)
Rich Set of Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	667	667	514	514	514	514	667

The table reports two-stage-least-squares fixed-effects OLS estimates, associating regional development with pre-colonial political centralization. In columns (1)-(2) the dependent variable is the log(0.01 + satellite light density at night) at the ethnicity-country level. In columns (3)-(4) the dependent variable is the log(satellite light density at night) at the ethnicity-country level. In columns (5)-(6) the dependent variable is the log(satellite light density at night)-log(population density) [Luminosity per capita]. In column (7) the dependent variable is the log(population density in 2000) at the ethnicity-country level. The set of controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The specifications in even-numbered columns also control for log population density in 2000. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

**Appendix Table 4: Pre-colonial Political Centralization and Ethnic Development
Accounting for Measurement Error and Population
2SLS Estimates**

Excluding	<u>1%</u> (1)	<u>5%</u> (2)	<u>10%</u> (3)	<u>25%</u> (4)	<u>50%</u> (5)
Jurisdictional Hierarchy [Murdock (1967)]	0.9563***	0.7396**	0.7199**	0.7440***	0.7301**
Double-clustered s.e.	(0.3439)	(0.2919)	(0.2904)	(0.2730)	(0.3254)
Log Population Density in 2000	-0.4050**	-0.2427*	-0.2284	-0.0955	0.0159
Double-clustered s.e.	(0.1624)	(0.1350)	(0.1500)	(0.1785)	(0.1979)
Geography and Location Controls	Yes	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	514	509	502	466	368

The table reports country fixed-effects two-stage-least-squares (2SLS) estimates, associating regional development with pre-colonial political centralization, as reflected in Murdock's (1967) jurisdictional hierarchy beyond the local level, that is instrumented with an indicator that takes on the value of one when the country-ethnic homeland falls within the boundaries of a large pre-colonial kingdom-empire (using the maps of Besley and Reynal-Querol (2014), who in turn use data from O' Brian (1999). In all specifications the dependent variable is the log(satellite light density at night)-log(population density) [Log Luminosity per capita]. The set of controls includes the log of land area, the log of (1 + land area under water) (lakes, rivers, and other streams), distance of the centroid of each country-ethnic homeland from the respective capital, from the sea coast, from the national border, an index of land suitability for agriculture, mean elevation, a malaria stability index, a diamond mine indicator, and an oil field indicator. The specifications in even-numbered columns also control for log population density in 2000. This set of controls follows exactly Michalopoulos and Papaioannou (2013). The Data Appendix gives detailed variable definitions and data sources. All specifications include country fixed effects (constants not reported). The table reports in parentheses double-clustered standard errors at the country and the ethno-linguistic family dimensions. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.