Questionable Inference on the Power of Pre-Colonial Institutions in Africa

Denis Cogneau, Yannick Dupraz

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2014 – 25

Questionable Inference on the Power of Pre-Colonial Institutions in Africa

Denis Cogneau
Yannick Dupraz

JEL Codes: O10, O43, N17, R12, Z13

Keywords: Institutions, Africa, Population, Development, Light intensity at night

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQU

Questionable Inference on the Power of Pre-Colonial Institutions in Africa

Denis Cogneau*, Yannick Dupraz†

May 2014

Abstract

In their paper "Pre-Colonial Ethnic Institutions and Contemporary African Development" [Econometrica 81(1): 113-152], Stelios Michalopoulos and Elias Papaioannou claim that they document a strong relationship between pre-colonial political centralization and regional development, by combining Murdock's ethnographic atlas (1967) with light density at night measures at the local level. We argue that their estimates do not properly take into account population effects. Among lowly populated areas, luminosity is dominated by noise, so that with linear specifications the coefficient of population density is biased downwards. We reveal that the identification of the effect of ethnic centralization very much relies on these areas. We implement a variety of models where the effect of population density is non-linear, and/or where the bounded or truncated nature of luminosity is taken into account. We conclude that the impact of ethnic-level political centralization on development is all contained in its long-term correlation with population density. We also abstract from the luminosity-population nexus by analyzing survey data for 33 countries. We show that individual-level outcomes like access to utilities, education, asset ownership etc. are not correlated with ethnic-level political centralization.

Keywords: Institutions, Africa, Population, Development, Light intensity at night.
JEL classification codes: O10, O43, N17, R12, Z13

*Denis Cogneau (corresponding author) is research fellow and associate professor at Paris School of Economics - IRD; his e-mail address is denis.cogneau@psemail.eu.

†Yannick Dupraz is a PhD student at Paris School of Economics (PSE) and Ecole des Hautes Etudes en Sciences Sociales (EHESS).

In their paper "Pre-Colonial Ethnic Institutions and Contemporary African Development" [Econometrica 81(1): 113-152], Stelios Michalopoulos and Elias Papaioannou (M&P from now on) document a relationship between pre-colonial ethnic political centralization and regional development. For this purpose, they put together two original data sources. First, the ethnographic atlas from George Peter Murdock (1967), in the revised and extended version proposed by J. Patrick Gray (1999), is combined with a map of African ethnicities drawn by Murdock himself in 1959, so that 441 ethnic homelands are matched with information on pre-colonial political institutions and other ethnic features.[1] By superimposing present-day country boundaries on Murdock's map, they create a dataset of 683 country-ethnicity areas, where a large number of ethnic homelands (193 out of 441) appear to have been split (in up to six parts), out of the colonial scramble for Africa. Second, they match this country-ethnicity database with satellite imagery data providing, for each 0.125×0.125 decimal degrees pixel, a measure of light intensity at night in the years 2007 and 2008. They draw from previous works of geographers (Elvidge et al., 1997), ecologists (Doll et al., 2006), and economists (Henderson et al., 2012; Chen and Nordhaus, 2010 and 2011) and argue that light intensity at night is a good proxy for economic development at the local level, if only through the correlation between connection to electricity and income. Then they notice that homelands corresponding to more politically centralized groups display above average light density today, or that pixels lying inside the homelands of centralized groups are more often lit than dark. Using a large set of geographic, demographic and economic controls, they try to convince the reader that the cen-

[1] In their paper p. 119, M&P say 490, but their replication database only contains 441 distinct areas from Murdock's map.

tralization effect is not confounded by other long term determinants, so that they truly identify the long-lasting effect of this specific ethnic feature on present day economic development.

The critique of the historical and anthropological accuracy of Murdock's atlas is out of the scope of our contribution. Being the only large-scale quantitative database on pre-colonial African institutions, it is increasingly used by researchers who wish to test continent-wide conclusions on how long-term history affects contemporary development (Nunn, 2008; Nunn and Puga, 2012; Nunn and Wantchekon, 2011; Fenske, 2012; Michalopoulos and Papaioannou, 2012, 2013). Yet this map of ethnic homelands can only be indicative, as even structured pre-colonial kingdoms drew no maps and had no formal borders (Herbst 2000). In M&P's dataset, some country-ethnicity areas display very small population sizes, and the smallest ones almost all correspond to splits of Murdock areas by international borders. It is unclear whether these small jigsaw puzzle pieces have some reality on the ground.[2] The presence of these underpopulated splits would not necessarily be an issue, provided that they do not carry too much weight in the analysis; unfortunately they do.

With regards to the use of satellite light intensity at night as a proxy for development, we note that geographers or ecologists first used light density to measure population density or more specifically urbanization. Then economists considered using it for measuring either the growth or the level of economic activity, as more traditionally measured by GDP. However, any indicator like GDP carries more population effects than *per capita* measures. If only for this reason we think that

[2]Conversely, we note that the Senufo area is located within the borders of Côte d'Ivoire, whereas it should surely extend to southern Mali as well.

controlling for population is a must, if the aim is to capture economic development or welfare. Furthermore, Chen and Nordhaus (2010, 2011) concluded on African cross-sectional data that the relationship between luminosity and output is uncertain outside of highly populated areas. Although they present light density *per se* as a rightful measure of economic development, M&P introduce population density as a control variable in many of their specifications. This could solve the problem if the impact of population on light density was correctly accounted for: here lies our main disagreement.

In the remainder of this paper, we first argue that one cannot use luminosity data to capture cross-sectional differences in economic development, without carefully accounting for the spatial distribution of population (section 1). Second, we show that sparsely populated areas are responsible for the apparent correlation between light density and pre-colonial political centralization, whereas these are areas for which noise dominates luminosity, both at the country-ethnicity and pixel levels (section 2). We also reveal that a whole range of non-linear econometric models, taking into account both the non-linear effect of population density and the bounded/truncated nature of the luminosity data, cancels out the impact of centralization (section 3). We argue that the use of survey data by M&P is similarly biased by population effects; using an extended database covering 33 countries, we confirm the absence of correlation between the welfare of individuals and the centralization indicator for the place of residence (section 4). In conclusion, we discuss the correlation of political centralization with population density, that finally subsumes all its relation to economic development. While causality could have run both ways, common geographical variables could also have determined the joint evolution of demography and of political institutions (section 5).

I. Economic Development, Population, and Luminosity

We start by going back to the seminal contributions of Henderson et al. (2012) and Chen and Nordhaus (2010, 2011), in order to examine what luminosity data should tell us about economic development or welfare. Henderson et al. argue that the *time variation* of light density (i.e. light intensity per square kilometers) is well correlated with the growth of GDP at country level - as electricity production in kWh also is. They also extend the use of light intensity to within-country regional growth, when they compare the growth of coastal areas to the growth of hinterlands among African countries. However they never analyze cross-sectional differences (compare levels of GDP, or of luminosity). Chen and Nordhaus acknowledge that luminosity data can be useful for measuring income level or growth in countries where national accounts are of low quality, in particular African ones. However they notice that in cross-section "luminosity at the low end is dominated by noise and errors [...] rather than economic activity" (p.7, 2010 version). For Africa, this means that the relationship between luminosity and income is uncertain, outside of urban areas where output and population are concentrated.

Using the 2008 cross-section of Henderson et al. data, we first confirm that at the country-level luminosity *per capita* is a much better predictor of GDP *per capita*, compared to light density. In the top panel of Table 1, we analyze a worldwide sample of 163 countries, whereas in the bottom panel we focus on the 50 African countries. In columns (1) to (3), we successively regress log GDP per capita in international dollars (i.e. PPP-corrected) on log luminosity per square kilometers (what M&P call "light density"), log luminosity per capita (i.e. light density minus log population density), and then add log population density as an additional control variable. The right part of the table (col. 4 to 6) provides the estimates of the same three regressions, but with population weights instead of unitary weights for each country. Unsurprisingly, the R-squared of the regressions, as well as the coefficient of the luminosity variable, are unambiguously higher with per capita measures, whether we consider the world as a whole or Africa only. For the African sample, the share of between-country variance explained by luminosity increases from 36% to 77% when shifting to per capita luminosity, or 56% to 81% with population-weighted estimates. Log population density adds little in terms of explanatory power (columns 3 and 6). In the African sub-sample, its coefficient is not statistically significant at 90% confidence and the adjusted R-squared even slightly decreases.

From these elementary results at the country level, we deem luminosity per capita to be a better proxy for income per capita differences than light density. Henderson et al. do not make any cross-sectional analysis but rather study differences in output growth as proxied by variations of light density across time, so that variations in population density are not involved. In contrast, Pinkovskiy (2011) studies whether levels of economic development vary discontinuously at international borders. For this, he explicitly transforms light density into light per capita, by using pixel-level estimates of population density. Likewise, we see no reason for using light density instead of luminosity per capita when the aim is the analysis of cross-sectional differences in economic development. In estimates that do not control for population density, we rather argue that using light density is irrelevant.

We then turn to M&P data at the country-ethnicity level, and estimate models that replace the light density outcome by luminosity per capita (Table 2). To be as transparent as possible, we exclude the 166 (out of 683) zero luminosity areas, hence focusing on what M&P call the intensive margin estimates. We also necessarily exclude the 5 unpopulated areas.[3] As control variables, we introduce country fixed effects, because the strength of M&P results precisely stems from the fact that ethnic centralization explains within-country variation in light density. Of course, here we do not include a linear effect of population density, as it would bring us back to M&P intensive margin estimates.[4] We also exclude other location or geographic controls. Each column reports the coefficient of one specific measure of pre-colonial ethnic political centralization as recoded from the Ethnographic Atlas by M&P: first jurisdictional hierarchy, an ordinal variable taking integer values from 0 to 4; second, a binary recoding of the former, with 0/1 coded as 0 and 2/3/4 coded as 1; third a series of three binary variables coding for petty chiefdoms (juris. hierarchy equal to 1), paramount chiefdoms (2), and pre-colonial states (3/4). Like in Table 1, the right part of the table reports population-weighted estimates.

As M&P never consider this latter reweighing, the left part of Table 2 is the most comparable to their own estimates using light density as an outcome. Columns (1) to (3) of Table 2 could be respectively compared to columns (1), (5) and (9) of table III, panel B (intensive margin) in M&P's article (p.129). The comparison shows that the impact of political centralization is much decreased when shifting to luminosity per capita. The coefficient of jurisdictional hierarchy goes from a very significant +0.3279 point estimate (s.e.= 0.1238) down to an in-

[3]M&P always include them thanks to "winsorizing", i.e. adding 0.01 to population density before applying the log transformation. In the pixel-level analysis, they withdraw unpopulated pixels.

[4]Except for the exclusion of zero population areas and for the above mentioned winsorizing of population density, which both make little difference.

significant +0.1109 (s.e.= 0.0725), meaning a threefold reduction. For the binary variable, the coefficient remains significant at 90% confidence, although it is also largely reduced from +0.4819 to +0.1879. The third comparison reveals that the reduction affects the coefficient of paramount chiefdoms, and overwhelmingly the one of pre-colonial states.

The right part of Table 2 (columns (4) to (6)) also reveals that giving each country-ethnicity its present-day population weight cancels out any effect of ethnic centralization, whose magnitude turns very close to zero. Population weighting can be debated. In its favor, one could argue that the relevance of ethnic centralization for economic development should be judged in the light of its importance for the majority of today's African people, and not for the average inhabitant of hardly populated areas. Then, from this standpoint ethnic centralization could already appear meaningless for development. Yet in the remainder of this paper we remain agnostic with regards to population weighting, although using population weights cancels out the effect of centralization in each and every specification. Our analysis reveals the spurious role played by sparsely-populated areas in M&P estimates, and population-weighted estimates is only one of the means through which the impact of such outliers can be diminished.

Besides, we also relax the equivalence-scale assumption that is implicit in per capita measures, by allowing total welfare (or luminosity) to have a non-unitary elasticity with respect to population. We only argue that the effects of population and of surface area should be carefully accounted for, before using light density as a meaningful indicator for economic development and before drawing strong inferences on the impact of correlates of luminosity that are also correlates of population density.

II. Noisy Luminosity and Underpopulated Areas

We exhibit a non-linear relationship between log luminosity and log population: in sparsely populated areas, light intensity does not move with population, which suggests that it is dominated by noise. At the country-ethnicity level, Figure 1 shows a plot of the two variables and reports the curve of a locally weighted regression ; Figure 2 adds location and geographic controls, and in particular land area. Both figures suggest that at low population levels the relationship between luminosity and population is essentially flat, while it is strongly increasing at higher levels.[5] Although they examine the relationship between light density and output density, Chen and Nordhaus (2010, 2011) obtain a comparable figure for Africa: a flat curve at low levels of output, then a strong increasing relationship.[6]

In order to account for this non-linear relationship between luminosity and population, we augment M&P's specification with a simple spline function kinked at 50,000 inhabitants. 143 country-ethnicity areas, out of 683, lie below the 50,000 threshold. There is little chance for a large city to be found in these areas: using a georeferenced database for cities of at least 20,000 inhabitants, we indeed show that only 2% of these areas below 50,000 contain a city, whereas 51% of the more populated areas have at least one.[7] Column (2) of Table 3 shows that the impact of ethnic centralization on light density is halved and loses significance

[5]In Figure 1, the increasing part is partly driven by land area. In Figure 2, land area is one of the controls, but both log luminosity and log population happen to display the same correlation with log land area, the coefficients of this latter variable being respectively 0.7360 (s.e.=0.0755) and 0.7190 (s.e.=0.0533) in each of the two regressions on controls.

[6]See Figure 5, p.30, in Chen and Nordhaus (2010), and Figure SI-2, p.SI-21 of the Supporting Information of Chen and Nordhaus (2011).

[7]This database was constructed from population census official documents collected on the website `http://www.citypopulation.de/`. The data are coarse. For each country, we retained a census year as close as possible to the year 1990. However, available years vary between 1970 (Angola) to 2005 (Somalia).

when making this distinction between unpopulated and populated areas through a spline function. Focusing on intensive margins estimates (bottom panel), we observe that for areas below 50,000 inhabitants log light density does not vary with population and decreases with log land area at a rate close to one (0.8759), suggesting that the numerator of light density is dominated by noise. Above the 50,000 threshold, light intensity increases with population density, again with an elasticity close to one.[8]. When adopting the $ln(0.01 + x)$ transformation used by M&P to include zero luminosity areas (top panel), the coefficients of population and land area are changed, but the same result holds regarding the coefficient of ethnic centralization.[9]

In other columns of Table 3, we reveal with even simpler means the sensitivity of M&P results to the presence of sparsely populated country-ethnicity areas. In column (4), we drop the 72 areas whose total population lies below 10,000 people, i.e. 10.5% of the total sample; when focusing on areas with positive luminosity, the sample reduction is even much lower: 20 areas only are withdrawn, i.e. 3.9% of the sample. This is enough to downsize the coefficient of the centralization variable by half and make it lose statistical significance. When dropping all areas below 50,000 inhabitants, the coefficient further decreases (column (5)). Population-weighted OLS estimates even produce a negative coefficient (column (6)). In each case, the coefficient of population density correlatively raises, going from 0.4375 in M&P's estimate to 1.0191 in population-weighted estimates.[10]

[8]Indeed, the coefficient of log population is 1.0287, while the one of log land area is -0.8759

[9]In columns (1) and (2), 16 zero population areas are not included. If we add them back by applying the $ln(0.01+x)$ transformation to the population variable, results do not change at all.

[10]Appendix Table A1 provides more detailed population-weighted estimates. The coefficient of ethnic centralization already gets insignificant with only country fixed effects, i.e. before controlling for population density or other variables. It is also shown that reweighing ethnicity areas within each country, while giving each country the same unitary weight, is sufficient to

Then M&P pixel-level estimates are equally sensitive to low-population units, as shown in Table 4. The coefficient falls down to a low and insignificant value when discarding the pixels with less than four people per square kilometer or when implementing population weights. The population density threshold is inspired from Min (2008), who calculates a minimum population threshold above which one can reliably assume that the lack of visible nighttime light output indicates a lack of electrification and outdoor lights; based on the median population in the most dimly lit cells across the entire globe, he proposes a threshold of 28 people for 2.7 km$\times$2.7 km pixels, i.e. around four people for one square kilometer. Columns (2) and (5) of Table 4 show that considering the sub-sample of sufficiently dense pixels, i.e. withdrawing the 30.8% less densely populated pixels from the analysis, again leads to a halved and insignificant coefficient for pre-colonial centralization, irrespectively of luminosity being measured in discrete form (lit/unlit, col. (2)) or in continuous form (ln(0.01+light density), col. (5)). Keeping all pixels but giving to each its population weight leads to similar results (columns (3) and (6)).[11] Likewise, appendix Tables A3 and A5 show that shifting to population-weights systematically cancels out the effect of pre-colonial centralization in the more sophisticated "contiguous ethnic homeland" estimates that M&P consider in the last section of their article (Table VII p.142 and Table VIII p.145-147).

obtain this result.

[11] Appendix Table A2 provides more detailed pixel-level population-weighted estimates. Here again, the coefficient already gets insignificant with only country fixed effects, i.e. before controlling for population density or other variables.

III. Non-Linear Models for Luminosity

In this section, we additionally argue that the way zero luminosity units are accounted for leads to upwardly biased estimates for the coefficient of pre-colonial centralization. This is again because OLS linear estimators impose a downward bias to the effect of population on luminosity, and hence an upward bias on any variable that is positively correlated with population. This is especially true for pixel-level estimates. More than 83% pixels are unlit, so that log light density is very much truncated on the left side.

When analyzing the lit/unlit dummy variable, the linear probability model (LPM) used by M&P produces many predicted probabilities outside of the [0;1] interval, as the marginal effect of population density cannot remain constant all along the support of light density. When using instead logit, probit or Lewbel-special-regressor estimators, the average marginal effect of population density increases while the one of pre-colonial centralization collapses. This is shown in Table 5. Column (1) merely reproduces M&P result with the LPM. The five following columns report logit and probit estimates. Column (2) for logit and columns (4) and (6) for probit report "pseudo" maximum likelihood estimates, where the double-clustering of errors by ethno-linguistic family and country is not plugged into the maximum likelihood, standard errors being computed ex-post from a sandwich estimator based on scores; column (6) additionally models a multiplicative heteroscedasticity related to population density.[12] Columns (3) for logit and column (5) for probit are mathematically more consistent maximum likelihood estimates, where gaussian random effects are modeled at the ethnicity level. All

[12] $\sigma_i^2 = exp(\theta.V_i)$, where V is log population density.

those estimates provide very consistent results: the ratio between the coefficient of ethnic centralization and the coefficient of population density, which is around 0.40 according to the LPM, goes down to very low values ranging from 0.004 to 0.118; correlatively, the average marginal effect of ethnic centralization is very much reduced and loses all significance, as it ranges from 0.03 to 0.80 percentage points, compared to 2.65 pp with the LPM.[13] In column (7), using the Lewbel (2000) special regressor estimator allows us to relax the logistic or normal assumptions for the distribution of errors. Population density is assumed to be a valid special regressor: firstly exogenous, i.e. independent from the unobservable determinants of luminosity; secondly with large support, i.e. its variation covering the distribution of these unobservables. Even if strict exogeneity of population density is not granted, the large support assumption holds fairly well: light density indeed runs from zero in desert places to capped values in very densely populated areas like capital cities. The Lewbel estimate for the impact of centralization stands in line with the logit or probit estimates; even if the estimated coefficient is higher (at 1.2 pp), it is not statistically significant.[14] Appendix Table A4 additionally shows that shifting from the LPM to logit again cancels out the effect of pre-colonial centralization in the more sophisticated "contiguous ethnic homeland" estimates that M&P consider in the last section of their article.

Table 6 then extends the same kind of non-linear estimators to light density in continuous form, alternatively to a simple OLS using the "winsorized" trans-

[13]We also used the Paapke and Woolridge (1996) "fractional" logit or probit estimators to model the proportion of lit pixels at the ethnicity-country level. Results obtained are very close to the pixel-level logits and probits. See also Table 4 footnote.

[14]As recommended by Lewbel, a minimal trimming of extreme values of the conditional density is applied, limited to 0.1% of sample size. Additional trimming leads to a further reduction of the ethnic centralization coefficient.

formation of light density, i.e. ln(0.01+light density). Column (1) reproduces the OLS estimate obtained by M&P. Columns (2) and (3) then report a tobit estimate, assuming that unlit pixels lie below the observed minimum in terms of light density and that residuals are normally distributed. Column (2) reports a pseudo-maximum likelihood with a sandwich estimator for double-clustered standard errors, and column (3) a true maximum-likelihood with gaussian random effects at ethnicity level. In columns (4) and (5) we then implement the Lewbel (2007) special regressor estimator for truncated data, again making use of log population density as a special regressor, under the same assumptions as above; in column (4) we use an Epachenikov (quadratic) kernel estimator, and in column (5) the Lewbel and Schennach (2007) ordered data estimator for inverse conditional density. The two tobit point estimates for the marginal effect of ethnic centralization differ considerably in magnitude : +11.43 percentage points with the pseudo-maximum likelihood, -4.68 pp for the random effect version. The two Lewbel estimators also deliver coefficients of opposed signs: -3.87 pp with the quadratic kernel, and +4.20 with the ordered data estimator. However none of those point estimates are significantly different from zero at 90% confidence. The last column of Table 6 considers the luminosity per capita outcome already introduced in Table 2. In this case, if we are willing to accept the exclusion restriction that log population density does not impact the intensive margin of luminosity per capita and only the likelihood of being unlit, then we can implement a Heckman selection estimator where log population density plays the role of the identifying variable for the selection part, aside to the joint normality assumption. Here again we obtain a negative and insignificant point estimate for the coefficient of pre-colonial centralization (-5.38 pp, see Table 6 column (8)).

IV. Pre-Colonial Centralization and Development using Survey Data

The main reason for using luminosity at night as a proxy for development is the absence of "geocoded high resolution measures of economic development spanning all Africa" (Michalopoulos and Papaioannou 2012). Although this is an undeniable fact, geocoded survey data spanning several African countries are available: in their appendix, M&P present results using 2005 Afrobarometer surveys, covering 17 Sub-Saharan African countries; aggregating Demographic and Health Surveys from different waves, we are able to construct a dataset spanning 33 African countries and comprising 245,044 households. A detailed description of this dataset's construction and its limitations is provided in appendix B. Working with survey data has its drawbacks (less countries are covered and geolocation is given with error), but it offers many advantages, allowing us to work with more direct measures of development than luminosity at night. Using our dataset, we find no correlation between various welfare indicators and the centralization index of the place of residence.

Our merged DHS dataset spans 33 African countries and 447 of the 683 country-ethnicity areas used by M&P. Among the 236 absent country-ethnicity areas, 145 are from 16 countries for which we could not find geocoded DHS surveys[15]. 91 are so sparsely populated that none of their inhabitants were sampled in the DHS surveys (the average population of these areas is 85).

When working with survey data (in their appendix), M&P consider the country-

[15]These countries are Botswana, Congo, Djibouti, Algeria, Eritrea, Western Sahara, Gambia, Guinea-Bissau, Equatorial Guinea, Libya, Mauritania, Sudan, Somalia, Tchad, Tunisia and South Africa.

ethnicity area as the unit of analysis, giving each one a unit weight. This is questionable, because there are very large variations in population size across country-ethnicity areas (from a few individuals to tens of thousands). Not weighting by population amounts to putting too much weight on individuals living in sparsely populated regions, while it is very likely that these regions will be specific both in terms of pre-colonial centralization and in terms of development (it would not be a problem if we were able to control perfectly for any confounding factor, but this is obviously not the case). The problem is made worse by the fact that sparsely populated areas are often situated at the countries' margins: the splitting of Murdocks' ethnic groups by international borders therefore increased the number of sparsely-populated country-ethnicity areas. Since we are considering questions of welfare, considering individuals or households as the unit of analysis seems more intuitive. Our main regressions are run on individuals, households or villages using the DHS sampling weights.[16] However, we also consider the specification where each country-ethnicity area is given a unit weight.

One advantage of using survey-based data instead of luminosity at night is that controlling for population density is not mandatory anymore. Controlling for population density (or urbanization) in the regression of pre-colonial centralization on contemporary development can be questioned, as population density could be one of the channels through which pre-colonial centralization influences contemporary development (it is a "bad control", Angrist & Pishke, 2009). Obviously, contemporary population density is very likely correlated with population density at the time when pre-colonial institutions were set up, so that the causal relation-

[16]Transformed to take into account differences in population sizes across countries, see appendix B); abstaining from this transformation does not affect our findings.

ship between density and centralization could go in the other way, but, to avoid the problem of the "bad control", in this section we can afford not controlling for population density[17].

We investigate the correlation between pre-colonial centralization and the following variables: the probability that a village is connected to electricity (we consider that a village is connected to electricity if more than 50% of households have electricity), presence of electricity in the household, an asset-based wealth index (normalized to have unit variance), whether an individual has ever been to school and whether she completed primary schooling. For binary variables, we use a linear probability model, which allows for double-clustering of the standard errors using the method of Cameron, Gelbach and Miller (2011). Results using a probit model are presented in appendix Table A7 (logit gives comparable results); as already mentioned in the previous section, ex-post clustering with a sandwich estimator on scores is not fully justified, since point estimates are potentially biased when observations are not independently and identically distributed. Yet, results of LPM and pseudo-probit are very close, suggesting that non-linearity is not an issue here.

As can be seen in Table 7, the effect of pre-colonial centralization, be it the categorical or binary variable, is practically never statistically significant (never ever when we add all controls) and always quite small. Let's consider the preferred M&P variable, i.e. jurisdictional hierarchy ranging from 1 to 4: its effect is 0.64 percentage points on village electrification, -0.23 pp on household access to electricity, -0.03 standard deviations on the asset-based wealth index, -0.45 pp on

[17] In the case of the light intensity data examined in the previous sections, we argued that the only way to avoid using this potentially "bad control" is to analyze luminosity per capita directly.

school attendance and -0.12 pp on primary school completion.

When working at the country-ethnicity level, giving each area a unit weight, the effect of pre-colonial centralization, although never significant when adding all controls, is bigger (Table 8). This is because not weighting country-ethnicity areas amounts to giving more weight to individual or households located in marginal, sparsely-populated areas, likely to have both lower pre-colonial centralization and worse contemporary development outcomes (if we were controlling perfectly for any confounding factor, the two estimations would yield similar results, but we are not.). It is worth noting that the very small, very sparsely populated country-ethnicity areas, the ones that are driving M&P's results on luminosity at night, are simply not present in our dataset, because they are so small that none of their inhabitants was sampled in the DHS survey.

The same kind of population effects explain why M&P find significant effects of pre-colonial centralization on development indicators when working with the 2005 Afrobarometer survey. Appendix Table A8 shows that the effects of centralization are much reduced and lose statistical significance when each country-ethnicity area is weighted by its 2000 population.

V. Pre-Colonial Centralization, Population Density and Geography

We showed that the link between precolonial centralization and luminosity at night was explained by population effects and that the correlation disappeared when population density was correctly accounted for. When using survey data, even though we did not control for population density directly, we considered individuals, households or villages as the unit of analysis, thus giving less importance to the very sparsely populated regions of Africa. It does not seem that pre-colonial centralization is important in explaining contemporary differences in development in Africa. However, the positive correlation between centralization and population density remains. In their appendix, M&P consider explicitly population density as an outcome of pre-colonial centralization and find significant effects. However, the causal relationship between the two variables goes both ways, making identification difficult. The causal link from population density to pre-colonial institutions has been hypothesized by the "land abundance" literature (Hopkins 1973, Iliffe 1995, Austin 2008, Fenske 2009).

If we import 1960 population density from the same source that M&P used for the 2000 variable, we can check whether changes in population density during the post-colonial era are correlated with pre-colonial centralization. Appendix Tables A9 and A10 rather suggest this is not the case. Whether at the country-ethnicity level or at the pixel level, once population density in 1960 is controlled for, the regression coefficient of centralization in the 2000 population density equation gets small and statistically insignificant.[18] Yet, we acknowledge that regional population figures are coarse estimates, given the scarcity and inaccuracy of underlying population census data, and that noise might dominate their variation across time.

Further, for the post-colonial period as well as for the most distant past, a quantity of geographic factors might determine jointly centralization and density. In the remainder of this section, we show that a slightly modified geographic control

[18]The same is obtained if we rather regress the first difference in log population density over forty years, as indeed the coefficient of lagged population density is close to one in the regression with levels.

set absorbs most of the correlation between centralization and population density.

We modify M&P's control set in the following way: distance to the coast is controlled for with a log specification instead of a linear one (the effect of distance to the coast is indeed likely to be quite strong at first and to weaken as one penetrates inland). Instead of considering only distance to the sea, we define distance to the coast as distance to the nearest sea or lake coast (the importance of the great lakes in determining both population density and precolonial centralization is evidenced in figure 3)[19]. While M&P take water bodies into account through the log area covered by water in the country-ethnicity and a dummy indicating whether there is a river or lake in the pixel, we think that the importance of rivers for trade is better captured through log distance to the nearest river (at the pixel level) and a dummy indicating whether the country-ethnicity is crossed by a river[20]. Finally, we control for land suitability using a different variable, capturing constraints on rain-fed agriculture, as in Fenske (2014)[21].

The regression with modified geographic controls was run at the country-ethnicity level (Table 9, to be compared with M&P Appendix Table 4A) and at the pixel level (Table 10, to be compared with M&P Appendix Table 4B). Logging distance to the coast and considering main lakes as well as sea coasts (columns (2) and (6)) reduces the effect of centralization by between a quarter and a third; it loses statistical significance at the country-ethnicity level, and is only significant at

[19]Important lakes are lakes ranking 0 or 1 in the Natural Earth dataset (`http://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-lakes/`): lakes Albert, Malawi, Tanganyika, Victoria, Chad and Tana.

[20]Our data on rivers comes from `http://www.naturalearthdata.com/../../10m-rivers-lake-centerlines/`.

[21]M&P's land suitability variable, taken from the *Atlas of the Biosphere*, is "the product of two component capturing the climatic and soil suitability for farming." Fenske's variable is "an index of combined climate, soil and terrain slope constrains on rainfed agriculture", taken from the FAO-GAEZ project.

the 10% level at the pixel-level. Using different variables to control for the presence of rivers and land quality (columns (3) and (5)) reduces the effect even more: it is now less than half of what is was in column (1) and loses all significance. The Nile region is very specific in Africa: there, the proximity to the river is more crucial than anywhere else. We might therefore fear that our geographic variables do not capture fully the effect of geography in this region. Indeed, the effect is reduced when the "Egyptian" ethnic group is dropped from the sample (columns (4) and (8)).[22]

It might be that good geographic conditions, favoring trade and agriculture, lead to the creation of centralized institutions which then allowed population to increase in the long-run, although the aforementioned evidence does not seem to support such a causal line for the most recent forty years. It might also be that centralized institutions emerged because population densities were high and coordination problems many. Most likely, political centralization and population density progressed hand in hand over the course of history. Disentangling the elements of a multi-secular causal chain with cross-sectional data is an impossible dream.

[22]Even though the effect of pre-colonial centralization on population density loses significance when we modify only slightly the geographic control set, we might still consider it is quite sizeable in magnitude. Take for instance column (7) of table 10 (pixel-level, new geographic controls, Egyptians included). A coefficient of 0.2 means that population density increases by 20% when the binary variable for political centralization goes from zero to one. However, this is mainly due to the fact that density starts at very low values (0.11 inhabitant per squared kilometer). A coefficient of 0.2 represents only 10% of a standard deviation.

VI. Conclusion

We found little evidence that the political centralization of pre-colonial institutions, as recorded in Murdock's ethnographic atlas, is a strong independent determinant of regional development in Africa. The correlation identified by Michalopoulos and Papaioannou seems all contained in population density differences. We argued that their use of satellite luminosity data does not properly take into account the interaction between population density (or urbanization) and economic development, and proposed a variety of estimates that do a better job at this. Their use of Murdock's map of ethnic groups amounts to delineating many sparsely populated areas where luminosity is dominated by noise. Furthermore representative survey data abstracting from the luminosity-population nexus still provide little support to the explanatory power of pre-colonial institutions. Even the causal channel going from political centralization to population density is questionable. Centralization does not seem to impact regional population growth between 1960 and 2000, and we show it is plausible that geography accounts for its correlation with present population density.

Let us tentatively offer lessons to be drawn from this replication exercise. First, luminosity data, and more generally satellite imagery, are certainly a promising source for improving our knowledge of the patterns of economic development. However, great caution should be granted to measurement issues. Second, Murdock's data also make a unique piece for comparative works, if one does not ask too much of them; in particular they do not provide a precise mapping of pre-colonial polities before the Scramble for Africa. Third, cross-sectional correlations carry strong limitations, especially on heterogeneous samples with a large spatial extension (continent wide here). The conditional independence assumptions that are required for the quasi-experimental inference of causality are unlikely to hold, and hence always remain under threat. The exploration of historical causality is better obtained with genuine historical data, where individual trajectories can be observed. Fourth, Africa is a continent in the process of populating, hence demographic forces, in particular migrations and urbanization, are very much active in long-term economic development; if only as potential confounding factors, they should be seriously taken into account in any historical narrative.

Table 1: GDP per capita and Luminosity in Cross-Sections of Countries

				Population-Weighted		
	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A: World 2008			
Log Light Intensity	0.4974***			0.5027***		
	(0.0337)			(0.0496)		
Log Luminosity per capita		0.7035***	0.7280***		0.7953***	0.8154***
		(0.0298)	(0.0277)		(0.0257)	(0.0274)
Log Population Density			0.1777***			0.0606**
			(0.0318)			(0.0306)
Adjusted R-squared	0.573	0.774	0.810	0.386	0.855	0.858
Observations	163	163	163	163	163	163
			Panel B: Africa 2008			
Log Light Intensity	0.4183***			0.4606***		
	(0.0782)			(0.0578)		
Log Luminosity per capita		0.6729***	0.6853***		0.6136***	0.6194***
		(0.0551)	(0.0578)		(0.0428)	(0.0434)
Log Population Density			0.0480			0.0553
			(0.0650)			(0.0643)
Adjusted R-squared	0.360	0.751	0.749	0.560	0.807	0.806
Observations	50	50	50	50	50	50

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on data from Henderson, Storeygard and Weil (2012).
Notes: The dependent variable is the logarithm of GDP per capita in constant 2005 international dollars, for the year 2008. OLS regressions. Ln Luminosity is the logarithm of the area-weighted average lights digital number per cell. Ln Luminosity per capita is Ln Luminosity minus Ln Population Density, all measured for the year 2008. Population-weighted are weighted OLS estimates with individual weight equal to country's population.

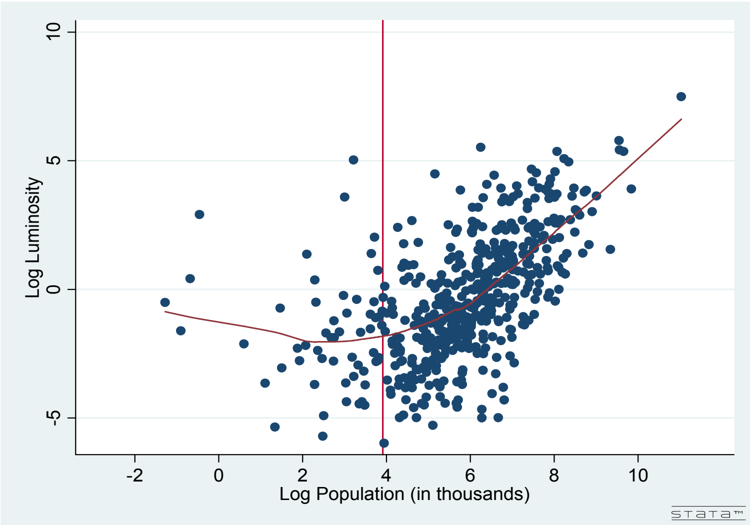
Table 2: Luminosity per Capita as an Alternative Outcome

	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Colonial Ethnic Institutions and Luminosity per capita Within African Countries						
Focusing on the Intensive Margin of Luminosity						
Jurisdictional Hierarchy	0.1109			0.0324		
	(0.0715)			(0.0751)		
Binary Political Centralization		0.1879*			0.0466	
		(0.1037)			(0.1345)	
Petty Chiefdoms			0.0640			-0.0926
			(0.2044)			(0.3595)
Paramount Chiefdoms			0.1578			-0.0685
			(0.2118)			(0.2250)
Precolonial States			0.4059			0.0909
			(0.2635)			(0.3284)
Adjusted R-squared	0.553	0.553	0.554	0.688	0.688	0.689
Observations	512	512	512	512	512	512
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	No	No	No	No
Geographic Controls	No	No	No	No	No	No
Population Density	No	No	No	No	No	No
Population Weights	No	No	No	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

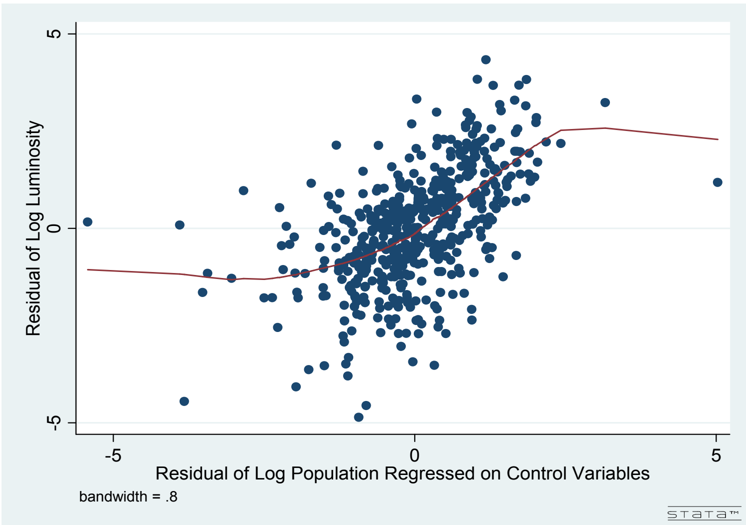
Source: Authors' analysis based on M&P data.
Notes: Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. OLS estimates. Dependent variable is ln(luminosity per capita), i.e. log(light density at night from satellite), the variable used in M&P, minus ln(population density). In M&P sample, 5 ethnic-country pairs (out of 517) have null population density and hence are not used here.

Figure 1: Non-Parametric Regression of Luminosity on Population



Note: Locally weighted (tricube) regression of log luminosity on log population. Luminosity is the total light intensity of each area, i.e. light intensity times land area. The vertical bar corresponds to population equal to 50,000 inhabitants (ln(50)=3.912).

Figure 2: Non-Parametric Regression of Luminosity on Population, with Controls



Note: Log luminosity and log population were first separately regressed on the location and geographic controls used in M&P OLS regressions, including land area. The graph then depicts the locally weighted (tricube) regression of the residual of the log luminosity regression on the residual of the log population regression.

Table 3: Accounting for Low-Population Areas: Country-Ethnicity Level

	M&P	Spline Pop.	M&P	$> 10,000$	$> 50,000$	Pop.Weights
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-Colonial Ethnic Institutions and Luminosity per capita Within African Countries						
Panel A: With ln(0.01+Light Density) as Outcome						
Jurisdictional Hierarchy	0.1588***	0.0715	0.1766***	0.0768	0.0433	-0.0508
	(0.0491)	(0.0463)	(0.501)	(0.0493)	(0.0487)	(0.0609)
Log Population	0.5343***					
	(0.0701)					
Log Population ($< 50,000$)		0.2026**				
		(0.0795)				
Log Population ($> 50,000$)		0.7865***				
		(0.0873)				
Log Population Density[b]			0.4375***	0.7662***	0.9325***	1.0191***
			(0.0622)	(0.0677)	(0.0875)	(0.0949)
Log Land Area	-0.3822***	-0.3926***	0.1293***	0.2533***	0.3318***	0.2408***
	(0.0775)	(0.0785)	(0.0451)	(0.0559)	(0.0653)	(0.0731)
Adjusted R-squared	0.678	0.712	0.661	0.719	0.750	0.898
Observations	666[a]	666[a]	682	610	523	666[a]
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
Population weights	No	No	No	No	No	Yes
Pre-Colonial Ethnic Institutions and Luminosity per capita Within African Countries						
Panel B: Focusing on the Intensive Margin of Luminosity						
Jurisdictional Hierarchy	0.1374*	0.0609	0.1493**	0.0790	0.0413	-0.0319
	(0.0713)	(0.0778)	(0.0727)	(0.0799)	(0.0792)	(0.0698)
Log Population	0.7434***					
	(0.1383)					
Log Population ($< 50,000$)		-0.0573				
		(0.1727)				
Log Population ($> 50,000$)		1.0287***				
		(0.1354)				
Log Population Density[b]			0.6820***	0.9397***	1.1357***	1.1812***
			(0.1355)	(0.1281)	(0.1426)	(0.1321)
Log Land Area	-0.8118***	-0.8759***	-0.0728	0.0714	0.2383***	0.1743*
	(0.1166)	(0.1127)	(0.0935)	(0.0864)	(0.0926)	(0.0900)
Adjusted R-squared	0.679	0.713	0.671	0.699	0.713	0.877
Observations	512[a]	512[a]	517	497	455	512[a]
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
Population weights	No	No	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. OLS estimates. Dependent variable is log(0.01 + light density at night from satellite), as in M&P Table III. Spline Pop.= Spline function of log population, with a kink-point located at 50,000 people in the ethnic-country pair (i.e. log population = 3.91).
a: 16 (out of 682) ethnic-country pairs have zero population in M&P sample. Among these 16, 11 have zero luminosity.
b: Same variable as in M&P, i.e. ln(0.01 + population density).

Table 4: Accounting for Low-Population Areas: Pixel-Level

	Lit/Unlit Pixels			ln(0.01+Light Density)		
	M&P	$>4/km^2$	Pop.Weights	M&P	$>4/km^2$	Pop.Weights
	(1)	(2)	(3)	(4)	(5)	(6)
Jurisdictional Hierarchy	0.0265***	0.0121	0.0029	0.1559***	0.0824	-0.0201
	(0.0073)	(0.0093)	(0.0123)	(0.0481)	(0.0582)	(0.0818)
Log Population Density	0.0657***	0.1159***	0.1406***	0.3287***	0.6121***	1.1074***
	(0.0098)	(0.0115)	(0.0162)	(0.0657)	(0.0787)	(0.1027)
Log Land Area	0.0473***	0.0575***	0.1120***	0.1021**	0.1277**	0.2231*
	(0.0108)	(0.0126)	(0.0183)	(0.0409)	(0.0499)	(0.1242)
Ajdusted R-squared	0.379	0.422	0.527	0.456	0.528	0.743
Observations	66173	45762	66173	66173	45762	66173
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Pixel Level	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Ethnic-Country Level	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes
Population weights	No	No	Yes	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. OLS estimates. Same outcomes as in M&P Table V, pp.136-137. Col. (1) and (4) respectively reproduce col. (5) and col. (10) of Table V p.136 in M&P.

Table 5: Pixel-Level Models for Luminosity in Discrete Form (Lit/Unlit)

	LPM	Logit		Probit			Lewbel
		Simple	RE	Simple	RE	Het.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Jurisdictional Hierarchy							
Average Marginal effect	0.0265***	0.0049	0.0003	0.0080	0.0016	0.0044	0.0121
	(0.0073)	(0.0060)^a	(0.0019)^a	(0.0057)^a	(0.0019)^a	(0.0039)^a	(0.0104)
Ratio to Coeff. Pop. Density	0.4041***	0.0702	0.0042	0.1181	0.0247	0.0503	0.1687
	(0.1200)^a	(0.0885)^a	(0.0279)^a	(0.0884)^a	(0.0289)^a	(0.0451)^a	(0.1893)
Predicted Probabilities out of [0;1]	14656	0	0	0	0	0	0
Observations	66173	66173	66173	65843	66173	66173	66173
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Pixel Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Ethnic-Country Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population weights	No	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Dependent variable is the Lit/Unlit dummy. Except for Logit RE, Probit RE & Probit Het., double-clustered standard errors at the country and the ethno-linguistic family level. Double-clustered estimates are obtained by computing three estimates, the first one clustered at the country level, the second at the ethno-linguistic family level, the third at the intersection of both. We then compute $Var(\hat{\beta}) = Var_1(\hat{\beta}) + Var_2(\hat{\beta}) - Var_{1\cap 2}(\hat{\beta})$, as in Cameron, Gelbach and Miller (2011). When non-positive semi-definite, the matrix is transformed by replacing all negative eigenvalues with zero (this transformation is actually innocuous for the coefficient of interest).
LPM: M&P estimates; column (1) reproduces column (5) in Table V p. 136.
Logit & Probit Simple, (2) & (4): Clustered standard errors are computed as a sandwich estimator on the scores, assuming no bias on pseudo-likelihood point estimates. We also tried fractional logit or probit (Paapke & Woolridge 1996) on pixel-level data averaged at the ethnic-country pairs level, and obtained very close estimates; for the average marginal effects: 0.0058 (s.e. 0.069) with the logit link, and 0.0067 (s.e. 0.0064) with the probit link.
Logit & Probit RE, (3) & (5): Gaussian random effects for ethno-linguistic family.
Probit Het. (6): Multiplicative heteroskedasticity as a function of population density; clustered standard errors at the ethno-linguistic family level are computed as a sandwich estimator on the scores, assuming no bias on pseudo-likelihood point estimates.
Lewbel estimator, (7): See Lewbel (2000). Population density is used as a special regressor which is exogenous, i.e. independent from the unobservable determinants of luminosity, and whose large support covers all the distribution of these unobservables (i.e. light density runs from zero in desert places to capped values in very densely populated areas like capital cities). The inverse conditional density of population density conditional on controls is estimated using the Lewbel and Schennach (2007) non-parametric ordered data estimator. Then, the average marginal effect is computed as $\hat{\beta}_1.M'(X\hat{\beta})$, where $\hat{\beta}_1$ is the estimated coefficient of jurisdictional hierarchy (the coefficient of population density, the special regressor, being normalized to 1), and where M' is the first derivative of an Epanechikov kernel regression of the dependent variable D (Lit/Unlit) on the index variable $X\hat{\beta}$. Clustered bootstrap standard errors (50 replications). The 0.1% highest absolute values for $[D - I(V > 0)]/f(V|X)$, the transformed dependent variable, are trimmed, V being the demeaned special regressor (log population density).
a: The delta method was used.

Table 6: Censored Data Models for Luminosity in Continuous Form at Pixel-Level

	Luminosity per km^2					Lum. per capita
	OLS	Tobit		Lewbel		Heckman
		Simple	RE	Quadratic	Ordered	
	(1)	(2)	(3)	(4)	(5)	(6)
Jurisdictional Hierarchy	0.1559***	0.1143	-0.0468	-0.0387	0.0420	-0.0538
	(0.0481)	(0.1422)	(0.0467)	(0.0713)	(0.0834)	(0.0847)
Observations	66173	66173	66173	66173	66173	66173
Uncensored Observations	66173	11028	11028	11028	11028	11028
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Pixel Level	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Ethnic-Country Level	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	(Yes)
Population weights	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Except for Tobit RE, double-clustered standard errors at the country and the ethno-linguistic family level. Double-clustered estimates are obtained by computing three estimates, the first one clustered at the country level, the second at the ethno-linguistic family level, the third at the intersection of both. We then compute $Var(\hat{\beta}) = Var_1(\hat{\beta}) + Var_2(\hat{\beta}) - Var_{1\cap 2}(\hat{\beta})$, as in Cameron, Gelbach and Miller (2011). When non-positive semi-definite, the matrix is transformed by replacing all negative eigenvalues with zero (this transformation is actually innocuous for the coefficient of interest).
OLS: M&P estimates with ln(0.01+Luminosity): columns (1) reproduces columns (10) in Table V p. 136.
Tobit: Dependent variable is ln(Luminosity per km^2), zero luminosity is treated as lying below the observed minimum. Simple: Clustered standard errors are computed as a sandwich estimator on the scores, assuming no bias on pseudo-likelihood point estimates. RE: Gaussian random effects for ethno-linguistic family.
Lewbel: See Lewbel (2007). Dependent variable is ln(Luminosity per km^2), zero luminosity is treated as unobserved luminosity, and population density is used as a special regressor which is exogenous, i.e. independent from the unobservable determinants of luminosity, and whose large support covers all the distribution of these unobservables (i.e. luminosity runs from zero in desert places to capped values in very densely populated areas like capital cities). The estimate of the coefficient of jurisdictional hierarchy stems from a weighted OLS regression, with weights $W = 1/f(V|X)$, where $f(V|X)$ is the density of log. population density (V) conditional to the regressors (X) considered by M&P. W is estimated on the whole sample of unlit and lit pixels, using either a quadratic Epanechikov kernel or the ordered data estimator from Lewbel and Schennach (2007). Clustered bootstrap standard errors (50 replications). The 0.1% highest values for the inverse density of log population density (conditional to the other explanatory variables) are trimmed, i.e. only 99.9% of uncensored observations are used.
Heckman: Heckman (1979) model of selection. Unobserved luminosity per capita is modeled as a function of jurisdictional hierarchy and of all controls including ln. population density. This latter variable is excluded from the continuous part (assuming population density has no impact on luminosity *per capita*). Clustered standard errors are computed as a sandwich estimator on the scores, assuming no bias on pseudo-likelihood point estimates.

Table 7: Results Using the DHS Dataset: Linear Probability Model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: village connection to electricity							
Jur. Hiererarchy	0.1341*	0.0152	0.0163	0.0064				
Double-clustered p-val.	(0.0857)	(0.4787)	(0.3852)	(0.7639)				
0-1 pol. centralization					0.2250	0.0229	0.0427*	0.0364
Double-clustered p-val.					(0.1687)	(0.4314)	(0.0885)	(0.2501)
Observations	10,291	10,291	10,160	10,160	10,291	10,291	10,160	10,160
	Dependent variable: household connection to electricity							
Jur. Hiererarchy	0.0952	0.0054	0.0067	-0.0023				
Double-clustered p-val.	(0.2054)	(0.7836)	(0.6825)	(0.9037)				
0-1 pol. centralization					0.1454	0.0064	0.0258	0.0193
Double-clustered p-val.					(0.2927)	(0.8049)	(0.2065)	(0.4795)
Observations	244,808	244,808	241,874	241,874	244,808	244,808	241,874	241,874
	Dependent variable: asset-based wealth index							
Jur. Hiererarchy	0.2963	0.0082	0.0102	-0.0032				
Double-clustered p-val.	(0.1877)	(0.8602)	(0.7987)	(0.9449)				
0-1 pol. centralization					0.4420	0.0083	0.0512	0.0473
Double-clustered p-val.					(0.2859)	(0.8929)	(0.3233)	(0.4592)
Observations	241,096	241,096	238,195	238,195	241,096	241,096	238,195	238,195
	Dependent variable: individual ever went to school							
Jur. Hiererarchy	-0.0192	-0.0019	-0.0016	-0.0045				
Double-clustered p-val.	(0.5178)	(0.9166)	(0.9297)	(0.6998)				
0-1 pol. centralization					-0.0608	-0.0280	-0.0132	-0.0087
Double-clustered p-val.					(0.3649)	(0.5068)	(0.6762)	(0.6762)
Observations	236,220	236,220	232,709	232,709	236,220	236,220	232,709	232,709
	Dependent variable: individual completed primary							
Jur. Hiererarchy	-0.0185	0.0020	0.0044	-0.0012				
Double-clustered p-val.	(0.5679)	(0.8905)	(0.7033)	(0.9087)				
0-1 pol. centralization					-0.0467	-0.0167	0.0015	0.0032
Double-clustered p-val.					(0.4863)	(0.6676)	(0.9492)	(0.8211)
Observations	236,220	236,220	232,709	232,709	236,220	236,220	232,709	232,709
Country F.E.	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Pixel controls	No	No	Yes	Yes	No	No	Yes	Yes
Country-ethnicity controls	No	No	No	Yes	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis on merged DHS surveys from 33 countries (only 32 for education variables). For each country, most recent geocoded survey among phases 3 to 6 (no geocoded surveys in phases 1 and 2). Each cluster was allocated to its country-ethnicity area and to the nearest M&P pixel.
Notes: Village connection to electricity equal to one if more than 50% of households in the cluster have electricity, zero otherwise. Asset-based wealth index built by the authors using principal component analysis (DHS wealth index not suitable for cross-country comparisons). OLS estimates. DHS sampling weights used in all regressions, transformed to take into account differences in population sizes across countries. Controls are the same as in M&P. For schooling variables, gender of the individual added as a control in all regressions. Within parentheses, double-clustered p-values at the country and the ethno-linguistic family level.

Table 8: DHS Dataset: Results on Unweighted Country-Ethnicity Areas

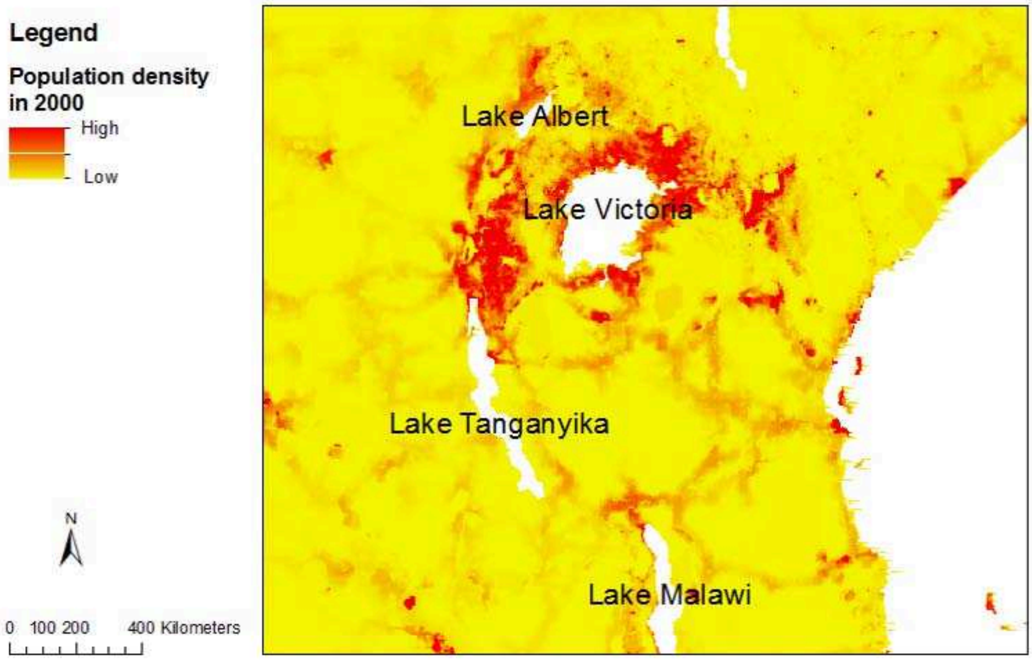
	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: % of villages connected to electricity					
Jur. Hiererarchy	0.0133	0.0390**	0.0190			
Double-clustered p-val.	(0.5090)	(0.0495)	(0.2979)			
0-1 pol. centralization				0.0236	0.0700*	0.0455
Double-clustered p-val.				(0.4854)	(0.0601)	(0.1178)
Observations	374	374	374	374	374	374
	Dependent variable: % of households having electricity					
Jur. Hiererarchy	0.0072	0.0290*	0.0138			
Double-clustered p-val.	(0.7091)	(0.0863)	(0.3955)			
0-1 pol. centralization				0.0081	0.0496	0.0319
Double-clustered p-val.				(0.8017)	(0.1060)	(0.1991)
Observations	374	374	374	374	374	374
	Dependent variable: average asset-based wealth index					
Jur. Hiererarchy	0.0377	0.0548*	0.0219			
Double-clustered p-val.	(0.4106)	(0.0599)	(0.4106)			
0-1 pol. centralization				0.0495	0.0977	0.0594
Double-clustered p-val.				(0.5140)	(0.1148)	(0.1673)
Observations	374	374	374	374	374	374
	Dependent variable: % of individuals who went to school					
Jur. Hiererarchy	0.0057	0.0135*	-0.0020			
Double-clustered p-val.	(0.7801)	(0.0590)	(0.8366)			
0-1 pol. centralization				0.0117	0.0118	-0.0033
Double-clustered p-val.				(0.7838)	(0.5988)	(0.8956)
Observations	371	371	371	371	371	371
	Dependent variable: % of primary completion					
Jur. Hiererarchy	-0.0042	0.0130	-0.0029			
Double-clustered p-val.	(0.7889)	(0.1354)	(0.7194)			
0-1 pol. centralization				-0.0081	0.0068	-0.0080
Double-clustered p-val.				(0.8107)	(0.7944)	(0.6910)
Observations	371	371	371	371	371	371
Country F.E.	No	Yes	Yes	No	Yes	Yes
Country-ethnicity controls	No	No	Yes	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis on merged DHS surveys from 33 countries (only 32 for education variables). For each country, most recent geocoded survey among phases 3 to 6 (no geocoded surveys in phases 1 and 2). Each cluster was allocated to its country-ethnicity area.
Notes: Villages connected to electricity are villages where more than 50% of households have electricity. Asset-based wealth index built by the authors using principal component analysis (DHS wealth index not suitable for cross-country comparisons). Country-ethnicity means computed using DHS weights (transformed to take into account differences in population sizes across countries). Unweighted OLS estimates. Controls are the same as in M&P. Within parentheses, double-clustered p-values at the country and the ethno-linguistic family level.

Figure 3: Population Density and Pre-Colonial Centralization in the Great Lakes Region

(a) Population density in 2000



(b) Precolonial centralization

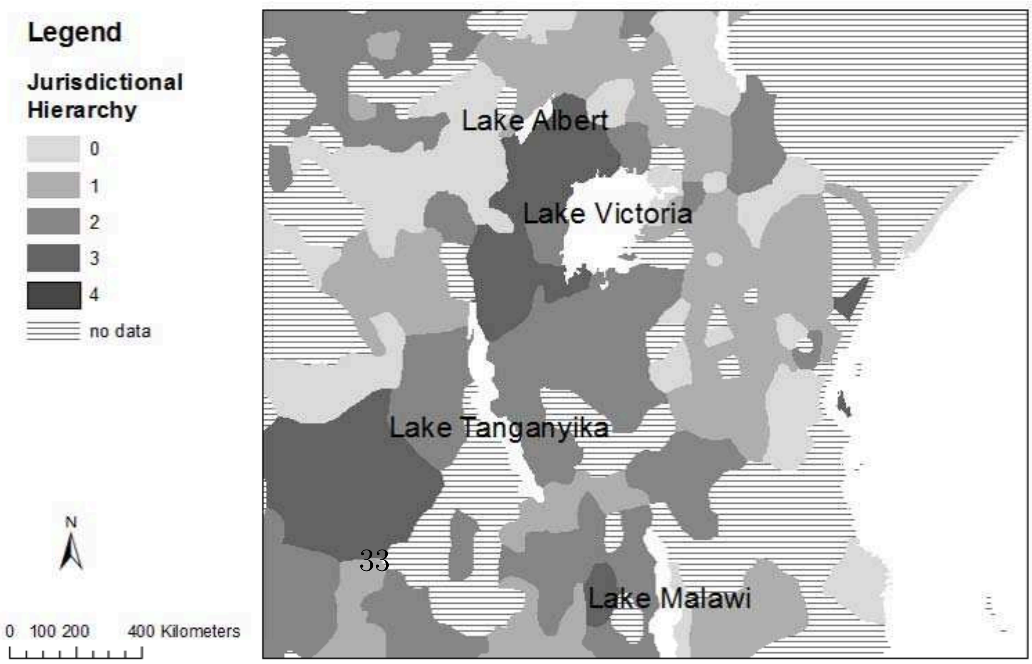


Table 9: Centralization and Log Country-Ethnicity Population Density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jur. Hiererarchy	0.2348**	0.1536	0.1064	0.0904				
Double-clustered p-val.	(0.0329)	(0.1246)	(0.2847)	(0.3388)				
0-1 pol. centralization					0.4461**	0.2864	0.1829	0.1536
Double-clustered p-val.					(0.0424)	(0.1587)	(0.3384)	(0.4036)
Observations	683	683	683	682	683	683	683	682
Country F.E.	Yes							
Geographic controls	Yes							
Log dist. to sea and lake coast	No	Yes	Yes	Yes	No	Yes	Yes	Yes
New river and land quality variables	No	No	Yes	Yes	No	No	Yes	Yes
Egyptians	Yes	Yes	Yes	No	Yes	Yes	Yes	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Specifications (1) and (5) are specifications (4) and (8) in M&P Appendix Table 4A. In other specifications, the control set is modified. (2) and (6): log distance to the nearest sea or lake coast instead of distance to the sea coast. (3) and (7): dummy indicating whether the country-ethnicity area is crossed by a river rather than log area covered by water; constraint on rain-fed agriculture taken from Fenske (2014) rather than M&P land suitability variable. (4) and (8): exclusion of the "Egyptian" ethnic group. OLS estimates. Within parentheses, double-clustered p-values at the country and the ethno-linguistic family level.

Table 10: Centralization and Log Pixel Population Density

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jur. Hiererarchy	0.2105**	0.1289*	0.0954	0.0675				
Double-clustered p-val.	(0.0100)	(0.0740)	(0.1575)	(0.2636)				
0-1 pol. centralization					0.3985**	0.2630*	0.2009	0.1414
Double-clustered p-val.					(0.0152)	(0.0761)	(0.1326)	(0.2259)
Observations	66173	66173	66428	65937	66173	66173	66428	65937
Country F.E.	Yes							
Pixel and split controls	Yes							
Log dist. to sea and lake coast	No	Yes	Yes	Yes	No	Yes	Yes	Yes
New river and land quality variables	No	No	Yes	Yes	No	No	Yes	Yes
Egyptians	Yes	Yes	Yes	No	Yes	Yes	Yes	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Specifications (1) and (5) are specifications (4) and (8) in M&P Appendix Table 4B. In other specifications, the control set is modified. (2) and (6): log distance to the nearest sea or lake coast instead of distance to the sea coast. (3) and (7): dummy indicating whether the country-ethnicity area is crossed by a river rather than log area covered by water (country-ethnicity control set); log distance to the nearest river rather than log pixel area covered by water (pixel control set); constraint on rain-fed agriculture taken from Fenske (2014) rather than M&P land suitability variable. (4) and (8): exclusion of the "Egyptian" ethnic group. OLS estimates. Within parentheses, double-clustered p-values at the country and the ethno-linguistic family level.

References

Angrist, J. D., and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Austin, G. (2008): "Resources, techniques, and strategies south of the Sahara: revising the factor endowments perspective on African economic development, 1500-2000," *Economic History Review*, 61(3), 587–624.

Barbour, K. M. (1961): "A Geographical Analysis of Boundaries in Inter-Tropical Africa," in *Essays on African Population*, ed. by K. M. Barbour, and R. Prothero, pp. 303–23. Routledge and Kegan Paul, London.

Brownlie, I. (1979): *African Boundaries - A Legal and Diplomatic Encyclopaedia*. C. Hurst and Company, London.

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

Chen, X., and W. D. Nordhaus (2010): "The Value of Luminosity Data as a Proxy for Economic Statistics," NBER Working Papers 16317, National Bureau of Economic Research, Inc.

——— (2011): "Using luminosity data as a proxy for economic statistics," *Proceeedings of the National Academy of Sciences (US)*, 108(21), 8589–8594.

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

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

Fenske, J. (2009): "Does Land Abundance Explain African Institutions?," Working Papers 74, Yale University, Department of Economics.

——— (2014): "Ecology, trade and States in pre-colonial Africa," *Journal of the European Economic Association*.

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

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–61.

Henderson, J. V., 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: Comparative Lessons in Authority and Control*, Princeton Studies in International History and Politics. Princeton University Press.

Hopkins, A. G. (1973): *An economic history of West Africa*, The Columbia economic history of the modern world. Longman.

Iliffe, J. (1995): *Africans: The History of a Continent*, African studies series. Cambridge University Press.

LEWBEL, A. (2000): "Semiparametric Qualitative Response Model Estimation With Unknown Heteroscedasticity or Instrumental Variables," *Journal of Econometrics*, 97, 145–177.

——— (2007): "Endogenous Selection or Treatment Model Estimation," *Journal of Econometrics*, 141, 777–806.

LEWBEL, A., AND S. SCHENNACH (2007): "A Simple Ordered Data Estimator For Inverse Density Weighted Functions," *Journal of Econometrics*, 186, 189–211.

MICHALOPOULOS, S., AND E. PAPAIOANNOU (2012): "National Institutions and African Development: Evidence from Partitioned Ethnicities," Working Papers 18275, National Bureau of Economic Research, Cambridge (MA).

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

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

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

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

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

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

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

Papke, L. E., and J. M. Wooldridge (2013): "Econometric methods for fractional response variables with an application to 401(K) plan participation rates," *Journal of Applied Econometrics*, 11, 619–632.

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

Smits, J., and R. Steendijk (2012): "The International Wealth Index," Working Paper 12-107, Nijmegen Center for Economics (NiCE).

Appendices

A Additional results on luminosity at night

Table A1: Population-Weighted Country-Ethnicity Areas

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Pre-Colonial Ethnic Institutions and Luminosity per km^2 Within African Countries									
All Observations with Positive Population									
Jurisdictional Hierarchy	0.3079***	0.1087	0.1162	0.1932***	0.0264	0.0467	0.1588***	-0.0508	0.0040
	(0.0789)	(0.1295)	(0.1199)	(0.0517)	(0.0622)	(0.0627)	(0.0488)	(0.0610)	(0.0628)
Adjusted R-squared	0.425	0.722	0.558	0.615	0.863	0.764	0.680	0.898	0.828
Observations	666	666	666	666	666	666	666	666	666
Panel B: Pre-Colonial Ethnic Institutions and Luminosity per km^2 Within African Countries									
Focusing on the Intensive Margin of Luminosity									
Jurisdictional Hierarchy	0.3098***	0.0989	0.0945	0.1510**	0.0178	0.0230	0.1374*	-0.0319	0.0052
	(0.1108)	(0.1521)	(0.1520)	(0.0695)	(0.0778)	(0.0813)	(0.0703)	(0.0697)	(0.0799)
Adjusted R-squared	0.437	0.690	0.528	0.640	0.846	0.752	0.681	0.877	0.803
Observations	513	513	513	513	513	513	513	513	513
Population Weights	No	Raw	Within	No	Raw	Within	No	Raw	Within
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Controls	No	No	No	No	No	No	Yes	Yes	Yes
Geographic Controls	No	No	No	No	No	No	Yes	Yes	Yes
Population Density	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.

Notes: Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. OLS estimates. Panel A: Dependent variable is ln(0.01 + Luminosity per km^2). Panel B: Dependent variable is ln(Luminosity per km^2). Population Weights: Raw = Total population circa 2000; Within = Ratio of population to mean population in each country (sum of weights for each country equal to number of ethnic groups). In M&P sample, 16 ethnic-country pairs (among 682) have null population density and hence are not used here.

Table A2: Population-Weighted Pixel-Level Estimates

	Lit/Unlit pixels				ln(0.01 + Luminosity)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jurisdictional Hierarchy	0.0447**	0.0227	0.0265***	0.0029	0.2362**	0.1420	0.1559***	-0.0201
	(0.0177)	(0.0228)	(0.0073)	(0.0123)	(0.1036)	(0.1840)	(0.0481)	(0.0818)
Adjusted R-squared	0.272	0.355	0.379	0.527	0.320	0.511	0.456	0.743
Predicted Probabilities out of [0;1]	2638	166	14656	28101	-	-	-	-
Observations	66570	66570	66173	66173	66570	66570	66173	66173
Population weights	No	Yes	No	Yes	No	Yes	No	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	No	Yes	Yes	No	No	Yes	Yes
Controls at the Pixel Level	No	No	Yes	Yes	No	No	Yes	Yes
Controls at the Ethnic-Country Level	No	No	Yes	Yes	No	No	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. OLS estimates. Dependent variables are the same as in M&P Table V. Columns (1), (3), (5) and (7) respectively reproduce estimates of columns (2), (5), (7) and (10) in Table V, p.136 (except for the slight correction of standard errors explained in the text). Population Weights: Total population circa 2000.

Table A3: Contiguous Ethnic Homelands: Population-Weighted Pixel-Level Estimates

	All Observations			Difference in Jurisdictional Hierarchy Index > |1|			One Ethnic Group Was Part of a Pre-Colonial State		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Jurisdictional Hierarchy	0.0052	-0.0061	-0.0039	0.0054	-0.0029	0.0029	0.0143	0.0060	0.0153
	(0.0115)	(0.0066)	(0.0083)	(0.0142)	(0.0090)	(0.0101)	(0.0176)	(0.0136)	(0.0126)
Adjusted R-squared	0.450	0.525	0.534	0.497	0.555	0.565	0.608	0.641	0.653
Observations	78139	78139	77833	34180	34180	34030	16570	16570	16474
N Countries	38	38	38	32	32	32	21	21	21
N Ethnic Groups	260	260	260	154	154	154	72	72	72
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Controls at the Pixel Level	No	No	Yes	No	No	Yes	No	No	Yes
Controls at the Ethnic-Country Level	No	No	Yes	No	No	Yes	No	No	Yes
Population weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Estimates of this Table reproduce M&P estimates of Table VII p.142, with the only exception that each pixel is given its population weights. Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. OLS estimates. Population Weights: Total population circa 2000, from the data used in M&P.

Table A4: Contiguous Ethnic Homelands: Logit Models for Pixel-Level Luminosity (Lit/Unlit)

	All Observations			Difference in Jurisdictional Hierarchy Index $> \|1\|$			One Ethnic Group Was Part of a Pre-Colonial State		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Jurisdictional Hierarchy									
Average Marginal effect	0.0240**	0.0036	0.0042	0.0289**	0.0044	0.0049	0.0408***	0.0105	0.0092
	$(0.0101)^a$	$(0.0054)^a$	$(0.0049)^a$	$(0.0131)^a$	$(0.0067)^a$	$(0.0059)^a$	$(0.0152)^a$	$(0.0090)^a$	$(0.0085)^a$
Ratio to Coeff. Pop. Density	-	0.0490	0.0577	-	0.0538	0.0609	-	0.1415	0.1309
	-	$(0.0749)^a$	$(0.0704)^a$	-	$(0.0847)^a$	$(0.0773)^a$	-	$(0.1310)^a$	$(0.0979)^a$
Observations	78139	78139	77833	34180	34180	34030	16570	16570	16474
N Countries	38	38	38	32	32	32	21	21	21
N Ethnic Groups	260	260	260	154	154	154	72	72	72
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Controls at the Pixel Level	No	No	Yes	No	No	Yes	No	No	Yes
Controls at the Ethnic-Country Level	No	No	Yes	No	No	Yes	No	No	Yes
Population weights	No	No	No	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.

Notes: Estimates of this Table reproduce M&P estimates of Table VII p.142, with the only exception that a logit specification is used instead of a linear probability model. Dependent variable is the Lit/Unlit dummy. Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. Double-clustered estimates are obtained by computing three estimates, the first one clustered at the country level, the second at the ethno-linguistic family level, the third at the intersection of both. We then compute $Var(\hat{\beta}) = Var_1(\hat{\beta}) + Var_2(\hat{\beta}) - Var_{1\cap 2}(\hat{\beta})$, as in Cameron, Gelbach and Miller (2011). When non-positive semi-definite, the matrix is transformed by replacing all negative eigenvalues with zero (this transformation is actually innocuous for the coefficient of interest). Clustered standard errors are computed as a sandwich estimator on the scores, assuming no bias on pseudo-likelihood point estimates.

a: The delta method was used.

Table A5: Adjacent Ethnic Homelands Close to the Ethnic Border: Population-Weighted Estimates

	All Observations			Difference in Jurisdictional Hierarchy Index > |1|			One Ethnic Group Was Part of a Pre-Colonial State		
	<100km	<150km	<200km	<100km	<150km	<200km	<100km	<150km	<200km
	of ethnic border			of ethnic border			of ethnic border		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pre-Colonial Ethnic Institutions and Regional Development Within Contiguous Ethnic Homelands in the Same Country									
Panel A: Pixel-Level Analysis in Areas Close to the "Thick" Ethnic Border									
Panel 1: Border Thickness-Total 50 km (25 km from each side of the ethnic boundary)									
Jurisdictional Hierarchy	-0.0130	-0.0165	-0.0167	0.0002	-0.0043	-0.0060	-0.0076	-0.0026	0.0012
	(0.0116)	(0.0123)	(0.0165)	(0.0154)	(0.0149)	(0.0205)	(0.0249)	(0.0206)	(0.0239)
	-1.11	-1.34	-1.01	0.01	-0.29	-0.29	-0.30	-0.13	0.05
Adjusted R-squared	0.527	0.514	0.519	0.511	0.525	0.542	0.570	0.599	0.613
Observations	6830	10451	13195	3700	5421	6853	2347	3497	4430
N Countries	31	31	31	23	23	23	15	15	15
N Ethnic Groups	120	121	121	61	62	62	31	32	32
Panel 2: Border Thickness-Total 100 km (50 km from each side of the ethnic boundary)									
Jurisdictional Hierarchy	-0.0038	-0.0143*	-0.0144	0.0060	-0.0020	0.0000	-0.0021	0.0057	0.0152
	(0.0147)	(0.0082)	(0.0131)	(0.0147)	(0.0096)	(0.0174)	(0.0178)	(0.0154)	(0.0167)
	-0.26	-1.75	-1.10	0.41	-0.21	0.00	-0.12	0.37	0.91
Adjusted R-squared	0.551	0.527	0.534	0.552	0.557	0.582	0.623	0.651	0.677
Observations	4460	8081	10825	2438	4159	5591	1538	2688	3621
N Countries	31	31	31	23	23	23	15	15	15
N Ethnic Groups	115	116	116	59	60	60	31	32	32
Panel B: Same, plus Controlling for a Fourth-order RD-Type Polynomial in Distance to the Ethnic Border									
Panel 1: Border Thickness-Total 50 km (25 km from each side of the ethnic boundary)									
Jurisdictional Hierarchy	0.0248	0.0142	0.0199	-0.0124	-0.0112	0.0060	0.0060	-0.0032	0.0318
	(0.0216)	(0.0205)	(0.0254)	(0.0325)	(0.0357)	(0.0492)	(0.0315)	(0.0419)	(0.0437)
	1.15	0.69	0.78	-0.38	-0.31	0.12	0.19	-0.08	0.73
Adjusted R-squared	0.530	0.516	0.523	0.517	0.527	0.546	0.576	0.606	0.624
Observations	6830	10451	13195	3700	5421	6853	2347	3497	4430
Panel 2: Border Thickness-Total 100 km (50 km from each side of the ethnic boundary)									
Jurisdictional Hierarchy	0.0294	0.0004	0.0051	-0.0345	-0.0361	-0.0109	-0.0341	-0.0198	-0.0091
	(0.0287)	(0.0181)	(0.0199)	(0.0582)	(0.0439)	(0.0371)	(0.0696)	(0.0364)	(0.0355)
	1.02	0.02	0.25	-0.59	-0.82	-0.29	-0.49	-0.54	-0.26
Adjusted R-squared	0.553	0.528	0.537	0.558	0.559	0.584	0.632	0.657	0.684
Observations	4460	8081	10825	2438	4159	5591	1538	2688	3621
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population Density	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Pixel Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls at the Ethnic-Country Level	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Estimates of this Table reproduce M&P estimates of Table VIII pp.145-147, with the only exception that each pixel is given its population weights. Dependent variable is the Lit/Unlit dummy. OLS estimates. Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. Population Weights: Total population circa 2000, from the data used in M&P.

B Construction of an African Dataset from DHS Country Surveys

In order to study the correlation between pre-colonial centralization and contemporary development outcomes, we merged geocoded DHS data from different countries. For each country, we chose the most recent geocoded survey among phases 3 to 6 (there were no geocoded surveys in phases 1 and 2)[23].

Although most surveys were conducted between 2005 and 2012, there is a span of 18 years between the oldest (1994) and most recent (2012) survey. The country fixed effects used in almost all regressions will absorb the effect of time, as households of the same country were surveyed in the same year. As far as education variables are concerned, we limited the sample to people born between 1959 and 1979. These people were at least 15 years old in 1994 and were therefore too old to still be in primary school; they were at most 53 years old in 2012, so that selection by mortality is not too much of a concern.

When conducting statistical analysis using DHS surveys, it is necessary to use sample weights to take into account the under- and over-sampling of some areas. Since we merged different DHS surveys, we needed to transform the weights to take into account the size of the different countries in our dataset (the number of households surveyed is roughly the same for Swaziland and Nigeria). We multiplied

[23]The sample comprises 33 countries. From phase 6: Angola, Burkina-Faso, Burundi, Cameroon, Côte d'Ivoire, Ethiopia, Gabon, Lesotho, Liberia, Madagascar, Malawi, Mozambique, Nigeria, Rwanda, Senegal, Tanzania, Uganda and Zimbabwe; from phase 5: Democratic Republic of the Congo, Egypt, Ghana, Guinea, Kenya, Mali, Namibia, Sierra Leone, Swaziland and Zambia; from phase 4: Benin and Morocco; from phase 3: Central African Republic, Niger and Togo. When considering education variables, Liberia and Madagascar are taken from phase 5 (as the education variables of the household members are missing from phase 6 surveys) and Angola is missing (as there is no geocoded survey with household member education variables=.

each weight by the ratio of the country population in 2000 over the number of cases in the country's DHS survey, before normalizing the weights to have the total sum of weights be equal to the number of cases in the Africa-wide dataset (abstaining from this weight transformation does not affect our findings)[24].

The DHS wealth index is country specific and therefore cannot be used for international comparisons. We therefore built a new wealth index, following the methodology used by Smits and Steendijk (2012). We included in the index only those assets for which the relevant variable was present in all of our 33 surveys. Asset weights were computed using principal component analysis (PCA). Table A6 presents the list of assets and the weights obtained by PCA. PCA was undertaken weighting each household by the transformed survey weights (following Smits and Steendijk and weighting each country by the square root of population size in 1985 instead of population size hardly changes the results). The wealth index is normalized to have zero mean and unit variance.

Each DHS enumeration area was allocated to its ethnic group on the Murdock map. We also matched each enumeration area with the closest pixel in M&P's dataset. We are therefore able to control at the country-ethnicity level and at the local level using the exact same variable they used. Geographical coordinates in DHS surveys are randomly displaced up to a 5 km maximum of positional error for rural clusters (10 km for 1 % of the rural clusters). This displacement will create some measurement errors, notably in the allocation of enumeration areas to ethnic groups and the centralization variables, biasing our coefficient towards zero. However, excluding the villages situated less than 5 or 10 km from an ethnic

[24]Since for education variables we consider only individuals born between 1959 and 1979, we use the population aged 20 to 39 in 2000. Source: United Nations Population Information Network, `http://www.un.org/popin/data.html`.

Table A6: Asset Indicators used in the Construction of the Wealth Index

	mean	std. deviation	raw indicator weight	normalized indicator weight
Consumer durables				
Television	0.36	0.48	0.3584	0.3687
Radio	0.61	0.49	0.1499	0.1542
Refrigerator	0.22	0.42	0.3525	0.3626
Car	0.06	0.23	0.1446	0.1488
Motorcycle	0.10	0.30	0.0124	0.0128
Bicycle	0.23	0.42	-0.0584	-0.0601
Housing characteristics				
Floor material:				
low quality	0.10	0.50	-0.3077	-0.3165
medium quality	0.48	0.46	0.0673	0.0692
high quality	0.22	0.41	0.2973	0.3058
Toilet facility:				
low quality	0.26	0.44	-0.1682	-0.1730
medium quality	0.49	0.50	-0.1644	-0.1692
high quality	0.25	0.43	0.3588	0.3691
Public utilities				
Access to electricity	0.41	0.49	0.3586	0.3689
Water source:				
low quality	0.52	0.50	-0.2831	-0.2912
medium quality	0.22	.42	-0.0222	-0.0229
high quality	0.26	.44	0.3456	0.3555

border hardly changes our results.

C Additional Results Using Survey Data

Table A7: DHS Dataset: Probit Model, Average Marginal Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: village connection to electricity							
Jur. Hiererarchy	0.1302*	0.0150	0.0173	0.0110				
Double-clustered p-val.	(0.0645)	(0.4759)	(0.3473)	(0.5797)				
0-1 pol. centralization					0.2250	0.0232	0.0471*	0.0462
Double-clustered p-val.					(0.1687)	(0.4203)	(0.0679)	(0.1177)
Observations	10,291	9,174	9,067	9,067	10,291	9,174	9,067	9,067
	Dependent variable: household connection to electricity							
Jur. Hiererarchy	0.0935	0.0050	0.0067	0.0005				
Double-clustered p-val.	(0.1939)	(0.7692)	(0.6118)	(0.9751)				
0-1 pol. centralization					0.1454	0.0068	0.0253	0.0213
Double-clustered p-val.					(0.2927)	(0.7674)	(0.1404)	(0.3117)
Observations	244,808	244,808	241,874	241,874	244,808	244,808	241,874	241,874
	Dependent variable: individual ever went to school							
Jur. Hiererarchy	-0.0194	-0.0026	-0.0054	-0.0092				
Double-clustered p-val.	(0.5212)	(0.8832)	(0.7494)	(0.4097)				
0-1 pol. centralization					-0.0609	-0.0267	-0.0165	-0.0133
Double-clustered p-val.					(0.3653)	(0.0395)	(0.5796)	(0.4943)
Observations	236,220	236,220	232,709	232,709	236,220	236,220	232,709	232,709
	Dependent variable: individual completed primary							
Jur. Hiererarchy	-0.0183	0.0019	0.0054	-0.0013				
Double-clustered p-val.	(0.5676)	(0.8984)	(0.6485)	(0.9023)				
0-1 pol. centralization					-0.0466	-0.0169	0.0051	0.0056
Double-clustered p-val.					(0.4869)	(0.6609)	(0.8213)	(0.6875)
Observations	236,220	236,220	232,709	232,709	236,220	236,220	232,709	232,709
Country F.E.	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Pixel controls	No	No	Yes	Yes	No	No	Yes	Yes
Country-ethnicity controls	No	No	No	Yes	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis on merged DHS surveys from 33 countries (only 32 for education variables). For each country, most recent geocoded survey among phases 3 to 6 (no geocoded surveys in phases 1 and 2). Each cluster was allocated to its country-ethnicity area and to the nearest M&P pixel.
Notes: Village connection to electricity equal to one if more than 50% of households in the cluster have electricity, zero otherwise. Asset-based wealth index built by the authors using principal component analysis (DHS wealth index not suitable for cross-country comparisons).
Probit estimates. Double-clustered estimates are obtained by computing three estimates, the first one clustered at the country level, the second at the ethno-linguistic family level, the third at the intersection of both. We then compute $Var(\hat{\beta}) = Var_1(\hat{\beta}) + Var_2(\hat{\beta}) - Var_{1\cap 2}(\hat{\beta})$, as in Cameron, Gelbach and Miller (2011). When non-positive semi-definite, the matrix is transformed by replacing all negative eigenvalues with zero (this transformation is actually innocuous for the coefficient of interest). Clustered p-values are computed as a sandwich estimator on the scores, assuming no bias on pseudo-likelihood point estimates. Standard errors and p-values of the average marginal effects computed using the delta method.
DHS sampling weights used in all regressions, transformed to take into account differences in population sizes across countries. For schooling variables, gender of the individual added as a control in all regressions.
In Egypt, all villages are connected to electricity; therefore, in all specifications with country fixed effects, the Egyptian fixed effect perfectly predicts connection to electricity. To increase numerical stability of the optimization process, Egyptian villages were dropped in panel 1, columns (2)-(4) and (6)-(8); the number of observations is thus lower than in Table 7, panel 1.

Table A8: 2005 Afrobarometer: Replication of M&P Results Weighting by Population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Variables in logs, no weighting							
	living conditions index				education variable			
0-1 pol. centralization	0.0740***	0.0743***	0.0753**	0.0656**	0.1275**	0.1230*	0.1469**	0.1064*
Double-clustered p-val.	(0.0057)	(0.0065)	(0.0255)	(0.0191)	(0.0446)	(0.0750)	(0.0215)	(0.0501)
	Variables in level, no weighting							
	living conditions index				education variable			
0-1 pol. centralization	0.1627***	0.1666***	0.1662**	0.1630**	0.2406	0.2110	0.2723*	0.2046*
Double-clustered p-val.	(0.0047)	(0.0054)	(0.0234)	(0.0141)	(0.1068)	(0.1574)	(0.0543)	(0.0709)
	Variables in level, weighting by population size							
	living conditions index				education variable			
0-1 pol. centralization	0.0889**	0.0887*	0.0702	0.0782*	-0.1516	-0.1241	0.0475	0.0868
Double-clustered p-val.	(0.0467)	(0.0672)	(0.1365)	(0.0624)	(0.2915)	(0.2926)	(0.5414)	(0.2429)
Observations	194	194	194	194	194	194	194	194
Country F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log population density	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Location controls	No	No	Yes	Yes	No	No	Yes	Yes
Geographic controls	No	No	No	Yes	No	No	No	Yes

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis on 2005 Afrobarometer Surveys matched to M&P country-ethnicity and pixel datasets.
Notes: The first panel tries to replicate M&P (2012) Appendix Table 5 (coefficients are slightly different because we had to rebuild the dataset). The living condition index is a categorical variable indicating how respondent view their present living conditions, from "very bad" (1) to "very good" (5). The education variable used by M&P is not interpretable as years of schooling. It is a categorical variable taking 10 values from 0 (no education) to 9 (post-graduate). The number of observations for the schooling regression is 193 instead of 194, because we work in logs and one of the country-ethnicity area has mean education of zero). Because there is no specific reason for considering variables in logs (the distribution of the living condition and schooling variables is roughly normal) we present the same specification in levels in panel 2. In panel 3, we weight each country-ethnicity area by its population size in 2000 (computed as population density in 2000 times surface area).

D Additional Results on Density

Table A9: Pre-Colonial Institutions and Changes in Regional Population Density 1960-2000

	Jurisdictional Hierarchy (1-4)				Political Centralization (0-1)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Institutions	0.0124	0.0165	0.0198	0.0245	0.0381	0.0534	0.0374	0.0570
	(0.0249)	(0.0290)	(0.0307)	(0.0344)	(0.0577)	(0.0607)	(0.0653)	(0.0667)
Log Population Density 1960	1.0169***	0.9705***	0.9945***	0.9609***	1.0168***	0.9696***	0.9949***	0.9607***
	(0.0256)	(0.0356)	(0.0383)	(0.0409)	(0.0256)	(0.0353)	(0.0381)	(0.0404)
Adjusted R-squared	0.932	0.937	0.946	0.948	0.932	0.937	0.946	0.948
Observations	683	683	682	682	683	683	682	682
Country Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
Location controls	No	Yes	No	Yes	No	Yes	No	Yes
Geographic controls	No	Yes	No	Yes	No	Yes	No	Yes
Population weights	No	No	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Dependent variable is ln(population density in 2000 + 0.01), as in M&P Table A.IVA, p.9 of the Supplemental Appendix. OLS estimates. Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. In comparison with the estimates of Table A.IVA, log population density in 1960 (with the same +0.01 transformation) is introduced as an additional control.

Table A10: Pre-Colonial Institutions and Changes in Regional Population Density 1960-2000: Pixel-Level Analysis

	Jurisdictional Hierarchy (1-4)				Political Centralization (0-1)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ethnic Institutions	0.0514**	0.0562	0.0529	0.0334	0.0657	0.0551	0.0745	0.0480
	(0.0254)	(0.0386)	(0.0336)	(0.0292)	(0.0490)	(0.0692)	(0.0634)	(0.0548)
Log Population Density 1960	0.9283***	0.8943***	0.8749***	0.8732***	0.9315***	0.8967***	0.8759***	0.8738***
	(0.0114)	(0.0121)	(0.0129)	(0.0124)	(0.0118)	(0.0125)	(0.0128)	(0.0126)
Adjusted R-squared	0.809	0.825	0.830	0.833	0.808	0.825	0.830	0.833
Observations	59618	59618	59277	59277	59618	59618	59277	59277
Country Fixed Effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Controls Pixel Level	No	No	Yes	Yes	No	No	Yes	Yes
Controls Ethnic-Country Level	No	No	No	Yes	No	No	No	Yes
Population weights	No	No	No	No	No	No	No	No

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Source: Authors' analysis based on M&P data.
Notes: Dependent variable is Log Population Density in 2000, as in M&P Table A.IVB, p.10 of the Supplemental Appendix. OLS estimates. Within parentheses, double-clustered standard errors at the country and the ethno-linguistic family level. In comparison with the estimates presented in Table A.IVB, log population density in 1960 is introduced as an additional control; as in M&P, only pixels with non-zero population density (both in 1960 and 2000 here) are kept.